

Effective thermodynamics of two interacting underdamped Brownian particles

Tim Herpich^{✉,*}, Kamran Shayanfard, and Massimiliano Esposito^{✉,†}

*Complex Systems and Statistical Mechanics, Physics and Materials Science Research Unit, University of Luxembourg,
L-1511 Luxembourg, Luxembourg*



(Received 30 October 2019; accepted 25 January 2020; published 12 February 2020)

Starting from the stochastic thermodynamics description of two coupled underdamped Brownian particles, we showcase and compare three different coarse-graining schemes leading to an effective thermodynamic description for the first of the two particles: marginalization over one particle, bipartite structure with information flows, and the Hamiltonian of mean force formalism. In the limit of time-scale separation where the second particle with a fast relaxation time scale locally equilibrates with respect to the coordinates of the first slowly relaxing particle, the effective thermodynamics resulting from the first and third approach are shown to capture the full thermodynamics and to coincide with each other. In the bipartite approach, the slow part does not, in general, allow for an exact thermodynamic description as the entropic exchange between the particles is ignored. Physically, the second particle effectively becomes part of the heat reservoir. In the limit where the second particle becomes heavy and thus deterministic, the effective thermodynamics of the first two coarse-graining methods coincide with the full one. The Hamiltonian of mean force formalism, however, is shown to be incompatible with that limit. Physically, the second particle becomes a work source. These theoretical results are illustrated using an exactly solvable harmonic model.

DOI: [10.1103/PhysRevE.101.022116](https://doi.org/10.1103/PhysRevE.101.022116)

I. INTRODUCTION

Over the past two decades, stochastic thermodynamics established the tools to formulate thermodynamics for small systems subjected to significant fluctuations and driven far from equilibrium [1–7]. This theory has been successful in various contexts, e.g., Brownian particles [8,9], electronic systems [10], chemical reaction networks [11,12], active matter [13,14], and information processing [15]. In a nutshell, stochastic thermodynamics consistently builds a thermodynamic structure on top of a stochastic process described by master equations [16] or Fokker-Planck equations [17], implicitly assuming that the traced out degrees of freedom always stay at equilibrium.

In this paper we want to address two apparently distinct questions within the framework of underdamped Fokker-Planck dynamics. First, we want to shed light on the nature of heat and work by understanding how a subset of degrees of freedom from the system can start to behave as a thermal bath or a work source, respectively. For systems characterized by master equations, it was proven that if there is a time-scale separation (TSS) between the slow and the fast degrees of freedom the latter equilibrate with respect to the slow coordinates and represent an ideal heat reservoir the slow degrees of freedom are coupled with [18]. Instead, the conditions under which a subset of degrees of freedom can generate a stochastic driving on the system energies that can be treated as a work source have been identified in Ref. [19]. However, the limit of a smooth deterministic driving of the energies requires a

limit that only an underdamped Fokker-Planck equation can provide. We aim therefore at reconsidering these questions in this paper within the framework of underdamped Brownian dynamics. Secondly, we want to consider various coarse-graining schemes preserving thermodynamic consistency that have been proposed in the literature [20–36]. In particular, we want to focus on three different well-established approaches that have been considered for stochastic dynamics governed by master equations. First, we consider the most straightforward approach where a subset of states is explicitly coarse-grained and the effective thermodynamics is defined for that reduced dynamics as one formally would for the full dynamics [18,37]. Next, we consider an approach based on splitting the full system in two parts resulting in effective second laws for each part which are modified by a term describing the transfer of mutual information between each part. This approach provides a convenient framework to describe how a Maxwell demon [38] mechanism can produce an information flow that is consumed by the system to drive processes against their spontaneous direction [39–42]. Finally, we consider the so-called Hamiltonian of mean force (HMF) approach which introduces a notion of energy for a system strongly coupled to its environment [43–45]. In this paper, we will consider these various coarse-grainings for underdamped Brownian particles and discuss how they are related. As we will see, far from being distinct, the question of the connection between the different coarse-graining schemes will provide us with a good framework to get insight into the nature of heat and work.

To achieve these goals, we will consider two coupled underdamped particles as this model already contains the key ingredients to generalize to multiple underdamped particles. Besides interesting formal connections between entropic contributions appearing in the different coarse-graining schemes,

*tim.herpich@uni.lu

†massimiliano.esposito@uni.lu

we will find that the effective thermodynamics based on marginalization and the Hamiltonian of mean force become equivalent and capture the correct global thermodynamics in the limit of time-scale separation. In this limit, the second particle is so much faster than the first one that it instantaneously relaxes to a local equilibrium corresponding to the coordinates of the first particle. Conversely, the thermodynamics based on the slow part of the bipartite structure does not agree with the full thermodynamics. The mismatch corresponds to the entropic contribution due to the coupling of the second particle. Physically, the coarse-grained particle becomes part of the heat reservoir. Moreover, in the limit where one particle has an exceedingly large mass compared to the other one, we will find that the former becomes a work source acting on the latter. In that case, the effective thermodynamics emerging from the first two coarse-graining schemes, marginalization and bipartite structure, again captures the correct global thermodynamics (at least up to a trivial macroscopic friction term in the work source). In contrast, we will show that the Hamiltonian is incompatible with that limit. These theoretical predictions will be confirmed using an analytically tractable model made up of two linearly coupled harmonic oscillators.

The plan of the paper is as follows. In Sec. II A the stochastic thermodynamics for both a single underdamped particle and two interacting and underdamped particles is formulated. Next, in Sec. III we formulate and compare the three different coarse-graining approaches—marginalization, bipartite perspective, and Hamiltonian of mean force—for our underdamped two-particle system. The respective effective thermodynamic description is furthermore compared with the full one in the two aforementioned limits. As an example, we consider an analytically solvable model in Sec. IV. We conclude with an outlook to potential future works in Sec. V.

II. STOCHASTIC THERMODYNAMICS

A. Single underdamped particle

We consider a particle of mass m with the phase-space coordinate $\mathbf{\Gamma} = (\mathbf{x}, \mathbf{v})^\top \in \mathbb{R}^6$, where $\mathbf{x} \in \mathbb{R}^3$ and $\mathbf{v} \in \mathbb{R}^3$ denote position and velocity of the particle, respectively. The particle moves in a time-dependent potential $V(\mathbf{x}, t)$, hence its Hamiltonian reads

$$e(\mathbf{\Gamma}, t) = \frac{m}{2} \mathbf{v}^2 + V(\mathbf{x}, t). \quad (1)$$

The particle is furthermore subjected to a generic force $\mathbf{g}(\mathbf{\Gamma}, t)$. If the force is conservative it derives from a potential, $\mathbf{g}(\mathbf{x}, t) = -\partial_{\mathbf{x}} \hat{V}(\mathbf{x}, t)$. In this case, one can define heat and work in different ways depending on whether $\hat{V}(\mathbf{x}, t)$ is part of the system Hamiltonian (1) or not (see Ref. [46]). Conversely, if the force \mathbf{g} is nonconservative, it does not derive from a potential. For generality, and since it will be useful later, we assume that the force may be velocity dependent, $\mathbf{g}(\mathbf{\Gamma}, t)$.

The system is coupled to a heat reservoir at inverse temperature β , giving rise to zero-mean delta-correlated Gaussian white noise:

$$\langle \boldsymbol{\eta}_i(t) \rangle = 0, \quad \langle \boldsymbol{\eta}_i(t) \boldsymbol{\eta}_j(t') \rangle = 2\xi \beta^{-1} \delta_{ij} \delta(t - t'), \quad (2)$$

for $i, j = 1, 2, 3$. We denote by ξ the friction the particle experiences and set $k_B \equiv 1$ in the following. Then, the stochastic

dynamics of the system is governed by the following Langevin equation:

$$\begin{pmatrix} \dot{\mathbf{x}} \\ \dot{\mathbf{v}} \end{pmatrix} = \begin{pmatrix} \mathbf{v} \\ \frac{1}{m} [-\partial_{\mathbf{x}} V(\mathbf{x}, t) + \mathbf{g}(\mathbf{\Gamma}, t) - \xi \mathbf{v} + \boldsymbol{\eta}(t)] \end{pmatrix}, \quad (3)$$

and the equivalent Fokker-Planck equation ruling the time evolution of the probability density $P(\mathbf{\Gamma}, t)$ reads

$$\partial_t P = -\nabla \cdot (\boldsymbol{\mu} P) + \nabla \cdot (\mathbf{D} \cdot \nabla P), \quad (4)$$

with the drift and diffusion matrices

$$\boldsymbol{\mu} = \begin{pmatrix} \mathbf{v} \\ \frac{1}{m} [-\partial_{\mathbf{x}} V(\mathbf{x}, t) + \mathbf{g}(\mathbf{\Gamma}, t) - \xi \mathbf{v}] \end{pmatrix}, \quad (5)$$

$$D_{ij} = \frac{\xi \delta_{ij}}{\beta m^2} \sum_{n=4}^6 \delta_{in}, \quad (6)$$

and the nabla operator $\nabla \equiv (\partial_{\mathbf{x}}, \partial_{\mathbf{v}})^\top$. The Fokker-Planck Eq. (4) can be cast into a continuity equation:

$$\partial_t P = -\nabla \cdot \mathbf{J} = -\nabla \cdot (\mathbf{L}^{\text{det}} + \mathbf{L}^{\text{diss}}) P. \quad (7)$$

Here, the probability current \mathbf{J} is split into a deterministic contribution

$$\mathbf{L}^{\text{det}} = \begin{pmatrix} \mathbf{v} \\ \frac{1}{m} [-\partial_{\mathbf{x}} V(\mathbf{x}, t) + \mathbf{g}(\mathbf{\Gamma}, t)] \end{pmatrix} \quad (8)$$

and a dissipative one

$$\mathbf{L}^{\text{diss}} = -\frac{\xi}{m^2} \begin{pmatrix} 0 \\ m\mathbf{v} + \beta^{-1} \partial_{\mathbf{v}} \ln P \end{pmatrix}. \quad (9)$$

The average energy of the particle is

$$E = \int d\mathbf{\Gamma} e P, \quad (10)$$

and its rate of change

$$d_t E = \dot{Q} + \dot{W} \quad (11)$$

can be decomposed into a work current

$$\dot{W} = \int d\mathbf{\Gamma} P \partial_t e + \int d\mathbf{\Gamma} P \mathbf{g} \cdot \mathbf{v} \quad (12)$$

and into a heat current

$$\dot{Q} = \int d\mathbf{\Gamma} e \partial_t P - \int d\mathbf{\Gamma} P \mathbf{g} \cdot \mathbf{v}. \quad (13)$$

Equation (11) constitutes the first law of thermodynamics ensuring energy conservation [47]. Using Eq. (7), the heat current can be written as follows:

$$\dot{Q} = -\xi \int d\mathbf{\Gamma} P \left(\mathbf{v} + \frac{1}{\beta m} \partial_{\mathbf{v}} \ln P \right) \cdot \mathbf{v}. \quad (14)$$

The nonequilibrium system entropy associated with the particle at $\mathbf{\Gamma}$ is defined as [48]

$$s(\mathbf{\Gamma}) = -\ln P, \quad (15)$$

where the ensemble average coincides with the Shannon entropy:

$$S = - \int d\mathbf{\Gamma} P \ln P. \quad (16)$$

Its time derivative

$$d_t S = \int d\Gamma [\nabla \cdot \mathbf{L}^{\text{diss}}] P + \dot{I}_F = \beta \dot{Q} + \dot{\Sigma} + \dot{I}_F \quad (17)$$

can be split into the entropy flow from the bath to the system, $\beta \dot{Q}$, and the entropy production rate

$$\dot{\Sigma} = \beta \xi \int d\Gamma P \left(\mathbf{v} + \frac{1}{\beta m} \partial_{\mathbf{v}} \ln P \right)^2 \geq 0, \quad (18)$$

the non-negativity of which constitutes the second law of thermodynamics. Since it will be useful later, we introduced the notation

$$\dot{I}_F \equiv \frac{1}{m} \int d\Gamma P \partial_{\mathbf{v}} \cdot \mathbf{g}. \quad (19)$$

Defining the nonequilibrium free-energy density $f(\Gamma) = e(\Gamma) - \beta^{-1} s(\Gamma)$, one has for the average nonequilibrium free energy

$$F = \int d\Gamma P f = E - \beta^{-1} S. \quad (20)$$

Equation (20) allows us to rewrite the work and heat current in Eqs. (12) and (14) as

$$\begin{aligned} \dot{W} &= \int d\Gamma P \partial_t f + \int d\Gamma P \mathbf{g} \cdot \mathbf{v}, \\ \dot{Q} &= d_t(F + \beta^{-1} S) - \dot{W} \end{aligned} \quad (21)$$

and the entropy production rate in Eq. (18) as

$$\dot{\Sigma} = \beta(\dot{W} - d_t F) - \dot{I}_F \geq 0. \quad (22)$$

The additional term \dot{I}_F in Eqs. (17) and (22) illustrates that the presence of the velocity-dependent nonconservative force \mathbf{g} modifies the thermodynamics as noted in Refs. [49,50].

B. Special cases

1. Standard stochastic thermodynamics

Owing to the velocity dependence of $g(\Gamma, t)$, Eqs. (17) and (22) constitute a generalized entropy balance and a generalized second law, respectively. The standard thermodynamic formulation

$$d_t S = \beta \dot{Q} + \dot{\Sigma}, \quad T \dot{\Sigma} = \dot{W} - d_t F \geq 0 \quad (23)$$

is recovered for velocity-independent or nonconservative Lorentz forces, that is, forces that are orthogonal to the velocity, $\partial_{\mathbf{v}} \cdot \mathbf{g} = 0$. In one dimension, this is only true for velocity-independent forces $\partial_v g = 0$.

2. Deterministic limit

The dynamics is deterministic if $\xi = 0$, which physically corresponds to a decoupling of the particle from the thermal reservoir. According to Eq. (14), one has $\dot{Q} = 0$ and $d_t E = d_t W$. It follows furthermore from Eq. (18) that $\dot{\Sigma} = 0$, hence it holds, using Eq. (17), that

$$d_t S = \frac{1}{m} \int d\Gamma P \partial_{\mathbf{v}} \cdot \mathbf{g}. \quad (24)$$

Again, if \mathbf{g} is velocity independent or a Lorentz force, the deterministic dynamics becomes Hamiltonian and the rate of

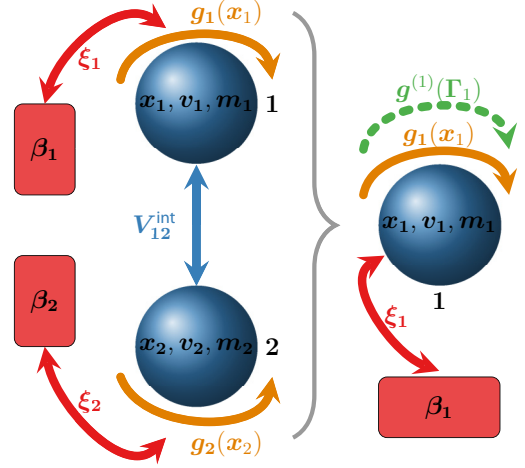


FIG. 1. On the left, schematics of the two underdamped and via V_{12}^{int} interacting particles 1 and 2 that are in contact with heat reservoirs at inverse temperatures β_1 and β_2 , respectively, are illustrated. It is furthermore assumed that both particles are subjected to nonconservative forces $g_i(x_i)$. The right depicts the coarse-grained description of solely the first particle in the presence of an additional nonconservative force $g^{(1)}(\Gamma_1)$ (green, dashed vector) that encodes the interaction with the second particle.

entropy change is identically zero, $d_t S = 0$. In this case the second law is a triviality.

3. Heavy particle

Finally, we consider the limit where the mass of the particle diverges, $m \rightarrow \infty$. We suppose that the conservative force scales with the mass, i.e., $O(\partial_{x_i} V/m) = 1 \forall i$, to avoid the trivial case of a particle in a flat potential. If ξ and \mathbf{g} are finite, so that $\xi/m \rightarrow 0$ and $\mathbf{g}/m \rightarrow 0$, one finds using Eqs. (14), (17), and (18) that

$$\dot{\Sigma} = -\beta \dot{Q} = \beta \xi v_i^2 \geq 0, \quad d_t S = 0, \quad (25)$$

where v_i is the solution of the deterministic equations

$$d_t x_i = v_i, \quad d_t v_i = -\frac{1}{m} \partial_{x_i} V(\mathbf{x}, t)|_{\mathbf{x}=x_i}. \quad (26)$$

According to Eq. (25), the heavy particle corresponds to the limit of macroscopic friction.

C. Two coupled underdamped particles

We now consider two particles labeled by $i = 1, 2$ of mass m_i with the phase-space coordinate $\Gamma_i = (x_i, v_i)^\top$, as depicted in Fig. 1. The particles move in a time-dependent potential

$$V(\mathbf{x}_1, \mathbf{x}_2, t) = V_1(\mathbf{x}_1, t) + V_2(\mathbf{x}_2, t) + V_{12}^{\text{int}}(\mathbf{x}_1, \mathbf{x}_2, t), \quad (27)$$

that contains the interaction potential $V_{12}^{\text{int}}(\mathbf{x}_1, \mathbf{x}_2, t)$, and the Hamiltonian therefore reads

$$\begin{aligned} e(\Gamma, t) &= \frac{m_1}{2} v_1^2 + \frac{m_2}{2} v_2^2 + V(\mathbf{x}_1, \mathbf{x}_2, t) \\ &= \sum_i e_i(\Gamma_i, t) + V_{12}^{\text{int}}(\mathbf{x}_1, \mathbf{x}_2, t), \end{aligned} \quad (28)$$

where we denote the bare Hamiltonian of each particle by $e_i(\mathbf{\Gamma}_i, t) = m_i \mathbf{v}_i^2/2 + V_i(\mathbf{x}_i, t)$ with $i = 1, 2$. Moreover, we assume that both particles are subjected to velocity-independent nonconservative forces $\mathbf{g}_i(\mathbf{x}_i, t)$ [51].

Each of the particles is connected to a heat reservoir at inverse temperature β_i giving rise to uncorrelated zero-mean Gaussian white noise:

$$\langle \eta_j^{(i)}(t) \rangle = 0, \quad \langle \eta_j^{(i)}(t) \eta_{j'}^{(i)}(t') \rangle = 2 \xi_i \beta_i^{-1} \delta_{j,j'} \delta(t - t'), \quad (29)$$

where ξ_i refers to the friction the particle i experiences. The stochastic dynamics of the two-body system is ruled by the following Langevin equation:

$$\begin{pmatrix} \dot{\mathbf{x}}_i \\ \dot{\mathbf{v}}_i \end{pmatrix} = \begin{pmatrix} \mathbf{v}_i \\ \frac{1}{m_i} [-\partial_{\mathbf{x}_i} V(\mathbf{x}_1, \mathbf{x}_2, t) + \mathbf{g}_i(\mathbf{x}_i, t) - \xi_i \mathbf{v}_i + \boldsymbol{\eta}^{(i)}(t)] \end{pmatrix}, \quad (30)$$

and the equivalent Fokker-Planck equation governing the time evolution of the probability density $P(\mathbf{\Gamma}, t)$ reads

$$\partial_t P = -\nabla \cdot \mathbf{J} = -\nabla \cdot (\mathbf{L}^{\text{det}} + \mathbf{L}^{\text{diss}})P, \quad (31)$$

with $\nabla = (\partial_{\mathbf{x}_1}, \partial_{\mathbf{v}_1}, \partial_{\mathbf{x}_2}, \partial_{\mathbf{v}_2})^\top$. The probability current \mathbf{J} can be split into a deterministic part

$$\mathbf{L}^{\text{det}} = \begin{pmatrix} \mathbf{v}_1 \\ \frac{1}{m_1} [-\partial_{\mathbf{x}_1} V(\mathbf{x}_1, \mathbf{x}_2, t) + \mathbf{g}_1(\mathbf{x}_1, t)] \\ \mathbf{v}_2 \\ \frac{1}{m_2} [-\partial_{\mathbf{x}_2} V(\mathbf{x}_1, \mathbf{x}_2, t) + \mathbf{g}_2(\mathbf{x}_2, t)] \end{pmatrix} \quad (32)$$

and a dissipative one

$$\mathbf{L}^{\text{diss}} = \begin{pmatrix} 0 \\ \frac{-\xi_1}{m_1^2} (m_1 \mathbf{v}_1 + \beta_1^{-1} \partial_{\mathbf{v}_1} \ln P) \\ 0 \\ \frac{-\xi_2}{m_2^2} (m_2 \mathbf{v}_2 + \beta_2^{-1} \partial_{\mathbf{v}_2} \ln P) \end{pmatrix}. \quad (33)$$

The average energy of the system is

$$E = \int d\mathbf{\Gamma} e P, \quad (34)$$

and the first law of thermodynamics reads

$$d_t E = \dot{Q} + \dot{W}, \quad (35)$$

with the heat and work current

$$\dot{Q} = \int d\mathbf{\Gamma} e \dot{P} - \int d\mathbf{\Gamma} P (\mathbf{g}_1 \cdot \mathbf{v}_1 + \mathbf{g}_2 \cdot \mathbf{v}_2), \quad (36)$$

$$\dot{W} = \int d\mathbf{\Gamma} \dot{e} P + \int d\mathbf{\Gamma} P (\mathbf{g}_1 \cdot \mathbf{v}_1 + \mathbf{g}_2 \cdot \mathbf{v}_2). \quad (37)$$

Using the Fokker-Planck Eq. (31), we can write the heat current in terms of additive contributions:

$$\dot{Q} = \sum_{i=1}^2 \dot{q}^{(i)}, \quad \dot{q}^{(i)} = -\xi_i \int d\mathbf{\Gamma} P \left(\mathbf{v}_i + \frac{1}{\beta_i m_i} \partial_{\mathbf{v}_i} \ln P \right) \mathbf{v}_i. \quad (38)$$

Like in the single-particle case (16), the nonequilibrium system entropy is defined as

$$S = - \int d\mathbf{\Gamma} P \ln P, \quad (39)$$

and the entropy balance is thus given by

$$d_t S = \sum_{i=1}^2 \beta_i \dot{q}^{(i)} + \dot{\Sigma}, \quad (40)$$

where the non-negative entropy production rate

$$\dot{\Sigma} = \sum_{i=1}^2 \dot{\sigma}^{(i)}, \quad \dot{\sigma}^{(i)} = \beta_i \xi_i \int d\mathbf{\Gamma} P \left(\mathbf{v}_i + \frac{1}{\beta_i m_i} \partial_{\mathbf{v}_i} \ln P \right)^2 \geq 0 \quad (41)$$

constitutes the second law of thermodynamics. In fact, Eq. (41) formulates a stronger statement: the additive contributions $\dot{\sigma}^{(i)}$ are separately non-negative. Yet, as will be shown further below, the additive contributions $\dot{\sigma}^{(i)}$ are, in general, not equal to the entropy production associated with the single particles.

III. COARSE GRAINING

A. Effective dynamics

We now shift our attention to the first particle alone. This formally amounts to integrating the Fokker-Planck Eq. (31) over the coordinates of the second particle $\mathbf{\Gamma}_2 = (\mathbf{x}_2, \mathbf{v}_2)$ such that we obtain the marginalized probability distribution of particle 1, $P_1 \equiv \int d\mathbf{\Gamma}_2 P$, that satisfies the following effective Fokker-Planck equation:

$$\partial_t P_1 = -\nabla_1 \cdot \mathbf{J}_1 = -\nabla_1 \cdot (\mathbf{L}_1^{\text{det}} + \mathbf{L}_1^{\text{diss}})P_1, \quad (42)$$

with $\nabla_1 = (\partial_{\mathbf{x}_1}, \partial_{\mathbf{v}_1})^\top$. The marginal probability current \mathbf{J}_1 can be split into a deterministic part

$$\mathbf{L}_1^{\text{det}} = \begin{pmatrix} \mathbf{v}_1 \\ \frac{1}{m_1} [-\partial_{\mathbf{x}_1} V_1(\mathbf{x}_1, t) + \mathbf{g}_1(\mathbf{x}_1, t) + \mathbf{g}^{(1)}(\mathbf{\Gamma}_1, t)] \end{pmatrix} \quad (43)$$

and a dissipative one

$$\mathbf{L}_1^{\text{diss}} = \begin{pmatrix} 0 \\ \frac{-\xi_1}{m_1^2} (m_1 \mathbf{v}_1 + \beta_1^{-1} \partial_{\mathbf{v}_1} \ln P_1) \end{pmatrix}. \quad (44)$$

By comparison with the exact single-particle Fokker-Planck Eq. (7), we note that the coarse-graining of the second particle encodes the interaction between the two particles in the effective and nonconservative force imposed on particle 1:

$$\mathbf{g}^{(1)}(\mathbf{\Gamma}_1, t) = - \int d\mathbf{\Gamma}_2 P_{2|1}(\mathbf{\Gamma}, t) \partial_{\mathbf{x}_1} V_{12}^{\text{int}}(\mathbf{x}_1, \mathbf{x}_2, t). \quad (45)$$

We note that the evolution Eq. (42) is not closed since $\mathbf{g}^{(1)}$ depends on $P_{2|1}$. Thus, solving the effective Fokker-Planck Eq. (42) is as difficult as treating the full-Fokker-Planck Eq. (31).

Moreover, for specific choices of the interaction potential, the first particle might be considered as an active Brownian particle. In this case, the velocity-dependent nonconservative force (45) is interpreted as an additional energy inflow leading to active motion. The latter is described effectively by negative dissipation in the direction of motion with velocity-dependent friction kernels. Some prominent models of active Brownian particles can, for instance, be found in Refs. [52–55].

B. Effective thermodynamics

1. Marginalization

In the following, we attempt to formulate a consistent thermodynamic description for this reduced dynamics. Naively, it is tempting to use as an educated guess the single-particle expressions in Sec. II A for the reduced dynamics. In this case, the naive entropy balance reads

$$d_t S_1 = \beta_1 \dot{q}^{(1)} + \dot{\Sigma}^{(1)} + \dot{I}_F^{(1)}, \quad (46)$$

where we use the notation from Eq. (19),

$$\dot{I}_F^{(1)} \equiv \frac{1}{m_1} \int d\Gamma_1 P_1 \partial_{v_1} \cdot \mathbf{g}^{(1)}, \quad (47)$$

and denote the single-particle Shannon entropy by

$$S_1 = - \int d\Gamma_1 P_1 \ln P_1, \quad (48)$$

which implies for the non-negative effective entropy production rate

$$\dot{\Sigma}^{(1)} = \beta_1 \xi_1 \int d\Gamma_1 P_1 \left(v_1 + \frac{1}{\beta_1 m_1} \partial_{v_1} \ln P_1 \right)^2 \geq 0. \quad (49)$$

For reasons that will become clear soon, we, however, define the effective entropy balance as follows:

$$d_t S = \beta_1 \dot{Q}^{(1)} + \dot{\Sigma}^{(1)} + \dot{I}_F^{(1)}, \quad (50)$$

where the effective heat

$$\dot{Q}^{(1)} = \dot{q}^{(1)} + \beta_1^{-1} S_{2|1} \quad (51)$$

is supplemented by the conditional Shannon entropy

$$S_{2|1} = S - S_1 = - \int d\Gamma_1 P_1 \int d\Gamma_2 P_{2|1} \ln P_{2|1}. \quad (52)$$

The difference between the full (38) and effective (51) heat current can be written as

$$\dot{Q} - \dot{q}^{(1)} = \dot{q}^{(2)} - \beta_1^{-1} S_{2|1}. \quad (53)$$

Moreover, the difference between the full (41) and the effective entropy production rate (49) is given by

$$\dot{\Sigma} - \dot{\Sigma}^{(1)} = \int d\Gamma_1 P_1 \dot{\Sigma}_1, \quad (54)$$

with the internal entropy production rate kernel

$$\dot{\Sigma}_1 = \dot{\Sigma}'_1 + \dot{\Sigma}''_1, \quad (55)$$

that can be split in the following two non-negative contributions:

$$\dot{\Sigma}'_1 = \beta_2 \xi_2 \int d\Gamma_2 P_{2|1} \left(v_2 + \frac{1}{\beta_2 m_2} \partial_{v_2} \ln P_{2|1} \right)^2 \geq 0, \quad (56)$$

$$\dot{\Sigma}''_1 = \frac{\xi_1}{\beta_1 m_1^2} \int d\Gamma_2 P_{2|1} (\partial_{v_1} \ln P_{2|1})^2 \geq 0. \quad (57)$$

The first contribution $\dot{\Sigma}'_1$ is the entropy production rate of the second particle if the coordinates of the first one are fixed [see Eq. (41)]. Conversely, the second contribution $\dot{\Sigma}''_1$ can be viewed as a contribution to the entropy production rate due to the correlation of the particles as we will see in Eq. (65).

An equivalent decomposition to Eq. (55) for Markovian master equations was found in Ref. [18]. From the last two

equations we deduce that the effective entropy production (rate) always underestimates the physical one:

$$\dot{\Sigma} \geq \dot{\Sigma}^{(1)}. \quad (58)$$

It is important to note that at this general level it is impossible to fully capture the full thermodynamics solely in terms of properties of the reduced dynamics. The missing contributions require knowledge about the conditional probability $P_{2|1}$.

2. Bipartite system

A second approach to formulate an effective thermodynamics is provided by a bipartite system where the two-particle system is split into two single-particle subsystems. The effective entropic expressions in both subsystems are defined in the same formal way as one would for a single particle. Subsequently, the sum of the effective entropy balances in both subsystems is compared with the full one of the two-particle system in order to identify the so-called information flows exchanged between the subsystems.

Physically, a bipartite system provides a simple and convenient representation of Maxwell's demon since the thermodynamic cost of the latter becomes fully accessible [39,41,42]. Mathematically, the bipartite structure identifies the nonadditive contributions of the full thermodynamic quantities for the two particles. We first note that the additive contributions to the two-particle heat current (38) can be rewritten in terms of marginalized probabilities only as follows:

$$\dot{q}^{(i)} = -\xi_i \int d\Gamma_i P_i \left(v_i + \frac{1}{\beta_i m_i} \partial_{v_i} \ln P_i \right) v_i, \quad (59)$$

where the marginal probability P_2 is obtained analogously as P_1 , that is, by marginalizing the two-point probability P over Γ_1 . Using the last equation along with Eqs. (13) and (38), we see that the following relation holds:

$$\dot{q}^{(i)} = \int d\Gamma_i e_i \dot{P}_i - \int d\Gamma_i P_i v_i \cdot (\mathbf{g}_i + \mathbf{g}^{(i)}), \quad (60)$$

with the nonconservative force $\mathbf{g}^{(2)}$:

$$\mathbf{g}^{(2)}(\Gamma_2, t) = - \int d\Gamma_1 P_{1|2}(\Gamma, t) \partial_{\mathbf{x}_2} V_{12}^{\text{int}}(\mathbf{x}_1, \mathbf{x}_2, t). \quad (61)$$

Conversely, the additive contributions $\dot{\sigma}^{(i)}$ to the entropy production rate in Eq. (73) cannot be represented by marginal distributions only. Therefore, the entropy-balance equations for the subsystems of the bipartite system cannot be expressed in terms of its associated degrees of freedom only. We proceed by deriving the nonadditive contribution to the entropy and identifying it as the information flow.

To this end, we first define the relative entropy (or Kulback-Leibler divergence) as a statistical measure of the distance between the distributions P and $P_1 P_2$ as follows:

$$I = D[P || P_1 P_2] = \int d\Gamma P \ln \frac{P}{P_1 P_2} \geq 0, \quad (62)$$

the non-negativity of which readily follows from the inequality $\ln P \leq P - 1$. From Eqs. (16) and (39) follows that the relative entropy is the nonadditive part of the two-particle system entropy, i.e.,

$$I = S_1 + S_2 - S. \quad (63)$$

Physically, this quantity corresponds to the mutual information that is a measure of correlations that quantifies how much one system knows about the other. If I is large, the two systems are highly correlated, whereas small values of I imply that the two systems know little about each other. The time derivative of the mutual information

$$d_t I = \dot{I}^{(2 \rightarrow 1)} + \dot{I}^{(1 \rightarrow 2)} \quad (64)$$

can be split into two directional information flows:

$$\dot{I}^{(2 \rightarrow 1)} = \int d\Gamma_1 P_1 \left(\frac{1}{m_1} \partial_{v_1} \cdot \mathbf{g}^{(1)} - \dot{\Sigma}_1'' \right), \quad (65)$$

$$\dot{I}^{(1 \rightarrow 2)} = \int d\Gamma_2 P_2 \left(\frac{1}{m_2} \partial_{v_2} \cdot \mathbf{g}^{(2)} - \dot{\Sigma}_2'' \right), \quad (66)$$

where we used Eqs. (45) and (57) in the first equation. In the second equation we used Eq. (61) and introduced the integral kernel specifying the difference between the full and the effective entropy production rate for the second particle:

$$\dot{\Sigma} - \dot{\Sigma}^{(2)} = \int d\Gamma_2 P_2 \dot{\Sigma}_2 = \int d\Gamma_2 P_2 (\dot{\Sigma}_2' + \dot{\Sigma}_2''), \quad (67)$$

with

$$\dot{\Sigma}_2' = \beta_1 \xi_1 \int d\Gamma_1 P_{1|2} \left(v_1 + \frac{1}{\beta_1 m_1} \partial_{v_1} \ln P_{1|2} \right)^2 \geq 0, \quad (68)$$

$$\dot{\Sigma}_2'' = \frac{\xi_2}{\beta_2 m_2^2} \int d\Gamma_1 P_{1|2} (\partial_{v_2} \ln P_{1|2})^2 \geq 0. \quad (69)$$

The directional information flows can be interpreted as follows: When $\dot{I}^{(i \rightarrow j)} > 0$, the dynamics of particle j increases the mutual information and thus the correlations between the two particles. In other words, j is learning about i and vice versa. Conversely, $\dot{I}^{(i \rightarrow j)} < 0$ corresponds to decreasing correlations between the two particles due to the evolution of particle j , which can be interpreted as either information erasure or the conversion of information into energy [39]. We furthermore point out that a positive directional information flow indicates that its force contribution

$$\dot{I}_F^{(i \rightarrow j)} \equiv \frac{1}{m_j} \int d\Gamma_j P_j \partial_{v_j} \cdot \mathbf{g}^{(j)} \quad (70)$$

dominates its entropic part

$$\dot{I}_S^{(i \rightarrow j)} \equiv - \int d\Gamma_j P_j \dot{\Sigma}_j'', \quad (71)$$

since the latter is nonpositive according to Eq. (57). Various other interpretations of these mutual information flows have been discussed in the literature [56–61].

An inspection of Eq. (47) reveals that the force contribution of the information flow, $\dot{I}_F^{(i \rightarrow j)}$, is the additional term that enters in the effective entropy balance due to the velocity-dependent nonconservative force $\mathbf{g}^{(j)}$:

$$d_t S_j = \beta_j \dot{q}^{(j)} + \dot{\Sigma}^{(j)} + \dot{I}_F^{(i \rightarrow j)}. \quad (72)$$

Using Eq. (71), we furthermore find that the difference between the effective (49) and the additive contribution to the two-particle entropy production rate (41) corresponds to the entropic part of the information flow:

$$\dot{I}_S^{(i \rightarrow j)} = \dot{\Sigma}^{(j)} - \dot{\sigma}^{(j)}. \quad (73)$$

The last two equations stipulate the following effective entropy balance equation for particle j :

$$d_t S_j = \beta_j \dot{q}^{(j)} + \dot{\sigma}^{(j)} + \dot{I}^{(i \rightarrow j)}. \quad (74)$$

It is important to note that Eq. (74) states that the directional information flows are the nonadditive quantities entering in the effective entropy balance. We emphasize that Eq. (74) is the underdamped Fokker-Planck analog of the result found for master equations in Ref. [39]. Moreover, using Eqs. (56) and (73), it holds that

$$\int d\Gamma_i P_i \dot{\Sigma}_i' = \dot{\sigma}^{(j)} = \dot{\Sigma}^{(j)} - \dot{I}_S^{(i \rightarrow j)}, \quad (75)$$

which because of Eq. (41) implies that

$$\dot{\Sigma} = \dot{\Sigma}^{(1)} + \dot{\Sigma}^{(2)} - \dot{I}_S^{(2 \rightarrow 1)} - \dot{I}_S^{(1 \rightarrow 2)}. \quad (76)$$

An identical result for bipartite master equations was found in Ref. [62] and recently for the more general case of systems undergoing a quantum dynamics formulated in terms of a density matrix, where the generator is additive with respect to the reservoirs [63].

3. Hamiltonian of mean force

Finally, we present a third approach to define an effective thermodynamics for the reduced dynamics of particle 1 in Fig. 1, where we set $\beta_{1,2} = \beta$ and $\mathbf{g}_2 = 0$. For reasons that will become clear soon, we furthermore consider an explicitly time-independent bare Hamiltonian of the second particle $\partial_{t_2} = \partial_{t_2} V_2 = 0$. As we will see, for this approach only a specific class of initial conditions can be considered.

The key concept is the so-called Hamiltonian of mean force, originally utilized in equilibrium thermostatics [64], which defines an effective energy for particle 1 that accounts for the strong coupling [44] to the second particle 2. Using it, this approach attempts to overcome the problem identified in the context of Eq. (50) that there is *a priori* no systematic way to embed the global energetics into the reduced dynamics.

The marginal of the global (Gibbs) equilibrium distribution over the second particle can be expressed as

$$P_1^{\text{HMF}} = \int d\Gamma_2 P^{\text{eq}} = \int d\Gamma_2 e^{-\beta(e - F^{\text{eq}})} = e^{-\beta(H^{\text{HMF}} - F_{\text{HMF}}^{\text{eq}})}, \quad (77)$$

where we introduced the effective free energy $F_{\text{HMF}}^{\text{eq}}$ of particle 1 which is defined as the difference between the full equilibrium free energy

$$F^{\text{eq}} = -\frac{1}{\beta} \ln \int d\Gamma e^{-\beta e} \quad (78)$$

and that of the second particle

$$F_2^{\text{eq}} = -\frac{1}{\beta} \ln \int d\Gamma_2 e^{-\beta e_2}, \quad (79)$$

that is, $F_{\text{HMF}}^{\text{eq}} = F^{\text{eq}} - F_2^{\text{eq}}$. Consequently the HMF is defined as

$$H^{\text{HMF}} \equiv e_1 - \beta^{-1} \ln \langle e^{-\beta V_{12}^{\text{int}}} \rangle_2^{\text{eq}}. \quad (80)$$

We denote by $\langle \cdot \rangle_2^{\text{eq}}$ and $\langle \cdot \rangle^{\text{eq}}$ an ensemble average over the equilibrium distribution of particle 2, $P_2^{\text{eq}} = \exp[-\beta(e_2 -$

$F_2^{\text{eq}}]$, and over the global equilibrium distribution, respectively.

The conditional equilibrium distribution $P_{2|1}^{\text{eq}}$ is obtained by dividing the global (Gibbs) equilibrium distribution by the marginal one in Eq. (77):

$$P_{2|1}^{\text{eq}} = \frac{P^{\text{eq}}}{P_1^{\text{HMF}}} = e^{-\beta(e - F_{2|1}^{\text{eq}})}, \quad (81)$$

where the free-energy landscape of particle 1 for a conditionally equilibrated particle 2 is

$$F_{2|1}^{\text{eq}} = e_1 - \beta^{-1} \ln \langle e^{-\beta V_{12}^{\text{int}}} \rangle_2^{\text{eq}} + F_2^{\text{eq}} = H^{\text{HMF}} + F_2^{\text{eq}}. \quad (82)$$

It is noteworthy that $F_{2|1}^{\text{eq}}$ is parametrically time dependent, whereas F_2^{eq} has no time dependence due to the choice of a time-independent Hamiltonian e_2 . Equation (82) shows that up to F_2^{eq} the HMF is equal to the free energy that the locally equilibrated second particle generates for given coordinates of the first particle.

Furthermore, we note the standard equilibrium identities

$$F_{2|1}^{\text{eq}} = E_{2|1}^{\text{eq}} - \beta^{-1} S_{2|1}^{\text{eq}}, \quad (83)$$

$$E_{2|1}^{\text{eq}} = \partial_\beta (\beta F_{2|1}^{\text{eq}}) = \int d\Gamma_2 P_{2|1}^{\text{eq}} e, \quad (84)$$

$$S_{2|1}^{\text{eq}} = \beta^2 \partial_\beta F_{2|1}^{\text{eq}} = - \int d\Gamma_2 P_{2|1}^{\text{eq}} \ln P_{2|1}^{\text{eq}}, \quad (85)$$

which, using Eq. (80), can be rewritten as

$$E_{2|1}^{\text{eq}} = \partial_\beta [\beta (H^{\text{HMF}} + F_2^{\text{eq}})], \quad (86)$$

$$S_{2|1}^{\text{eq}} = \beta^2 \partial_\beta (H^{\text{HMF}} + F_2^{\text{eq}}). \quad (87)$$

Inspired by [43], we employ the HMF (80) and its derived quantities in Eqs. (86) and (87) and average them over arbitrary *nonequilibrium* probabilities for particle 1, i.e.,

$$E^{\text{HMF}}(t) = \langle \partial_\beta (\beta H^{\text{HMF}}) \rangle(t) \quad (88)$$

and

$$S^{\text{HMF}}(t) \equiv S_1(t) + \beta^2 \langle \partial_\beta H^{\text{HMF}} \rangle(t), \quad (89)$$

where $\langle \cdot \rangle(t)$ refers to an ensemble average over a generic *nonequilibrium* distribution $P(t)$. We note that the definition of the entropy (89) also includes the single-particle Shannon entropy of particle 1 in addition to the contribution that stems from the HMF. Choosing a definition of work that coincides with the global one (37),

$$W^{\text{HMF}}(t) \equiv \int_0^t dt' \left[\langle \dot{e} \rangle(t') + \left(\int d\Gamma_1 P_1 v_1 \cdot g_1 \right)(t') \right], \quad (90)$$

the first law of thermodynamics imposes the following definition for heat:

$$Q^{\text{HMF}}(t) = -W(t) + \langle \partial_\beta (\beta H^{\text{HMF}}) \rangle(t) - \langle \partial_\beta (\beta H^{\text{HMF}}) \rangle(0). \quad (91)$$

Defining the nonequilibrium free energy to be of the same form as in the standard equilibrium case (83),

$$F^{\text{HMF}}(t) = E^{\text{HMF}}(t) - \frac{S^{\text{HMF}}(t)}{\beta} = \langle H^{\text{HMF}} \rangle(t) - \frac{S_1(t)}{\beta}, \quad (92)$$

we can rewrite the entropy balance

$$\Delta S^{\text{HMF}}(t) = \beta Q^{\text{HMF}}(t) + \Sigma^{\text{HMF}}(t), \quad (93)$$

in the form of a second law of thermodynamics as follows:

$$\Sigma^{\text{HMF}}(t) = \beta [W(t) - \Delta F^{\text{HMF}}(t)] \geq 0. \quad (94)$$

In order to prove the non-negativity of this definition for the entropy production [43,45], an initial condition of the form

$$P(0) = P_1(0) P_{2|1}^{\text{eq}} = P_1(0) e^{-\beta(e - H^{\text{HMF}} - F_2^{\text{eq}})} \quad (95)$$

is required. Indeed, using Eqs. (20) and (94), we have

$$\Sigma^{\text{HMF}}(t) - \Sigma(t) = \beta [\Delta F - \Delta F^{\text{HMF}}(t)]. \quad (96)$$

Due to the special choice for the initial condition (95), Eqs. (86) and (87) are valid at $t = 0$ so that

$$F(0) - F^{\text{HMF}}(0) = F_2^{\text{eq}}. \quad (97)$$

At later times, Eqs. (86) and (87) are no longer valid and we need to resort to the definitions (89) and (88) to obtain

$$F(t) - F^{\text{HMF}}(t) = \langle e \rangle(t) - \langle H^{\text{HMF}} \rangle(t) + \beta^{-1} [S_1(t) - S(t)]. \quad (98)$$

Since the HMF can also be expressed as

$$\langle H^{\text{HMF}} \rangle(t) = \langle e \rangle(t) + \beta^{-1} \langle \ln P_{2|1}^{\text{eq}} \rangle - F_2^{\text{eq}}, \quad (99)$$

we have

$$F(t) - F^{\text{HMF}}(t) = F_2^{\text{eq}} + \beta^{-1} \left\langle \ln \frac{P(t)}{P_{2|1}^{\text{eq}} P_1(t)} \right\rangle, \quad (100)$$

and finally arrive at

$$\Sigma^{\text{HMF}}(t) - \Sigma(t) = D[P(t) || P_{2|1}^{\text{eq}} P_1(t)] \geq 0. \quad (101)$$

Thus, the entropy production based on the HMF always overestimates the global two-particle entropy production which, because of Eq. (41), proves the inequality in Eq. (94). Furthermore, with Eq. (58) we obtain the following hierarchies of inequalities:

$$\Sigma^{\text{HMF}}(t) \geq \Sigma(t) \geq \Sigma^{(1)}(t), \quad (102)$$

where the equality signs hold in the limit of time-scale separation, as will be shown further below. The last equation is the Fokker-Planck analog of the result found for master equations in Ref. [45]. This reference also identifies the conditions under which the *rate* of the entropy production (94) is non-negative.

C. Limiting cases

As already pointed out above, the effective Fokker-Planck Eq. (42) is, in general, not closed because of the dependence on the conditional probability $P_{2|1}$. With the results of the preceding section at hand, we now study the three different coarse-graining schemes for two limiting cases in which the effective Fokker-Planck equation becomes closed and thus analytically tractable.

1. Fast-dynamics limit: The heat reservoir

First, we assume a time-scale separation between the stochastic dynamics of the two particles where particle 2 evolves much faster than particle 1. Hence, for fixed coordinates of the first particle, the second generically relaxes

towards a nonequilibrium steady state and the stationary conditional probability $P_{2|1}^{\text{TSS}}$ can be determined by solving the fast dynamics for fixed Γ_1 . As a consequence, the effective Fokker-Planck Eq. (42) becomes closed and the effective thermodynamics follows from replacing $P_{2|1}$ by $P_{2|1}^{\text{TSS}}$ in all expressions in Sec. III B. However, this effective thermodynamics naturally does not match with the full one, as we would neglect hidden degrees of freedom that are out of equilibrium.

The latter equilibrate only if $\mathbf{g}_2 = 0$ and $\beta_{1,2} = \beta$, that is, when the second particle instantaneously equilibrates with respect to each value of the slow coordinates of particle 1. Then, the conditional probability is given at any time by the Gibbs distribution [18]:

$$P_{2|1}^{\text{TSS}}(\mathbf{x}_1, \Gamma_2) \equiv P_{2|1}^{\text{eq}}(\mathbf{x}_1, \Gamma_2) = e^{-\beta(e - F_{2|1}^{\text{eq}})}. \quad (103)$$

As a result, the effective force $\mathbf{g}^{(1)}$ in Eq. (45) becomes a velocity-independent force that derives from an effective potential so that

$$[-\partial_{\mathbf{x}_1} V_1 + \mathbf{g}^{(1)}]_{\text{TSS}} = -\partial_{\mathbf{x}_1} F_{2|1}^{\text{eq}}, \quad (104)$$

where the notation $Z|_{\text{TSS}}$ corresponds to the conditional probability $P_{2|1}$ in the expression Z being substituted by the equilibrium one in Eq. (103). Hence, in the limit of TSS and local equilibrium, the particle is subjected to the effective potential given by the free-energy landscape of the first particle, $F_{2|1}^{\text{eq}}$.

a. Marginalization. Substituting Eq. (103) into Eq. (52) and accounting for probability conservation, we get

$$\frac{d_t S_{2|1}|_{\text{TSS}}}{\beta} = \int d\Gamma \dot{P}_1(t) P_{2|1}^{\text{eq}}(e - F_{2|1}^{\text{eq}}) \quad (105)$$

$$= \int d\Gamma \dot{P}_1(t) P_{2|1}^{\text{eq}} e - \int d\Gamma_1 \dot{P}_1(t) F_{2|1}^{\text{eq}}. \quad (106)$$

With Eqs. (60) and (82), we note the relation

$$\dot{q}^{(1)}|_{\text{TSS}} = \int d\Gamma_1 \dot{P}_1(t) F_{2|1}^{\text{eq}}, \quad (107)$$

from which along with Eq. (51) it follows that

$$\dot{Q}^{(1)}|_{\text{TSS}} = \dot{q}^{(1)}|_{\text{TSS}} + \beta^{-1} d_t S_{2|1}|_{\text{TSS}} = \dot{Q}|_{\text{TSS}}, \quad (108)$$

hence clarifying why the effective heat (51) was defined to contain the conditional Shannon entropy.

We have therefore proven that in the limit of TSS the effective (51) and the global heat current (36) coincide and the first law of thermodynamics remains formally the same as in Eq. (35):

$$d_t E|_{\text{TSS}} = \dot{Q}^{(1)}|_{\text{TSS}} + \dot{W}|_{\text{TSS}} = \dot{Q}|_{\text{TSS}} + \dot{W}|_{\text{TSS}}. \quad (109)$$

Furthermore, in the limit of TSS, the time dependence of all quantities stems only from the dynamics of particle 1 and the parametric time dependence of the Hamiltonian. Equation (108) proves that the second law of thermodynamics formally also remains the same as in Eq. (41):

$$\dot{\Sigma}^{(1)}|_{\text{TSS}} = d_t S|_{\text{TSS}} - \beta \dot{Q}^{(1)}|_{\text{TSS}} \quad (110)$$

$$= d_t S|_{\text{TSS}} - \beta \dot{Q}|_{\text{TSS}} = \dot{\Sigma}|_{\text{TSS}} \geq 0. \quad (111)$$

Hence, in the limit of TSS, the full thermodynamics of the two particles can be described solely by the reduced dynamics

of a single particle that is subjected to the potential $F_{2|1}^{\text{eq}}$. Physically, the second particle can be viewed as being part of the heat reservoir the first particle is coupled to.

b. Bipartite system. Furthermore, substituting (103) into Eqs. (65), (70), and (71) gives a vanishing directional information flow from the fast to the slow particle:

$$j_F^{(2 \rightarrow 1)}|_{\text{TSS}} = j_S^{(2 \rightarrow 1)}|_{\text{TSS}} = j^{(2 \rightarrow 1)}|_{\text{TSS}} = 0. \quad (112)$$

This means that in the limit of TSS the information flow is completely asymmetric, $d_t I|_{\text{TSS}} = d_t I^{(1 \rightarrow 2)}|_{\text{TSS}}$. From the last equation follows that the additive and effective entropy production rate (49) agrees with the global one (41):

$$\dot{\sigma}^{(1)}|_{\text{TSS}} = \dot{\Sigma}^{(1)}|_{\text{TSS}} = \dot{\Sigma}|_{\text{TSS}}, \quad (113)$$

which in turn implies that $\dot{\sigma}^{(2)}|_{\text{TSS}} = 0$, though there is a mismatch between the effective entropy balance of the slow particle (74) and the full entropy balance (40) given by the conditional Shannon entropy:

$$d_t S_1|_{\text{TSS}} = \beta \dot{q}^{(1)}|_{\text{TSS}} + \dot{\sigma}^{(1)}|_{\text{TSS}} = d_t S|_{\text{TSS}} - d_t S_{2|1}|_{\text{TSS}}. \quad (114)$$

Moreover, the effective entropy balance of the second particle reads

$$d_t S_2|_{\text{TSS}} = \beta d_t q^{(2)}|_{\text{TSS}} + d_t I^{(1 \rightarrow 2)}|_{\text{TSS}}, \quad (115)$$

that can be rewritten as

$$d_t S_2|_{\text{TSS}} - d_t S_{2|1}|_{\text{TSS}} = d_t I^{(1 \rightarrow 2)}|_{\text{TSS}}. \quad (116)$$

Equation (116) stipulates that the information flow $d_t I|_{\text{TSS}} = d_t I^{(1 \rightarrow 2)}|_{\text{TSS}}$ from the slow to the fast particle does, in general, not vanish. This is physically plausible since the particles are still correlated. The information flow $d_t I^{(1 \rightarrow 2)}|_{\text{TSS}}$ reflects time-varying correlations between the two particles due to the change of the probability distribution of both out-of-equilibrium particles. Consequently, the information flow is zero for a global equilibrium state characterized by $P^{\text{eq}} = P_{2|1}^{\text{eq}} P_1^{\text{eq}}$.

c. Hamiltonian of mean force. We now turn to the HMF formalism in the limit of TSS and local equilibrium, $\beta_{1,2} = \beta$ and $\mathbf{g}_2 = 0$. Further, as done above in the introduction of the HMF formalism, we assume that the bare Hamiltonian of the second particle is time independent, $\partial_t e_2 = 0$. Because of Eq. (103), the requirement of an initial equilibrium conditional probability distribution (95) is fulfilled at all times t . Hence Eqs. (86) and (87) are valid at any time t and a comparison with Eqs. (88) and (89), respectively, shows that

$$E^{\text{HMF}}(t)|_{\text{TSS}} = E(t)|_{\text{TSS}} - \partial_\beta(\beta F_2^{\text{eq}}), \quad (117)$$

$$S^{\text{HMF}}(t)|_{\text{TSS}} = S(t)|_{\text{TSS}} - \beta^2 \partial_\beta F_2^{\text{eq}}. \quad (118)$$

This explains the choice of a time-independent Hamiltonian e_2 , since in this case F_2^{eq} has no time dependence. As a result, the HMF definitions of the corresponding *currents* coincide with the global ones:

$$d_t E^{\text{HMF}}(t)|_{\text{TSS}} = d_t E(t)|_{\text{TSS}}, \quad (119)$$

$$d_t S^{\text{HMF}}(t)|_{\text{TSS}} = d_t S(t)|_{\text{TSS}}. \quad (120)$$

Moreover, we conclude that an agreement of the definitions for the time-integrated quantities would be achieved in the limit of TSS, if the HMF was defined as $H^{\text{HMF}*} \equiv F_{2|1}^{\text{eq}}$ which corresponds to the definition $F_{\text{HMF}}^{\text{eq}*} \equiv F^{\text{eq}}$. In this case, the equivalence of definitions would still be true for a time-dependent Hamiltonian e_2 .

By construction, the definitions of work agree [cf. Eqs. (37) and (90)], thus it follows from Eq. (119) that the definitions of heat *current* also coincide:

$$\dot{Q}^{\text{HMF}}(t)|_{\text{TSS}} = d_t E^{\text{HMF}}(t)|_{\text{TSS}} - \dot{W}(t)|_{\text{TSS}} = \dot{Q}(t)|_{\text{TSS}}. \quad (121)$$

Since according to Eqs. (119) and (121) the entropy production rates are also identical,

$$\dot{\Sigma}^{\text{HMF}}(t)|_{\text{TSS}} = d_t S^{\text{HMF}}(t)|_{\text{TSS}} - \beta \dot{Q}^{\text{HMF}}(t)|_{\text{TSS}} = \dot{\Sigma}(t)|_{\text{TSS}}, \quad (122)$$

we find that at the differential level the Hamiltonian of mean-force formalism captures the full thermodynamics in the limit of TSS. Furthermore, we have proven that in the limit of TSS all definitions of the entropy production rate in Eqs. (41), (49), and (93) are equivalent, i.e.,

$$\dot{\Sigma}(t)|_{\text{TSS}} = \dot{\Sigma}^{(1)}(t)|_{\text{TSS}} = \dot{\sigma}^{(1)}(t)|_{\text{TSS}} = \dot{\Sigma}^{\text{HMF}}(t)|_{\text{TSS}}. \quad (123)$$

Together with Eq. (110), this proves the equality signs in Eq. (102) in the limit of TSS.

This constitutes our first main result: In the limit of TSS and local equilibrium, the effective thermodynamic descriptions resulting from marginalization and the HMF formalism fully capture the full thermodynamics. In contrast, the effective bipartite description does not match with the full thermodynamics since it neglects the correlations between the two particles.

2. Large-mass limit: The work source

We proceed by studying the limit of a diverging mass of the second particle, $m_2 \rightarrow \infty$, that has already been discussed in Sec. II B. In view of active Brownian motion, this limit is interesting since the heavy second particle could represent a passive cargo, while the light particle may be considered active. Again, in order to avoid any triviality we assume that the potentials scale with the mass m_2 as follows: $O(\partial_{x_2} V_2/m_2) = 1 \forall i$ while $\partial_{x_2} V^{\text{int}}/m_2 \rightarrow 0 \forall i$ as $m_2 \rightarrow \infty$. Because of the infinite mass of particle 2 its motion occurs deterministically such that we can neglect the influence of particle 1. Consequently, the marginal probabilities become statistically independent and the conditional distribution reads

$$P_{2|1}^{\text{det}}(\mathbf{\Gamma}_2, t) = P_2^{\text{det}}(\mathbf{\Gamma}_2, t) = \delta(\mathbf{x}_2 - \mathbf{x}_t) \delta(\mathbf{v}_2 - \mathbf{v}_t), \quad (124)$$

for all times t including the initial time $t = 0$. Here, \mathbf{x}_t and \mathbf{v}_t are the solutions of the deterministic equations of motion (26). As a result, the effective force (45) becomes conservative:

$$\mathbf{g}^{(1)}(\mathbf{x}_1, t)|_{\text{det}} = -\partial_{\mathbf{x}_1} V_{12}^{\text{int}}(\mathbf{x}_1, \mathbf{x}_2, t)|_{\mathbf{x}_2=\mathbf{x}_t}, \quad (125)$$

where the notation $Z|_{\text{TSS}}$ corresponds to the conditional probability $P_{2|1}$ in the expression Z being substituted by the delta-correlated one in Eq. (124). Thus, we are dealing with a closed

effective Fokker-Planck Eq. (42) for the light particle one that is externally driven by the deterministic motion of the heavy second particle.

a. Marginalization. Since the marginal probabilities are statistically independent, the conditional Shannon entropy (52) vanishes,

$$S_{2|1}|_{\text{det}} = 0, \quad (126)$$

such that the definition of the effective heat (51) reduces to the naive one (46), $\dot{Q}^{(1)}|_{\text{det}} = \dot{q}^{(1)}|_{\text{det}}$. Therefore, by inserting Eq. (124) into Eq. (36), we get

$$\dot{Q}|_{\text{det}} - \dot{q}^{(1)}|_{\text{det}} = \dot{q}^{(2)}|_{\text{det}} = -\xi_2 \mathbf{v}_t^2. \quad (127)$$

Thus, the first law of thermodynamics remains—up to a macroscopic frictional term related to the heavy particle—formally the same as in Eq. (35):

$$d_t E|_{\text{det}} = \dot{q}^{(1)}|_{\text{det}} + \dot{W}|_{\text{det}} - \xi_2 \mathbf{v}_t^2 = \dot{Q}|_{\text{det}} + \dot{W}|_{\text{det}}. \quad (128)$$

Here, the difference is that the time dependence of all quantities comes from the dynamical time dependence of particle 1 alone, the parametric time dependence of the Hamiltonian, and the deterministic trajectory of the second particle ($\mathbf{x}_t, \mathbf{v}_t$). Further, Eq. (126) implies that the definitions for the single-particle Shannon entropy (48) and the full system entropy agree (39):

$$d_t S_1|_{\text{det}} = d_t S|_{\text{det}}, \quad (129)$$

which, in turn, proves that the effective second law of thermodynamics (49)—up to a macroscopic frictional term of the heavy particle—formally also remains the same as in Eq. (41):

$$\dot{\Sigma}^{(1)}|_{\text{det}} = d_t S_1|_{\text{det}} - \beta_1 \dot{q}^{(1)}|_{\text{det}} = \dot{\Sigma}|_{\text{det}}, \quad (130)$$

where $\dot{\Sigma}|_{\text{det}} = \dot{\Sigma}|_{\text{det}} - \beta_2 \xi_2 \mathbf{v}_t^2$. The effective thermodynamic description for the two particles therefore reduces, up to a simple macroscopic term, to the standard one of a single particle that is subjected to an external driving. Consequently, the physical interpretation of this limit is that the second particle represents a work source that modulates the energy landscape of the first particle according to a protocol ($\mathbf{x}_t, \mathbf{v}_t$). If the deterministic particle is furthermore Hamiltonian, $\xi_2 = 0$, the work source is nondissipative and the effective description coincides with the full one.

b. Bipartite system. Owing to the statistical independence of the marginal distributions, the mutual information (62) and thus the information flow is identically zero:

$$I|_{\text{det}} = \dot{I}^{2 \rightarrow 1}|_{\text{det}} = \dot{I}^{1 \rightarrow 2}|_{\text{det}} = 0. \quad (131)$$

As a result, the effective entropy balance of the light particle coincides with the full one:

$$d_t S_1|_{\text{det}} = \beta_1 \dot{q}^{(1)}|_{\text{det}} + \dot{\sigma}^{(1)}|_{\text{det}} = \dot{Q}|_{\text{det}} + \dot{\Sigma}|_{\text{det}} = d_t S|_{\text{det}}, \quad (132)$$

while the corresponding effective entropy balance equation for the heavy particle takes the simple macroscopic form

$$\beta_2 \dot{q}^{(2)}|_{\text{det}} = -\dot{\sigma}^{(2)}|_{\text{det}} = -\beta_2 \xi_2 \mathbf{v}_t^2. \quad (133)$$

c. Hamiltonian of mean force. The large-mass limit represents a special case of systems away from TSS. Yet, the assumption of a conditional Gibbs state (95) is inconsistent with the independent single-particle distributions (124).

Therefore, the HMF formalism and the deterministic limit are incompatible.

We can therefore summarize our second main result: In the deterministic limit, the effective thermodynamics of the first two coarse-graining schemes—marginalization and bipartite structure—are, up to a simple macroscopic frictional term, equivalent to the full thermodynamics. In contrast, the HMF formalism is incompatible with the deterministic limit. In fact, the HMF thermodynamics only matches with the full one in the limit of TSS. This is not surprising since the HMF definitions [cf. Eqs. (80) and (95)] are motivated by equilibrium thermostatics. Notably, in the TSS limit there is a completely asymmetric information flow from the slow to the fast particle, while in the deterministic limit all information flows vanish.

IV. TWO LINEARLY COUPLED HARMONIC OSCILLATORS

A. Full solution

In this section, the results derived above are illustrated for an analytically solvable example. For this purpose, we consider an isothermal version of the setup in Fig. 1 in one dimension. Moreover, the Hamiltonian (28) is assumed time independent,

$$V(x_1, x_2) = (k_1 x_1^2)/2 + (k_2 x_2^2)/2 + \beta(x_1 x_2), \quad (134)$$

and the nonconservative forces \mathbf{g}_i are taken as zero. Consequently, there is no work done on or by the two-particle system, $d_t E = d_t Q$. The Fokker-Planck Eq. (31) reads

$$\partial_t P = -\nabla \cdot (\boldsymbol{\gamma} \cdot \boldsymbol{\Gamma} P) + \nabla^\top \cdot (\mathbf{D} \cdot \nabla P), \quad (135)$$

with $\boldsymbol{\Gamma} = (x_1, v_1, x_2, v_2)^\top$ and $\nabla \equiv (\partial_{x_1}, \partial_{v_1}, \partial_{x_2}, \partial_{v_2})^\top$. The constant drift coefficient and diffusion matrix read, respectively,

$$\boldsymbol{\gamma} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -\frac{k_1}{m_1} & -\frac{\xi_1}{m_1} & -\frac{\beta}{m_1} & 0 \\ 0 & 0 & 0 & 1 \\ -\frac{\beta}{m_2} & 0 & -\frac{k_2}{m_2} & -\frac{\xi_2}{m_2} \end{pmatrix}, \quad (136)$$

$$\mathbf{D} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{\xi_1}{\beta m_1^2} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{\xi_2}{\beta m_2^2} \end{pmatrix}. \quad (137)$$

This partial differential equation is supplemented by the initial condition $P(0) = \delta[\boldsymbol{\Gamma}(t) - \boldsymbol{\Gamma}(0)]$. The solution of this Fokker-

Planck equation is given by a Gaussian [65]:

$$P = \frac{1}{(2\pi)^2 \sqrt{\det \boldsymbol{\Upsilon}}} \exp \left[-\frac{1}{2} (\boldsymbol{\Gamma} - \langle \boldsymbol{\Gamma} \rangle)^\top \cdot \boldsymbol{\Upsilon}^{-1} \cdot (\boldsymbol{\Gamma} - \langle \boldsymbol{\Gamma} \rangle) \right], \quad (138)$$

where the average values of the coordinates are determined as follows:

$$\langle \boldsymbol{\Gamma} \rangle(t) = e^{\mathcal{M}t} \cdot \boldsymbol{\Gamma}(0), \quad (139)$$

and the covariance matrix is calculated as

$$\boldsymbol{\Upsilon}_{kl}(t) \equiv 2 \sum_{i,j} \frac{1 - e^{-(\lambda_i + \lambda_j)t}}{\lambda_i + \lambda_j} C_{ij} u_i^{(k)} u_j^{(l)}. \quad (140)$$

Here, we introduced the transformation matrix

$$\mathbf{C} = \mathbf{V} \cdot \mathbf{D} \cdot \mathbf{V}^\top, \quad \mathbf{V} = (\mathbf{v}^{(1)}, \mathbf{v}^{(2)}, \mathbf{v}^{(3)}, \mathbf{v}^{(4)}), \quad (141)$$

where λ_i and $\mathbf{u}^{(i)}$ ($\mathbf{v}^{(i)}$) denote the i th eigenvalue and right (left) eigenvector of the drift coefficient matrix $\boldsymbol{\gamma}$, respectively, i.e.,

$$\begin{aligned} \boldsymbol{\gamma} \cdot \mathbf{u}^{(i)} &= \lambda_i \mathbf{u}^{(i)}, \\ \mathbf{v}^{(i)} \cdot \boldsymbol{\gamma} &= \lambda_i \mathbf{v}^{(i)}, \end{aligned} \quad (142)$$

such that the left and right eigenvectors of $\boldsymbol{\gamma}$ constitute an orthonormal dual basis, $\mathbf{v}^{(i)} \cdot \mathbf{u}^{(j)} = \delta_{ij}$. Substituting Eq. (138) into Eqs. (38) and (41), we obtain for the heat current and the entropy production rate

$$\dot{Q} = \sum_{i=1}^2 \left[-\xi_i (\boldsymbol{\Upsilon}_{2i,2i} + \langle \boldsymbol{\Gamma}_{2i} \rangle^2) + \frac{\xi_i}{\beta m_i} \right] = \sum_{i=1}^2 \dot{q}^{(i)}, \quad (143)$$

$$\begin{aligned} \dot{\Sigma} &= \sum_{i=1}^2 \left[\beta \xi_i (\boldsymbol{\Upsilon}_{2i,2i} + \langle \boldsymbol{\Gamma}_{2i} \rangle^2) - 2 \frac{\xi_i}{m_i} + \frac{\xi_i}{\beta m_i^2} \boldsymbol{\Upsilon}_{2i,2i}^{-1} \right] \\ &= \sum_{i=1}^2 \dot{\sigma}^{(i)}, \end{aligned} \quad (144)$$

and because of Eq. (40)

$$d_t S(t) = \sum_{i=1}^2 \left(\frac{\xi_i}{\beta m_i^2} \boldsymbol{\Upsilon}_{2i,2i}^{-1} - \frac{\xi_i}{m_i} \right). \quad (145)$$

In the following, the distribution for particle 1 $P_1(t)$ is needed. The latter is readily determined by marginalizing Eq. (138) over the coordinates $\boldsymbol{\Gamma}_2$ of the second particle:

$$P_1 = \frac{1}{2\pi \sqrt{\det \tilde{\boldsymbol{\Upsilon}}}} \exp \left[-\frac{1}{2} (\tilde{\boldsymbol{\Gamma}} - \langle \tilde{\boldsymbol{\Gamma}} \rangle)^\top \cdot \tilde{\boldsymbol{\Upsilon}}^{-1} \cdot (\tilde{\boldsymbol{\Gamma}} - \langle \tilde{\boldsymbol{\Gamma}} \rangle) \right], \quad (146)$$

with $\tilde{\boldsymbol{\Gamma}} = (x_1, v_1)^\top$ and the inverse of the marginalized covariance matrix $\tilde{\boldsymbol{\Upsilon}}$ that is given by

$$\begin{aligned} \tilde{\boldsymbol{\Upsilon}}_{11}^{-1} &= \frac{1}{(\boldsymbol{\Upsilon}_{34}^{-1})^2 - \boldsymbol{\Upsilon}_{33}^{-1} \boldsymbol{\Upsilon}_{44}^{-1}} [(\boldsymbol{\Upsilon}_{14}^{-1})^2 \boldsymbol{\Upsilon}_{33}^{-1} - 2 \boldsymbol{\Upsilon}_{13}^{-1} \boldsymbol{\Upsilon}_{14}^{-1} \boldsymbol{\Upsilon}_{34}^{-1} + \boldsymbol{\Upsilon}_{11}^{-1} (\boldsymbol{\Upsilon}_{34}^{-1})^2 + (\boldsymbol{\Upsilon}_{13}^{-1})^2 \boldsymbol{\Upsilon}_{44}^{-1} - \boldsymbol{\Upsilon}_{11}^{-1} \boldsymbol{\Upsilon}_{33}^{-1} \boldsymbol{\Upsilon}_{44}^{-1}], \\ \tilde{\boldsymbol{\Upsilon}}_{12}^{-1} &= \frac{1}{(\boldsymbol{\Upsilon}_{34}^{-1})^2 - \boldsymbol{\Upsilon}_{33}^{-1} \boldsymbol{\Upsilon}_{44}^{-1}} [\boldsymbol{\Upsilon}_{14}^{-1} \boldsymbol{\Upsilon}_{24}^{-1} \boldsymbol{\Upsilon}_{33}^{-1} - \boldsymbol{\Upsilon}_{14}^{-1} \boldsymbol{\Upsilon}_{23}^{-1} \boldsymbol{\Upsilon}_{34}^{-1} - \boldsymbol{\Upsilon}_{13}^{-1} \boldsymbol{\Upsilon}_{24}^{-1} \boldsymbol{\Upsilon}_{34}^{-1} + \boldsymbol{\Upsilon}_{12}^{-1} (\boldsymbol{\Upsilon}_{34}^{-1})^2 + \boldsymbol{\Upsilon}_{13}^{-1} \boldsymbol{\Upsilon}_{23}^{-1} \boldsymbol{\Upsilon}_{44}^{-1} - \boldsymbol{\Upsilon}_{12}^{-1} \boldsymbol{\Upsilon}_{33}^{-1} \boldsymbol{\Upsilon}_{44}^{-1}], \\ \tilde{\boldsymbol{\Upsilon}}_{22}^{-1} &= \frac{1}{(\boldsymbol{\Upsilon}_{34}^{-1})^2 - \boldsymbol{\Upsilon}_{33}^{-1} \boldsymbol{\Upsilon}_{44}^{-1}} [(\boldsymbol{\Upsilon}_{24}^{-1})^2 \boldsymbol{\Upsilon}_{33}^{-1} - 2 \boldsymbol{\Upsilon}_{23}^{-1} \boldsymbol{\Upsilon}_{24}^{-1} \boldsymbol{\Upsilon}_{34}^{-1} + \boldsymbol{\Upsilon}_{22}^{-1} (\boldsymbol{\Upsilon}_{34}^{-1})^2 + (\boldsymbol{\Upsilon}_{23}^{-1})^2 \boldsymbol{\Upsilon}_{44}^{-1} - \boldsymbol{\Upsilon}_{22}^{-1} \boldsymbol{\Upsilon}_{33}^{-1} \boldsymbol{\Upsilon}_{44}^{-1}]. \end{aligned} \quad (147)$$

Inserting Eq. (146) into Eqs. (45) and (70) gives the force contribution to the information flow from particle 2 to particle 1:

$$j_F^{(2 \rightarrow 1)} = -\frac{\beta}{m_1}(\tilde{\Upsilon}_{12}^{-1} \Upsilon_{13} + \tilde{\Upsilon}_{22}^{-1} \Upsilon_{23}), \quad (148)$$

which can be seen by noting that

$$-\frac{\beta}{m_1} \int d\Gamma P_1 x_2 \partial_{v_1} P_{2|1} = \frac{\beta}{m_1} \int d\Gamma P x_2 \partial_{v_1} \ln P_1. \quad (149)$$

Moreover, from Eq. (49) it follows for the effective entropy production rate that

$$\dot{\Sigma}^{(1)} = \beta \xi_1 (\tilde{\Upsilon}_{22} + \langle \Gamma_2 \rangle^2) - 2 \frac{\xi_1}{m_1} + \frac{\xi_1}{\beta m_1^2} \tilde{\Upsilon}_{22}^{-1}, \quad (150)$$

from which via Eqs. (73) and (144) we get the entropic contribution to the information flow:

$$j_S^{(2 \rightarrow 1)} = \beta \xi_1 (\tilde{\Upsilon}_{22} - \Upsilon_{22}) + \frac{\xi_1}{\beta m_1^2} (\tilde{\Upsilon}_{22}^{-1} - \Upsilon_{22}^{-1}). \quad (151)$$

Combining the last three equations with Eqs. (50), (51), and (52) yields

$$j^{(2 \rightarrow 1)} = \beta \xi_1 (\tilde{\Upsilon}_{22} - \Upsilon_{22}) + \frac{\xi_1}{\beta m_1^2} (\tilde{\Upsilon}_{22}^{-1} - \Upsilon_{22}^{-1}) - \frac{\beta}{m_1} (\tilde{\Upsilon}_{12}^{-1} \Upsilon_{13} + \tilde{\Upsilon}_{22}^{-1} \Upsilon_{23}), \quad (152)$$

$$d_t S_{2|1} = \frac{\xi_2}{\beta m_2^2} \Upsilon_{44}^{-1} - \frac{\xi_2}{m_2} - \frac{\beta}{m_1} (\tilde{\Upsilon}_{12}^{-1} \Upsilon_{13} + \tilde{\Upsilon}_{22}^{-1} \Upsilon_{23}) - \beta \xi_1 (\tilde{\Upsilon}_{22} - \Upsilon_{22}) - \frac{\xi_1}{\beta m_1^2} (\tilde{\Upsilon}_{22}^{-1} - \Upsilon_{22}^{-1}), \quad (153)$$

$$\dot{Q}^{(1)} = \frac{\xi_2}{\beta m_2^2} \Upsilon_{44}^{-1} + \frac{\xi_1}{m_1} - \frac{\xi_2}{m_2} - \frac{\beta}{m_1} (\tilde{\Upsilon}_{12}^{-1} \Upsilon_{13} + \tilde{\Upsilon}_{22}^{-1} \Upsilon_{23}) - \beta \xi_1 (\tilde{\Upsilon}_{22} + \langle \Gamma_2 \rangle^2) - \frac{\xi_1}{\beta m_1^2} (\tilde{\Upsilon}_{22}^{-1} - \Upsilon_{22}^{-1}). \quad (154)$$

B. Fast-dynamics limit

Since $g_2 = 0$, the limit of TSS implies that the second particle is at local equilibrium conditioned on the coordinates of particle 1. Within TSS, the effective force (104) reads

$$g^{(1)} = \beta^2 \frac{x_1}{k_2}, \quad (155)$$

and closes the effective Fokker-Planck Eq. (42):

$$\partial_t P_1 = -\nabla_1 \cdot [(\gamma_1 \cdot \tilde{\Gamma}) P_1] + \nabla_1 \cdot (D_1 \cdot \nabla_1 P_1), \quad (156)$$

with $\nabla_1 \equiv (\partial_{x_1}, \partial_{v_1})^\top$. The drift coefficient and the diffusion matrix read

$$\gamma_1 = \begin{pmatrix} 0 & 1 \\ -\frac{k_1}{m_1} - \frac{\beta^2}{k_2 m_1} & -\frac{\xi_1}{m_1} \end{pmatrix}, \quad D_1 = \begin{pmatrix} 0 & 0 \\ 0 & \frac{\xi_1}{\beta m_1^2} \end{pmatrix}. \quad (157)$$

This Fokker-Planck equation implies that we are dealing with a bivariate Ornstein-Uhlenbeck process; thus, its solution is given by a bivariate Gaussian [65]:

$$P_1 = \frac{1}{2\pi \sqrt{\det \tilde{\Upsilon}}} e^{-\frac{1}{2}(\tilde{\Gamma} - \langle \tilde{\Gamma} \rangle)^\top \cdot \tilde{\Upsilon}^{-1} \cdot (\tilde{\Gamma} - \langle \tilde{\Gamma} \rangle)}, \quad (158)$$

where the covariance matrix $\tilde{\Upsilon}$ is specified by Eq. (140) and the averages of the coordinates $\tilde{\Gamma}$ are determined as follows:

$$\langle \tilde{\Gamma} \rangle(t) = e^{\gamma_1 t} \cdot \tilde{\Gamma}(0). \quad (159)$$

In the following of this subsection, we employ the numerical values $\xi_1 = 0.8$, $\beta = 0.05$, $k_1 = 1$, $m_1 = 1$, while we consider three different spring constants k_2 , masses m_2 , and friction coefficients ξ_2 : ($k_2 = 15$, $m_{2,a} = 5$, $\xi_{2,a} = 0.75$), ($k_2 = 25$, $m_{2,b} = 7.5$, $\xi_{2,b} = 0.25$), and ($k_2 = 50$, $m_{2,c} = 10$, $\xi_{2,c} = 0.1$). This choice of parameters corresponds to an increasing separation of the time scales between the different stochastic dynamics of the two particles. In the order a - b - c , the second particle approaches equilibrium conditioned on the coordinates of the first particle: Since the interaction potential scales linearly in the inverse temperature [Eq. (134)], we chose a relatively small value for β to implement a weak-coupling condition between the first and second particle—a crucial requisite for the second particle to behave like an ideal heat reservoir [66,67]. As k_2 and m_2 increase and ξ_2 decreases, the relaxation time scale of the second particle further shrinks, hence the time scales of the particles dynamics start to separate, as desired. Moreover, we prepare the initial condition (95) with $P_1(0) = \delta[\tilde{\Gamma} - \tilde{\Gamma}(0)]$ with $\tilde{\Gamma}(0) = (2, 1)^\top$.

Figure 2 depicts in (a) the difference between the global \dot{Q} and effective heat current $\dot{Q}^{(1)}$ and in (b) the scaled difference between the global $\dot{\Sigma}$ and effective entropy production rate $\dot{\Sigma}^{(1)}$ as a function of time t . We observe that both the effective heat current and entropy production rate converge to the corresponding full quantities in the limit of TSS. The overall system remains out of equilibrium as reflected by finite (effective) heat currents and (effective) entropy production rates of the first particle. Since the corresponding single-particle definition for the heat, $\dot{q}^{(1)}$, does not agree with definition of the effective one [not shown in (a)], it follows that the time derivative of the conditional Shannon entropy, $d_t S_{2|1}$, remains finite in the limit of TSS. We furthermore note that the effective heat current and entropy production rate are in agreement with the time derivative of the heat (91) and entropy production (93) using the HMF formalism. Moreover, Fig. 2(c) shows that the directional information flow $j^{(2 \rightarrow 1)}$ vanishes in the limit of TSS. This in turn implies first that the additive contribution $\dot{\sigma}^{(2)}$ to the full entropy production rate becomes zero while the inverse information flow $j^{(1 \rightarrow 2)}$ remains finite. It furthermore follows from the nonpositivity of $j^{(2 \rightarrow 1)}$ that the nonpositive entropic contribution $j_S^{(2 \rightarrow 1)}$ dominates over the non-negative force contribution $j_F^{(2 \rightarrow 1)}$.

C. Large-mass limit

In the large- m_2 limit, the effective force (125) reads

$$g^{(1)} = -\beta x_2 \big|_{x_2=x_1}, \quad (160)$$

and closes the effective Fokker-Planck Eq. (42):

$$\partial_t P_1 = -\nabla_1 \cdot [(\gamma_1 \cdot \tilde{\Gamma} + g^{(1)}) P_1] + \nabla_1 \cdot (D_1 \cdot \nabla_1 P_1). \quad (161)$$

The constant drift coefficient, the scaled effective force vector, and the diffusion matrix read

$$\gamma_1 = \begin{pmatrix} 0 & 1 \\ -\frac{k_1}{m_1} & -\frac{\xi_1}{m_1} \end{pmatrix}, \quad g^{(1)} = \begin{pmatrix} 0 \\ -\frac{\beta x_1}{m_1} \end{pmatrix}, \quad D_1 = \begin{pmatrix} 0 & 0 \\ 0 & \frac{\xi_1}{\beta m_1^2} \end{pmatrix}. \quad (162)$$

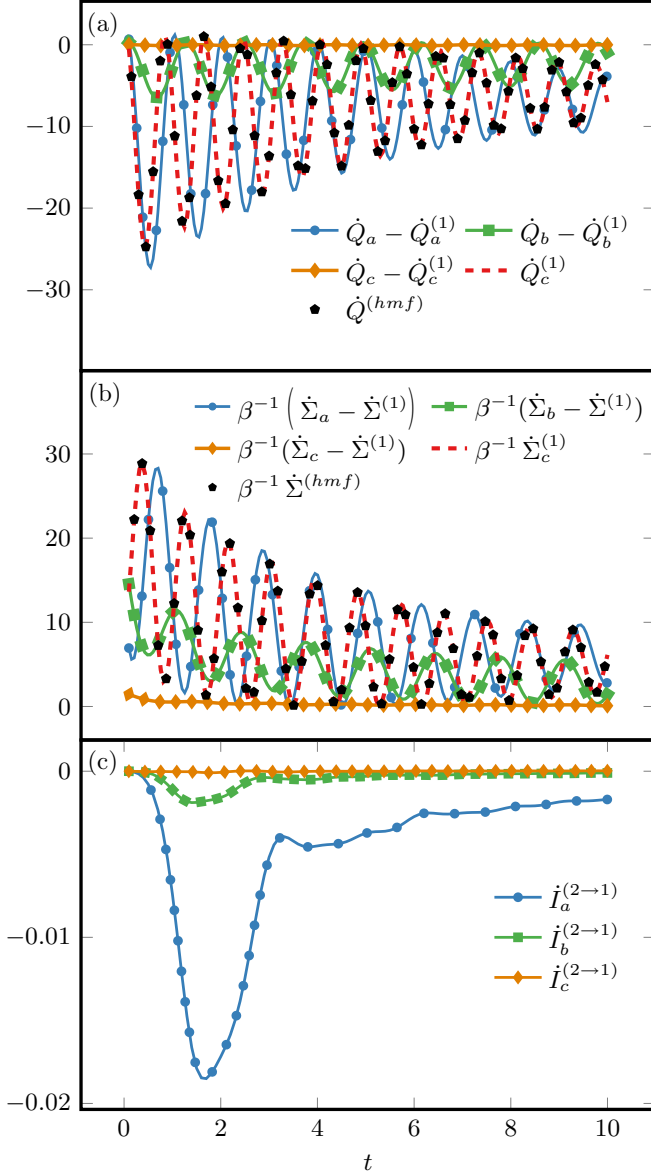


FIG. 2. Difference between the full \dot{Q} and effective heat current $\dot{Q}^{(1)}$ in (a) and between the scaled full $\beta \dot{\Sigma}$ and scaled effective entropy production rate $\beta \dot{\Sigma}^{(1)}$ in (b) as a function of time t . In these two figures, the parameter sets a , b , and c correspond to the solid blue line with circle markers, the solid green line with square markers, and the solid orange line with diamond markers, respectively. Moreover, the effective quantities for the parameters c and those based on the HMF are overlaid in panels (a) and (b) and correspond to the red dashed line and the black pentagon markers, respectively. The information flow $I^{(2 \rightarrow 1)}$ as a function of time and for the parameter sets a (solid blue line with circle markers), b (solid green line with square markers), and c (solid orange line with diamond markers) is depicted in panel (c).

This partial differential equation is supplemented by the initial condition $P_1(0) = \delta[\tilde{\Gamma} - \tilde{\Gamma}(0)]$ with $\tilde{\Gamma}(0) = (2, 1)^\top$. The averages are determined as follows:

$$\langle \tilde{\Gamma} \rangle(t) = e^{\gamma_1 t} \cdot \tilde{\Gamma}(0) + \int_0^t e^{\gamma_1(t-t')} \cdot \mathbf{g}^{(1)}(t') dt', \quad (163)$$

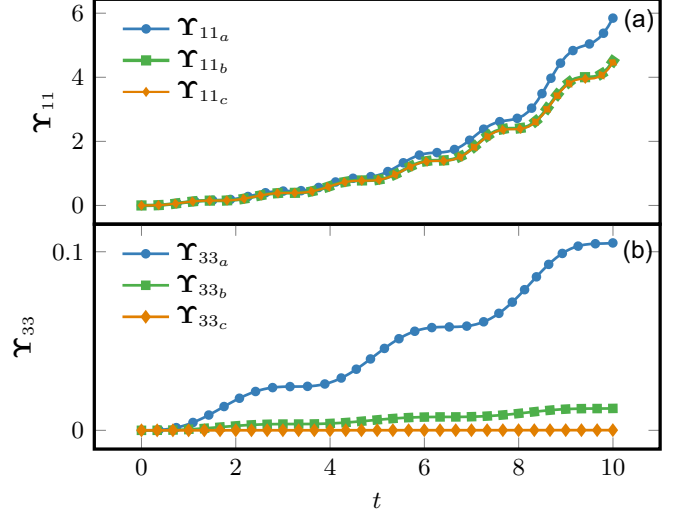


FIG. 3. Variance Υ_{11} in (a) and Υ_{33} in (b) of the positional degrees of freedom x_1 and x_2 , respectively, as a function of time t and for the different parameter sets a (solid blue line with circle markers), b (solid green line with square markers), and c (solid orange line with diamond markers).

while the coordinates (x_t, v_t) of the second particle follow the solution of the deterministic equation of motion (26):

$$\begin{aligned} x_t &= 2 \cos\left(\frac{k_2}{m_2} t\right) + \frac{m_2}{k_2} \sin\left(\frac{k_2}{m_2} t\right), \\ v_t &= \cos\left(\frac{k_2}{m_2} t\right) - 2 \frac{k_2}{m_2} \sin\left(\frac{k_2}{m_2} t\right), \end{aligned} \quad (164)$$

for the initial condition as chosen above. In the following, we employ the numerical values $\xi_1 = 0.3$, $\xi_2 = 1.5$, $\beta = 1$, $k_1 = 4$, $m_1 = 1$, while we consider three different masses m_2 and constants k_2 such that their ratio remains constant: $(m_{2,a} = 4, k_{2,a} = 3.8)$, $(m_{2,b} = 40, k_{2,b} = 38)$, and $(m_{2,c} = 400, k_{2,c} = 380)$. It is important to note that the set of parameters a , b , and c is chosen such that the ratio of m_2 and k_2 remains constant and thus leaves the deterministic trajectory of the second particle invariant according to Eq. (164).

Figure 3 depicts the variances Υ_{11} and Υ_{33} of the positional variables x_1 and x_2 , in panels (a) and (b), respectively. As expected, the fluctuations of the first particle do not exhibit striking qualitative changes since the variance of the second particle vanishes with growing mass m_2 . We verify that

$$\Upsilon_{ij} = 0, \quad \forall ij \neq \{11, 12, 21, 22\}, \quad (165)$$

thus confirming that the second particle behaves deterministically in the large- m_2 limit as prescribed by the equations of motion (164).

Next, Fig. 4(a) shows that the effective heat current, $\dot{Q}^{(1)}$, converges to the full one, \dot{Q} , minus the macroscopic dissipation of the heavy particle, $\xi_2 v_t^2$, as m_2 increases.

This macroscopic term is naturally non-negative and periodic with the frequency k_2/m_2 due to the choice of a harmonic potential (134). Furthermore, Fig. 4(b) illustrates the convergence of the effective entropy production $\dot{\Sigma}^{(1)}$ to the full one, $\dot{\Sigma}$, plus the macroscopic dissipation of the heavy particle with increasing m_2 . Since the single-particle definitions for

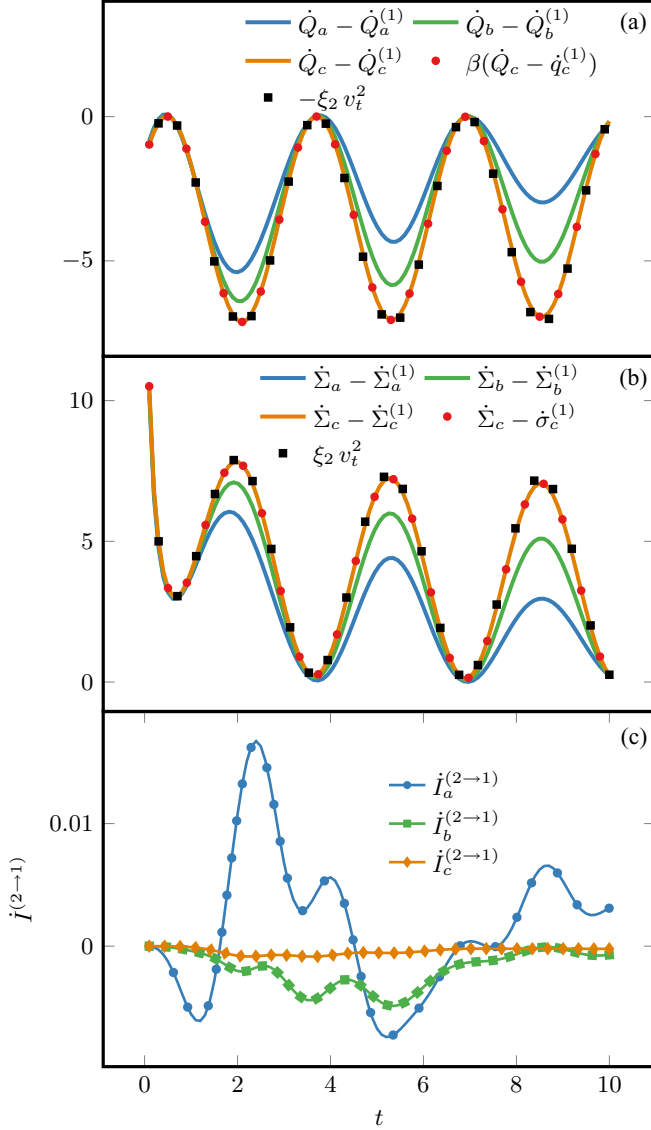


FIG. 4. Difference between the full \dot{Q} and effective heat current $\dot{Q}^{(1)}$ in (a) and between the full $\dot{\Sigma}$ and effective entropy production rate $\dot{\Sigma}^{(1)}$ in (b) as a function of time t . In panel (a) [(b)], the parameter sets a , b , and c correspond to the upper [lower] blue intermediate green and lower [upper] orange line, respectively. Moreover, the difference between the full \dot{Q} [$\dot{\Sigma}$] and the naive definition of the effective heat current [entropy production rate] $\dot{q}^{(1)}$ [$\dot{\sigma}^{(1)}$] is overlaid in panel (a) [(b)]. The information flow $\dot{I}^{(2 \rightarrow 1)}$ as a function of time and for the parameter sets a [solid blue line with circle markers], b [solid green line with square markers], and c [solid orange line with diamond markers] is depicted in panel (c).

the heat current, $\dot{q}^{(1)}$, and the entropy production rate, $\dot{\sigma}^{(1)}$, also converge to the full quantities, respectively, it follows that the time derivative of the conditional Shannon entropy, $d_t S_{2|1}$, and the information flow from the light to the heavy particle, $\dot{I}^{(1 \rightarrow 2)}$, vanish as m_2 grows.

Finally, in Fig. 4(c) the directional information flow from the heavy to the light particle $\dot{I}^{(2 \rightarrow 1)}$ is shown to decrease in modulus with increasing m_2 . It is interesting to note that

the vanishing directional flow becomes negative if m_2 is sufficiently large. This means that the nonpositive entropic contribution converges at a slower rate to zero than the force one does.

V. CONCLUSION

In this paper, we presented three coarse-graining approaches for the thermodynamics of two interacting underdamped Brownian particles: the observation of only one particle while the other one has been coarse-grained, the partitioning of the two-body system into two single-particle systems exchanging information flows, and the Hamiltonian of mean force formalism. We demonstrated that the effective thermodynamics of the first and third approach is equivalent to the correct global thermodynamics in the limit of time-scale separation between the two particles, where the faster evolving particle equilibrates with respect to the coordinates of the more slowly evolving particle. Conversely, we observed a mismatch between the effective and full thermodynamics in the bipartite case, since the entropic contribution due to the coupling of the two particles is not taken into account. Physically, in this limit the faster evolving particle becomes part of the heat reservoir to which the other particle is coupled. Conversely, if one particle becomes deterministic because of an exceedingly large mass compared to the other particle's mass, it acts as an additional work source on the lighter particle. In this case, the effective thermodynamics of the first two of the aforementioned three approaches agree, up to a simple macroscopic term related to the dissipation of the work source, with the correct global one. The Hamiltonian of mean force formalism, however, was shown to be incompatible with the large-mass limit. In fact, the same is true for any physical regime outside the time-scale separation limit. This reflects that the Hamiltonian of mean force formalism was originally motivated by and employed in equilibrium thermostatics. These theoretical predictions were confirmed via an analytically tractable model made up of two linearly coupled harmonic oscillators. We remark that the generalization to an arbitrary many-body system, where particles 1 and 2 are replaced by two subsets of interacting particles, is straightforward. Since the findings for systems with arbitrarily many bodies are identical to the results for the two-body setup reported above, an explicit presentation of the former is omitted.

In view of applications, we remark that due to the emergence of a velocity-dependent nonconservative force in the effective description our paper could be of significant methodological value for active particles, where the issue of coarse-graining is indeed very important. To conclude, we leave the study of effective fluctuating thermodynamics in underdamped systems for future works. In this context, it would also be interesting to explore if other coarse-graining schemes that were applied to jump processes, for instance, as proposed in Refs. [68,69], can also be utilized in underdamped Fokker-Planck systems. These studies could give rise to new strategies for systematically and thermodynamically consistently coarse-graining many-body systems.

ACKNOWLEDGMENTS

We gratefully acknowledge funding by the National Research Fund, Luxembourg (AFR Ph.D. Grant 2016 No.

11271777 and ATTRACT Project No. FNR/A11/02) and by the European Research Council (Project NanoThermo, ERC-2015-CoG Agreement No. 681456).

-
- [1] U. Seifert, Stochastic thermodynamics, fluctuation theorems and molecular machines, *Rep. Prog. Phys.* **75**, 126001 (2012).
 - [2] C. Van den Broeck and M. Esposito, Ensemble and trajectory thermodynamics: A brief introduction, *Physica A* **418**, 6 (2015).
 - [3] K. Sekimoto, *Stochastic Energetics*, 1st ed. (Springer-Verlag, Berlin, 2010).
 - [4] X.-J. Zhang, H. Qian, and M. Qian, Stochastic theory of nonequilibrium steady states and its applications. Part I, *Phys. Rep.* **510**, 1 (2012).
 - [5] H. Ge, M. Qian, and H. Qian, Stochastic theory of nonequilibrium steady states. Part II: Applications in chemical biophysics, *Phys. Rep.* **510**, 87 (2012).
 - [6] C. Jarzynski, Equalities and inequalities: Irreversibility and the second law of thermodynamics at the nanoscale, *Annu. Rev. Condens. Matter Phys.* **2**, 329 (2011).
 - [7] R. Rao and M. Esposito, Conservation laws shape dissipation, *New J. Phys.* **20**, 023007 (2018).
 - [8] S. Ciliberto, Experiments in Stochastic Thermodynamics: Short History and Perspectives, *Phys. Rev. X* **7**, 021051 (2017).
 - [9] K. Proesmans, Y. Dreher, M. Gavrilov, J. Bechhoefer, and C. Van den Broeck, Brownian Duet: A Novel Tale of Thermodynamic Efficiency, *Phys. Rev. X* **6**, 041010 (2016).
 - [10] J. P. Pekola, O.-P. Saira, V. F. Maisi, A. Kemppinen, M. Möttönen, Y. A. Pashkin, and D. V. Averin, Single-electron current sources: Toward a refined definition of the ampere, *Rev. Mod. Phys.* **85**, 1421 (2013).
 - [11] R. Rao and M. Esposito, Nonequilibrium Thermodynamics of Chemical Reaction Networks: Wisdom from Stochastic Thermodynamics, *Phys. Rev. X* **6**, 041064 (2016).
 - [12] R. Rao and M. Esposito, Conservation laws and work fluctuation relations in chemical reaction networks, *J. Chem. Phys.* **149**, 245101 (2018).
 - [13] C. Nardini, É. Fodor, E. Tjhung, F. van Wijland, J. Tailleur, and M. E. Cates, Entropy Production in Field Theories without Time-Reversal Symmetry: Quantifying the Non-Equilibrium Character of Active Matter, *Phys. Rev. X* **7**, 021007 (2017).
 - [14] L. Dabelow, S. Bo, and R. Eichhorn, Irreversibility in Active Matter Systems: Fluctuation Theorem and Mutual Information, *Phys. Rev. X* **9**, 021009 (2019).
 - [15] J. M. R. Parrondo, J. M. Horowitz, and T. Sagawa, Thermodynamics of information, *Nat. Phys.* **11**, 131 (2015).
 - [16] M. Esposito and C. Van den Broeck, Three faces of the second law. I. Master equation formulation, *Phys. Rev. E* **82**, 011143 (2010).
 - [17] C. Van den Broeck and M. Esposito, Three faces of the second law. II. Fokker-Planck formulation, *Phys. Rev. E* **82**, 011144 (2010).
 - [18] M. Esposito, Stochastic thermodynamics under coarse graining, *Phys. Rev. E* **85**, 041125 (2012).
 - [19] G. Verley, C. Van den Broeck, and M. Esposito, Work statistics in stochastically driven systems, *New J. Phys.* **16**, 095001 (2014).
 - [20] B. Altaner and J. Vollmer, Fluctuation-Preserving Coarse Graining for Biochemical Systems, *Phys. Rev. Lett.* **108**, 228101 (2012).
 - [21] A. Celani, S. Bo, R. Eichhorn, and E. Aurell, Anomalous Thermodynamics at the Microscale, *Phys. Rev. Lett.* **109**, 260603 (2012).
 - [22] G. Bulnes Cuetara, M. Esposito, and P. Gaspard, Fluctuation theorems for capacitively coupled electronic currents, *Phys. Rev. B* **84**, 165114 (2011).
 - [23] A. Wachtel, R. Rao, and M. Esposito, Thermodynamically consistent coarse graining of biocatalysts beyond Michaelis-Menten, *New J. Phys.* **20**, 042002 (2018).
 - [24] J. Mehl, B. Lander, C. Bechinger, V. Blickle, and U. Seifert, Role of Hidden Slow Degrees of Freedom in the Fluctuation Theorem, *Phys. Rev. Lett.* **108**, 220601 (2012).
 - [25] T. Herpich, J. Thingna, and M. Esposito, Collective Power: Minimal Model for Thermodynamics of Nonequilibrium Phase Transitions, *Phys. Rev. X* **8**, 031056 (2018).
 - [26] T. Herpich and M. Esposito, Universality in driven Potts models, *Phys. Rev. E* **99**, 022135 (2019).
 - [27] S. Sheng, T. Herpich, G. Diana, and M. Esposito, Thermodynamics of majority-logic decoding in information erasure, *Entropy* **21**, 284 (2019).
 - [28] K. Sekimoto, Microscopic heat from the energetics of stochastic phenomena, *Phys. Rev. E* **76**, 060103(R) (2007).
 - [29] A. Gomez-Marín, J. M. R. Parrondo, and C. Van den Broeck, Lower bounds on dissipation upon coarse graining, *Phys. Rev. E* **78**, 011107 (2008).
 - [30] T. Leonard, B. Lander, U. Seifert, and T. Speck, Stochastic thermodynamics of fluctuating density fields: Non-equilibrium free energy differences under coarse-graining, *J. Chem. Phys.* **139**, 204109 (2013).
 - [31] G. Bisker, M. Poletti, T. R. Gingrich, and J. M. Horowitz, Hierarchical bounds on entropy production inferred from partial information, *J. Stat. Mech.: Theory Exp.* (2017) 093210.
 - [32] F. Knoch and T. Speck, Cycle representatives for the coarse-graining of systems driven into a non-equilibrium steady state, *New J. Phys.* **17**, 115004 (2015).
 - [33] I. A. Martínez, G. Bisker, J. M. Horowitz, and J. M. R. Parrondo, Inferring broken detailed balance in the absence of observable currents, *Nat. Commun.* **10**, 3542 (2019).
 - [34] M. Kahlen and J. Ehrich, Hidden slow degrees of freedom and fluctuation theorems: An analytically solvable model, *J. Stat. Mech.: Theory Exp.* (2018) 063204.
 - [35] M. Uhl, P. Pietzonka, and U. Seifert, Fluctuations of apparent entropy production in networks with hidden slow degrees of freedom, *J. Stat. Mech.: Theory Exp.* (2018) 023203.
 - [36] D. Gupta and S. Sabhapandit, Partial entropy production in heat transport, *J. Stat. Mech.: Theory Exp.* (2018) 063203.
 - [37] S. Bo and A. Celani, Entropy production in stochastic systems with fast and slow time-scales, *J. Stat. Phys.* **154**, 1325 (2014).

- [38] H. S. Leff and A. F. Rex, *Maxwell's Demon: Entropy, Classical and Quantum Information, Computing* (Princeton University, Princeton, NJ, 1990).
- [39] J. M. Horowitz and M. Esposito, Thermodynamics with Continuous Information Flow, *Phys. Rev. X* **4**, 031015 (2014).
- [40] J. M. Horowitz and H. Sandberg, Second-law-like inequalities with information and their interpretations, *New J. Phys.* **16**, 125007 (2014).
- [41] J. M. Horowitz, Multipartite information flow for multiple maxwell demons, *J. Stat. Mech.: Theory Exp.* (2015) P03006.
- [42] 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.: Theory Exp.* (2014) P02016.
- [43] U. Seifert, First and Second Law of Thermodynamics at Strong Coupling, *Phys. Rev. Lett.* **116**, 020601 (2016).
- [44] C. Jarzynski, Stochastic and Macroscopic Thermodynamics of Strongly Coupled Systems, *Phys. Rev. X* **7**, 011008 (2017).
- [45] P. Strasberg and M. Esposito, Stochastic thermodynamics in the strong coupling regime: An unambiguous approach based on coarse graining, *Phys. Rev. E* **95**, 062101 (2017).
- [46] C. Jarzynski, Comparison of far-from-equilibrium work relations, *C. R. Phys.* **8**, 495 (2007).
- [47] K. Sekimoto, Langevin equation and thermodynamics, *Prog. Theor. Phys. Suppl.* **130**, 17 (1998).
- [48] U. Seifert, Entropy Production Along a Stochastic Trajectory and an Integral Fluctuation Theorem, *Phys. Rev. Lett.* **95**, 040602 (2005).
- [49] K. H. Kim and H. Qian, Entropy Production of Brownian Macromolecules with Inertia, *Phys. Rev. Lett.* **93**, 120602 (2004).
- [50] K. H. Kim and H. Qian, Fluctuation theorems for a molecular refrigerator, *Phys. Rev. E* **75**, 022102 (2007).
- [51] For velocity-dependent nonconservative forces, e.g., magnetic forces, the following procedure is analogous to the case of velocity-independent forces. The only formal modifications are the additional terms, $\sum_i b1/m_i \int d\Gamma_i P_i \partial_{v_i} \cdot \mathbf{g}_i$, that appear in the entropy balance equation (40) [cf. Eq. (17)].
- [52] P. Romanczuk, M. Bär, W. Ebeling, B. Lindner, and L. Schimansky-Geier, Active Brownian particles, *Eur. Phys. J.: Spec. Top.* **202**, 1 (2012).
- [53] C. Ganguly and D. Chaudhuri, Stochastic thermodynamics of active Brownian particles, *Phys. Rev. E* **88**, 032102 (2013).
- [54] D. Chaudhuri, Active Brownian particles: Entropy production and fluctuation response, *Phys. Rev. E* **90**, 022131 (2014).
- [55] R. García-García, P. Collet, and L. Truskinovsky, Guided active particles, *Phys. Rev. E* **100**, 042608 (2019).
- [56] T. Schreiber, Measuring Information Transfer, *Phys. Rev. Lett.* **85**, 461 (2000).
- [57] A. E. Allahverdyan, D. Janzing, and G. Mahler, Thermodynamic efficiency of information and heat flow, *J. Stat. Mech.: Theory Exp.* (2009) P09011.
- [58] X. S. Liang and R. Kleeman, Information Transfer between Dynamical System Components, *Phys. Rev. Lett.* **95**, 244101 (2005).
- [59] A. J. Majda and J. Harlim, Information flow between subspaces of complex dynamical systems, *Proc. Natl. Acad. Sci. USA* **104**, 9558 (2007).
- [60] A. C. Barato, D. Hartich, and U. Seifert, Information-theoretic versus thermodynamic entropy production in autonomous sensory networks, *Phys. Rev. E* **87**, 042104 (2013).
- [61] S. Goldt and U. Seifert, Stochastic Thermodynamics of Learning, *Phys. Rev. Lett.* **118**, 010601 (2017).
- [62] G. Diana and M. Esposito, Mutual entropy production in bipartite systems, *J. Stat. Mech.: Theory Exp.* (2014) P04010.
- [63] K. Ptaszyński and M. Esposito, Thermodynamics of Quantum Information Flows, *Phys. Rev. Lett.* **122**, 150603 (2019).
- [64] J. G. Kirkwood, Statistical mechanics of fluid mixtures, *J. Chem. Phys.* **3**, 300 (1935).
- [65] H. Risken, *The Fokker-Planck Equation: Methods of Solution and Applications* (Springer-Verlag, Berlin, 1989).
- [66] H.-P. Breuer and F. Petruccione, *The Theory of Open Quantum Systems* (Oxford University, New York, 2006).
- [67] K. Lindenberg and B. J. West, *The Nonequilibrium Statistical Mechanics of Open and Closed Systems* (Wiley, New York, 1990).
- [68] M. Polettini and M. Esposito, Effective Thermodynamics for a Marginal Observer, *Phys. Rev. Lett.* **119**, 240601 (2017).
- [69] M. Polettini and M. Esposito, Effective fluctuation and response theory, *J. Stat. Phys.* **176**, 94 (2019).