Approach to equilibrium and nonequilibrium stationary distributions of interacting many-particle systems that are coupled to different heat baths

Roland R. Netz *
Fachbereich Physik, Freie Universität Berlin, 14195 Berlin, Germany

(Received 26 August 2019; revised manuscript received 15 November 2019; accepted 23 January 2020; published 18 February 2020)

A Hamiltonian-based model of many harmonically interacting massive particles that are subject to linear friction and coupled to heat baths at different temperatures is used to study the dynamic approach to equilibrium and nonequilibrium stationary states. An equilibrium system is here defined as a system whose stationary distribution equals the Boltzmann distribution, the relation of this definition to the conditions of detailed balance and vanishing probability current is discussed both for underdamped as well as for overdamped systems. Based on the exactly calculated dynamic approach to the stationary distribution, the functional that governs this approach, which is called the free entropy $\mathcal{S}_{\rm free}(t)$, is constructed. For the stationary distribution $\mathcal{S}_{\rm free}(t)$ becomes maximal and its time derivative, the free entropy production $\dot{\mathcal{S}}_{\rm free}(t)$, is minimal and vanishes. Thus, $\mathcal{S}_{\rm free}(t)$ characterizes equilibrium as well as nonequilibrium stationary distributions by their extremal and stability properties. For an equilibrium system, i.e., if all heat baths have the same temperature, the free entropy equals the negative free energy divided by temperature and thus corresponds to the Massieu function which was previously introduced in an alternative formulation of statistical mechanics. Using a systematic perturbative scheme for calculating velocity and position correlations in the overdamped massless limit, explicit results for few particles are presented: For two particles localization in position and momentum space is demonstrated in the nonequilibrium stationary state, indicative of a tendency to phase separate. For three elastically interacting particles heat flows from a particle coupled to a cold reservoir to a particle coupled to a warm reservoir if the third reservoir is sufficiently hot. This does not constitute a violation of the second law of thermodynamics, but rather demonstrates that a particle in such a nonequilibrium system is not characterized by an effective temperature which equals the temperature of the heat bath it is coupled to. Active particle models can be described in the same general framework, which thereby allows us to characterize their entropy production not only in the stationary state but also in the approach to the stationary nonequilibrium state. Finally, the connection to nonequilibrium thermodynamics formulations that include the reservoir entropy production is discussed.

DOI: 10.1103/PhysRevE.101.022120

I. INTRODUCTION

Systems that even in their stationary state are not in equilibrium have in the last decades received renewed attention since standard concepts of thermodynamics and statistical mechanics do not work and because of their experimental relevance [1–14]. Some of the motivation for studying such systems comes from experimental observations of molecular processes in living systems, which are fundamentally nonequilibrium (NEQ) [15,16]. Apart from biological applications, experimental advances allow for the construction of NEQ systems either from biological components that are driven out of equilibrium by ATP consumption or by using macromolecular and colloidal assemblies that can be driven out of equilibrium by chemical fuels or by applying external forces [17,18].

Two prominent features of NEQ systems are departures from the canonical Boltzmann distribution and violation of the fluctuation-dissipation theorem, which were recently demonstrated to go hand in hand for a harmonically coupled particle system [19], the fundamental hallmark of NEQ systems is the violation of the detailed-balance condition [5,8]. A particularly striking example for the breakdown of equilibrium statistical mechanics is the occurrence of phase transitions induced by NEQ driving. As a result of external forces acting on particle systems, laning and phase separations have been observed [20–25]. Internal NEQ effects can generate particle propulsion that is sustained by chemical or biochemical means. NEQ effects can also be induced by stochastic forces due to the coupling of particles to different heat or chemical energy reservoirs. Such internal NEQ effects have also been shown to lead to clustering and phase separation [26–35]. Experimentally, symmetry-breaking transitions in suspensions of swimming bacteria and filament systems driven by motor proteins have indeed been demonstrated [36,37]. Also in more coarse-grained models, defined by migration rules, collective ordering can be obtained and describes the response of pedestrians to spatial confinement or the flocking and swarming of animals [12,38]. NEQ transformations also occur for polymers that are driven by externally applied torques [39,40].

The equilibrium fluctuation-dissipation theorem (FDT) describes the dynamical response of an equilibrium system to a small perturbation. Since it is a key concept for systems close

*rnetz@physik.fu-berlin.de

to equilibrium, and since any violation of the FDT clearly signals that a system is off equilibrium, the derivation of generalized fluctuation-dissipation relations that also hold for NEQ systems is central and has been discussed in the context of laser [41], chaotic [42], glassy [43], driven colloidal [44], sheared [45–47], and active systems [48]. Generalized NEQ fluctuation-dissipation relations were derived [41,49–54] and compared with experimental data for glasses [55], colloids [56,57], bundles of biological filaments [58,59], living cells [60], and biological gels that are driven by motor proteins in the presence of ATP [19,61].

Significant theoretical progress has been made in the characterization of NEQ systems by generalized NEQ fluctuation-dissipation relations, as discussed above, and by fluctuation relations that give bounds and exact relations involving trajectory ensembles [62–64]. The field of stochastic thermodynamics has linked the statistics of trajectories to entropy-production contributions [14,65]. A good fraction of contemporary work on NEQ systems is concerned with analyzing the solutions of governing dynamic equations either by simulation techniques or, for simple systems, by analytical methods. In essence, and as acknowledged by most workers in the field, NEQ theory is far from the predictive power and understanding furnished by, e.g., the usage of thermodynamic potentials in the context of equilibrium scenarios. For an isolated system the entropy is maximized, but this provides little help for the type of NEQ problems one is typically interested in, because they are not isolated. For a system coupled to a single heat reservoir and in the absence of external driving forces, which gives rise to the equilibrium canonical ensemble, the free energy is minimized. The free energy, when evaluated exactly or approximately, allows us to predict phase transitions, structures, and all static properties of an equilibrium system. In the search for a similarly useful framework, the first theoretical studies on NEQ systems formulated general extremal principles that express the system's tendency to extremize its dissipation, i.e., its entropy production [2,3]. These early extremal principles were limited to linear and homogeneous systems close to equilibrium and included interactions only indirectly in terms of phenomenological coefficients. Subsequently, the master equation approach allowed the derivation of extremal principles for general NEQ stationary states in terms of a generalized entropy [5,6,8,11]. More recently, Onsager's variational principle [1] was revisited and used for the study of various NEQ problems in soft condensed matter [66].

In this paper we start from a quadratic Hamiltonian model with general interactions. By adding friction terms and stochastic fields to the Hamilton equations, we arrive at the general linear Hamiltonian-based many-dimensional Langevin equation that, for suitably chosen friction and stochastic parameters, describes a many-body system of massive particles coupled to multiple heat baths with different and well-defined temperatures, a model that corresponds to the nonequilibrium version of the multidimensional Ornstein-Uhlenbeck process and has in certain limits been studied in literature [4,51,67–71].

The explicit presence of heat baths with different temperatures allows us to prepare the system in a unique NEQ stationary state and to calculate all contributions to the entropy production as the system approaches the stationary state. Instead of considering trajectories in phase space, we base our theory on the time-dependent distribution. The key point of our model is that we can exactly calculate the time derivative of the distribution entropy and from that construct, by comparison with the independently calculated relaxation of the distribution, the time-dependent functional that governs the approach to equilibrium as well to NEQ stationary distributions. In analogy to the relation between the energy and the free energy, this functional is called the free entropy, $\mathcal{S}_{\rm free}(t)$, since it contains the distribution entropy of the system and accounts for the interactions within the system and the coupling to the reservoirs. Using the free entropy $\mathcal{S}_{\rm free}(t)$, the total entropy $\mathcal{S}_{\rm tot}(t)$, which includes the entropy of the interacting particle system $\mathcal{S}(t)$ as well as the reservoir entropy $\mathcal{S}_{\rm res}(t)$, can be decomposed as

$$\mathcal{S}_{\rm tot}(t) = \mathcal{S}(t) + \mathcal{S}_{\rm res}(t) = \mathcal{S}_{\rm free}(t) + t\dot{\mathcal{S}}^{\circ}_{\rm res} \tag{1}$$

up to unimportant constants, where $\dot{\mathcal{S}}^{\circ}_{\rm res}$ denotes the entropy production due to heat transfer with all reservoirs in the unique stationary state. While $\mathcal{S}_{\rm res}(t)$ is difficult to calculate for the stochastic models used for the description of heat reservoirs and in fact increases boundlessly, $\mathcal{S}_{\rm free}(t)$ and $\dot{\mathcal{S}}^{\circ}_{\rm res}$ can be explicitly calculated. Similar to the free energy of statistical mechanics, different observables can be derived from the free entropy by taking suitable derivatives, as is shown in Sec. IV A.

For an equilibrium system, i.e., when all heat baths have the same temperature and $\dot{\mathcal{S}}^{\circ}_{\rm res} = 0$, and in the stationary state, the free entropy is time-independent and equals the negative free energy divided by temperature,

$$\mathcal{S}^{\bullet}_{\rm free} = -\mathcal{F}^{\bullet}/T = \mathcal{S}^{\bullet} - \mathcal{U}^{\bullet}/T, \tag{2}$$

where filled circles in our paper denote the equilibrium stationary state. In fact, the free entropy functional had been used by Massieu already in 1869 [72,73], a few years before Gibbs introduced his energy transforms. The advantage of the free entropy functional has been pointed out by Planck [74] and Schrödinger [75], while the name was introduced more recently in the mathematical literature [76,77]. The free entropy is central in the context of our model, since the presence of different heat bath temperatures does not allow the definition of a unique NEQ version of the free energy.

In a number of previous papers functionals were derived that in a NEQ stationary state are extremal and thereby allow the study of the relaxation and the stability of NEQ systems and the relation of these functionals to the Kullback-Leibler entropy [78] was pointed out [5–8,11,79–83]. Our model differs from those works since we introduce NEQ by the coupling to multiple heat baths with different temperatures. While the early approaches to NEQ thermodynamics were centered on the total entropy production $\dot{\mathcal{S}}_{\rm tot}(t)$, i.e., the time derivative of the total entropy including the reservoirs [2,3], in later developments the entropy production due to heat transfer from the reservoirs, which in a NEQ stationary state is constant and given by $\dot{\mathcal{S}}^{\circ}_{\rm res}$, has been separated off and called the house-keeping entropy [9,10,84]. One advantage of our model is that since the particles have finite masses, the heat fluxes between the particles and the heat reservoirs can be calculated from energy balance considerations including the kinetic energy and thus the stationary reservoir entropy production $\dot{\mathcal{S}}^{\circ}_{\rm res}$ can be derived straightforwardly. For this we introduce a systematic perturbation scheme using the particle masses as expansion parameters to calculate NEQ mixed position-velocity correlations.

We here show that for a system coupled to different temperature reservoirs, the free entropy functional $\mathcal{S}_{\rm free}(t)$ can be written explicitly and is for the NEQ stationary distribution maximal and constant in time. The free entropy production $\dot{\mathcal{S}}_{\rm free}(t)$ is positive except at the stationary NEQ state, for which it vanishes:

$$\dot{\mathcal{S}}_{\rm free}(t) \geqslant 0. \tag{3}$$

This shows that the NEQ stationary state is stable with respect to small perturbations and that in the NEQ stationary state the total entropy production is given solely by the reservoirs. The time derivative of the free entropy production, which would be the second time derivative of the free entropy, is not needed, unlike early NEQ approaches [2,3]. Our framework thus treats NEQ and equilibrium systems on the same footing, as for an equilibrium system the NEQ free entropy smoothly crosses over to the standard free energy divided by $-T$ [see Eq. (2)], and so the standard equilibrium and stability conditions of statistical mechanics are recovered.

Interestingly, the dynamic approach to the equilibrium and to the NEQ stationary distributions obey the same differential equation, as is shown in Sec. III C. In fact, while equilibrium and NEQ stationary distributions exhibit many fundamental differences, the relaxation times that characterize the approach to stationarity are independent of the heat bath temperatures and therefore do not allow distinguishing equilibrium from NEQ systems. While this might be a simplification due to our neglect of nonlinear interactions, we argue that nonlinear systems can typically be quadratically approximated around locally stable states, and thus our results should also apply to sufficiently well-behaved nonlinear systems.

We also demonstrate that the definition of equilibrium we are using in this paper, which is based on the Boltzmann distribution, is equivalent to the detailed-balance condition only if the friction matrix is symmetric and if the random fields couple separately to position and velocity degrees of freedom. For a simple system consisting of two coupled massive particles, we show that if the Boltzmann distribution is realized, the fluctuation-dissipation theorem is satisfied, even when the friction matrix is asymmetric (and thus the condition of detailed balance is not satisfied). This shows that equilibrium definitions based on Boltzmann statistics, on the fluctuation-dissipation theorem and on the detailed-balance condition are equivalent only for symmetric friction matrices and that the detailed-balance condition is the strictest of all three.

As a simple application of our model we present results for three interacting massive particles that are coupled to temperature reservoirs at different temperatures, for which we demonstrate that heat flows from a particle coupled to a cold reservoir to a particle coupled to a warm reservoir if the third reservoir is sufficiently hot. This, of course, does not constitute a violation of the second law of thermodynamics. Rather, this NEQ entrainment effect can be rationalized by the fact that the reservoirs are not coupled directly to each other but rather indirectly via the particles, and that the particles are not characterized solely by the heat bath temperatures. This point can be explained in more detail by considering just two particles that are coupled to different heat baths: We demonstrate that the concept of a NEQ effective temperature has only a rather limited value, since each covariance matrix entry of two coupled NEQ particles would have to be attributed a different effective temperature. For two particles we also demonstrate that NEQ effects give rise to localization effects both in position and in momentum space, which is reminiscent of attractive interactions. This reflects the tendency of NEQ systems to phase separate in both position and momentum space. Finally, we show how active particle models can be described using our general framework and discuss the connection between the free entropy production and the total entropy production that includes the reservoirs.

In the following sections we first treat general NEQ systems and derive the necessary conditions to reach a stationary state and a stationary equilibrium state, described by the Lyapunov and the Lyapunov-Boltzmann equations, respectively. In Sec. IV B we start treating the core model of this paper, where particles are coupled to heat reservoirs that are characterized by different temperatures, and present various explicit examples.

II. MANY-PARTICLE HAMILTONIAN MODEL

A. From Hamilton to Langevin equations

To proceed, we consider N massive particles in one dimension with positions x_α and momenta p_α that move according to the Hamilton equations

$$\dot{x}_\alpha(t) = \frac{\partial\mathcal{H}(\vec{x},\vec{p})}{\partial p_\alpha}, \tag{4}$$

$$\dot{p}_\alpha(t) = -\frac{\partial\mathcal{H}(\vec{x},\vec{p})}{\partial x_\alpha}, \tag{5}$$

where $\alpha = 1,\dots,N$ is an index that runs over all particles. The case of M interacting particles in three dimensions is described by $N = 3M$ particle coordinates and is implicitly included in our model. Using the antisymmetric matrix

$$U = \begin{pmatrix} 0 & 1 & 0 & 0 & \\ -1 & 0 & 0 & 0 & \\ 0 & 0 & 0 & 1 & \\ 0 & 0 & -1 & 0 & \\ & & & & \ddots \end{pmatrix} \tag{6}$$

and the state vector

$$\vec{z}(t) = (x_1(t), p_1(t), x_2(t), p_2(t)\dots)^T \tag{7}$$

the Hamilton equations can be written compactly as

$$\dot{z}_i(t) = U_{ij}\frac{\partial\mathcal{H}(\vec{z})}{\partial z_j}, \tag{8}$$

where $i = 1,\dots,2N$ is an index that runs over all position and momentum coordinates. Throughout this paper, greek indices denote particles (running from 1 to N), roman indices denote coordinates (running from 1 to $2N$), and indices that appear more than once are summed over except primed indices. We consider quadratic Hamiltonians of the form

$$\mathcal{H}(\vec{z}) = z_i H_{ij} z_j/2 \tag{9}$$

that are described by a general symmetric matrix H which we assume to be positive definite, i.e., $\mathcal{H}(\vec{z}) > 0$ for general $\vec{z}$. A possible linear term in the Hamiltonian can be absorbed into the definition of the state vector z_j and need not be considered explicitly. For quadratic Hamiltonians the Hamilton equations are linear and given by

$$\dot{z}_i(t) = U_{ij} H_{jk} z_k(t). \tag{10}$$

More specific forms of the Hamiltonian matrix H, in particular Newtonian Hamiltonians where momentum und position degrees of freedom are decoupled, will be discussed later, in the first part of this paper the discussion applies to general Hamiltonian models.

By adding linear friction terms and random fields to the Hamilton equations, which will be later shown to mimic the coupling to heat baths with in general different temperatures, we obtain the coupled linear Langevin equations

$$\dot{z}_i(t) = U_{ij} H_{jk} z_k(t) - \Gamma_{ik} z_k(t) + \Phi_{ik} F_k(t) \tag{11}$$

which by definition of the generally asymmetric coupling matrix

$$A_{ik} = -U_{ij} H_{jk} + \Gamma_{ik} \tag{12}$$

can be written more compactly as

$$\dot{z}_i(t) = -A_{ik} z_k(t) + \Phi_{ik} F_k(t). \tag{13}$$

Here Γ is the friction coefficient matrix and Φ is the random strength matrix that describes how random fields couple to different particle coordinates. For simplicity, we assume Gaussian white random fields with zero mean $\langle F_i(t)\rangle = 0$ and and variances $\langle F_i(t) F_j(t')\rangle = 2\delta_{ij}\delta(t-t')$. Non-Markovian models with colored noise can be obtained by integrating out degrees of freedom and need not explicitly be considered [85]. In standard friction models friction forces couple to the momentum degree of freedom and are proportional to particle velocities. In more elaborate models that include hydrodynamic interactions, the friction force acting on a given particle depends on the velocities of all particles. In the first part of this paper the matrices Γ and Φ are kept general, and Γ can also be asymmetric (which will be shown to have direct consequences for the detailed-balance condition in Sec. III D). In the second part of the paper, starting in Sec. IV B, we model heat baths with different temperatures and for this will assume Γ and Φ to be diagonal. The general form of the linear Langevin Eq. (13) does not directly reveal whether it describes an equilibrium or a NEQ system [70]; this point will be addressed further below.

B. Stationary distribution

The algebraic solution of the Langevin Eq. (13) is

$$z_i(t) = e^{-tA}_{ij} z_j(0) + \int_0^t dt' e^{-(t-t')A}_{ij} \Phi_{jk} F_k(t'), \tag{14}$$

where $z_j(0)$ denotes the initial particle positions and momenta at time zero. The average over the noise gives

$$\langle z_i(t)\rangle = e^{-tA}_{ij} z_j(0), \tag{15}$$

which can be viewed as the solution of the noise-averaged version of the Langevin Eq. (13)

$$\langle \dot{z}_i(t)\rangle = -A_{ik}\langle z_k(t)\rangle. \tag{16}$$

If all eigenvalues of the matrix A have positive real components, a unique stationary distribution exists and is characterized by a vanishing mean $\langle z_i(t\to\infty)\rangle = 0$.

The covariance matrix of the deviations from the mean $\Delta z_i(t) = z_i(t) - \langle z_i(t)\rangle$ follows from squaring the solution Eq. (14) and averaging over the noise, leading to [86]

$$E_{ij}(t) \equiv \langle \Delta z_i(t)\Delta z_j(t)\rangle = 2\int_0^t dt' e^{-(t-t')A}_{ik} e^{-(t-t')A}_{jl} C_{kl}, \tag{17}$$

where the random correlation matrix is defined by

$$C_{kl} = \Phi_{km}\Phi_{lm} = C_{lk} \tag{18}$$

and is symmetric by construction. Since

$$A_{ik}E_{kj} + A_{jk}E_{ki} = A_{ik}E_{kj} + A_{jk}E_{ik} = 2\int_0^t dt' \frac{d}{dt'} e^{-(t-t')A}_{ik} e^{-(t-t')A}_{jl} C_{kl}, \tag{19}$$

the stationary covariance matrix, denoted by an open circle and defined by

$$E^\circ_{ij} \equiv E_{ij}(t\to\infty), \tag{20}$$

is unique and given by the Lyapunov equation

$$2C_{ij} = A_{ik}E^\circ_{kj} + A_{jk}E^\circ_{ki} \tag{21}$$

if all eigenvalues of the matrix A have positive real parts, which we will assume to be true throughout this paper.

III. SYSTEMS THAT HAVE AN EQUILIBRIUM DISTRIBUTION

A. Distributions, entropy, and free energy

In equilibrium, the normalized distribution in terms of the state vector $\vec{z}$ is given by the Boltzmann distribution

$$\rho(\vec{z}) = e^{-\beta\mathcal{H}(\vec{z})}/\mathcal{Z}, \tag{22}$$

where $\beta = 1/(k_BT)$ denotes the inverse thermal energy and $\mathcal{Z} = \int d\vec{z} e^{-\beta\mathcal{H}(\vec{z})}$ is the partition function. We will discuss the connection of this definition of equilibrium to the conditions of detailed balance, vanishing probability current as well as the fluctuation-dissipation theorem in Sec. III D. Positive definiteness of the Hamiltonian matrix H guarantees that $\mathcal{Z}$ is finite (if the Hamiltonian is invariant with respect to one or few degrees of freedom they can be separated off to make the reduced Hamiltonian positive definite). For a quadratic Hamiltonian, the average state vector vanishes and all covariances can be calculated from the Boltzmann distribution, which only involves inversion of the Hamiltonian matrix.

For the later discussion of the NEQ scenario, it is instructive to derive the equilibrium distribution also via the thermodynamic route. From the thermodynamic definitions of the free energy $\mathcal{F} = -k_BT\ln\mathcal{Z}$ and the entropy $\mathcal{S} = -\partial\mathcal{F}/\partial T$, the Shannon expression for the entropy directly follows as

$$\mathcal{S}/k_B = -\int d\vec{z}\rho(\vec{z})\ln\rho(\vec{z}); \tag{23}$$

the derivation is shown in Appendix A. Note that the Shannon expression Eq. (23) can also be used to describe the distribution entropy for time-dependent distributions, i.e., for nonstationary and even NEQ situations, since the expression makes no reference to the equilibrium ensemble or to the presence of a heat bath. This will allow us to describe the time-dependent approach to equilibrium as well as to stationary NEQ distributions.

For the linear Langevin Eq. (13), the time-dependent probability distribution is Gaussian and can be written as

$$\rho(\vec{z},t) = \mathcal{N}^{-1}(t)\exp\{-[z_i - \langle z_i(t)\rangle] \times E_{ij}^{-1}(t)[z_j - \langle z_j(t)\rangle]/2\}, \tag{24}$$

where the time-dependent normalization constant is given by

$$\mathcal{N}(t) = \sqrt{(2\pi)^{2N}\det E(t)} \tag{25}$$

and exists only if the covariance matrix $E(t)$ is positive definite [note that by its definition $E(t)$ is also symmetric]. The proof is standard textbook material [86], in Appendix B we present a derivation based on random-field path integrals, which has the advantage that it can in principle be generalized to non-Gaussian colored noise; there we furthermore demonstrate that the expression (24) is in fact the Green's function of the general Langevin Eq. (13), i.e., the conditional probability distribution at time t for the case that the distribution is a delta function at time $t = 0$.

With the Gaussian form (24), the integral in Eq. (23) can be performed and yields the time-dependent distribution entropy as

$$\mathcal{S}(t)/k_B = N + \ln\mathcal{N}(t) = N + (1/2)\ln[(2\pi)^{2N}\det E(t)]. \tag{26}$$

The internal energy is given by

$$\mathcal{U}(t) = \langle\mathcal{H}[\vec{z}(t)]\rangle = H_{ij}\langle z_i(t)z_j(t)\rangle/2 = H_{ij}\langle z_i(t)\rangle\langle z_j(t)\rangle/2 + H_{ij}E_{ij}(t)/2. \tag{27}$$

With these results for the entropy and internal energy, the free energy

$$\mathcal{F}(t) = \mathcal{U}(t) - T\mathcal{S}(t) \tag{28}$$

follows as

$$\mathcal{F}(t) = H_{ij}\langle z_i(t)\rangle\langle z_j(t)\rangle/2 + H_{ij}E_{ij}(t)/2 - k_BTN - (k_BT/2)\ln[(2\pi)^{2N}\det E(t)]. \tag{29}$$

The extremum of the free energy is determined by $\langle\vec{z}(t)\rangle = 0$ and by

$$\frac{\partial\mathcal{F}(t)}{\partial E_{ij}} = H_{ij}/2 - k_BTE_{ji}^{-1}(t)/2 = 0, \tag{30}$$

the solution of which is time-independent and defines the equilibrium distribution (denoted by a filled circle) as

$$E_{ij}^{\bullet} = k_BTH_{ij}^{-1}. \tag{31}$$

In deriving Eq. (30) we used the basic algebraic relation $\partial\ln\det E/\partial E_{ij} = E_{ji}^{-1}$. The partial derivative denotes the derivative with respect to one matrix component while keeping all other components fixed. The equilibrium free energy follows by reinserting $\langle\vec{z}(t)\rangle = 0$ and the solution $E_{ij}^{\bullet}$ into the free-energy expression (29) and is given by

$$\mathcal{F}^{\bullet} = -(k_BT/2)\ln[(2\pi k_BT)^{2N}/\det H]. \tag{32}$$

We next want to show that the extremum of the free energy is in fact a minimum [for this we neglect the trivial quadratic dependence of Eq. (29) on the mean state vector $\langle\vec{z}(t)\rangle$]. We first realize that

$$\frac{\partial^2\mathcal{F}(t)}{\partial E_{kl}\partial E_{ij}} = k_BTE_{ik}^{-1}(t)E_{lj}^{-1}(t)/2, \tag{33}$$

where we used the basic algebraic relation $\partial E_{kn}^{-1}/\partial E_{ij} = -E_{ki}^{-1}E_{jn}^{-1}$. Around the equilibrium distribution $E_{ij} = E_{ij}^{\bullet}$ the free energy is to second order given by

$$\mathcal{F}(t) - \mathcal{F}^{\bullet} \simeq H_{ki}\big[E_{ij}(t) - H_{ij}^{-1}k_BT\big] \times H_{jl}\big[E_{lk}(t) - H_{lk}^{-1}k_BT\big]/(2k_BT), \tag{34}$$

which can be rewritten as

$$\mathcal{F}(t) - \mathcal{F}^{\bullet} \simeq \frac{k_BT[\delta_{il} - \beta H_{ik}E_{kl}(t)][\delta_{il} - \beta H_{lk}E_{ki}(t)]}{2}. \tag{35}$$

The latter form is quadratic and of the general form $\mathcal{F}(t) - \mathcal{F}^{\bullet} \simeq k_BTB_{il}B_{li}/2$ with $B_{il} \equiv \delta_{il} - \beta H_{ik}E_{kl}(t)$, but this by itself does not guarantee that $\mathcal{F}(t) - \mathcal{F}^{\bullet}$ is positive since B_{il} is not necessarily symmetric. In Appendix C we show by diagonalization that the positivity of the expression (34) for $\mathcal{F}(t) - \mathcal{F}^{\bullet}$ follows from the fact that H is symmetric and positive definite.

The Gaussian distribution (24) in conjunction with Eq. (31) is equivalent to the Boltzmann distribution (22), which we have thus rederived by minimizing the time-dependent free-energy functional (29). But the free-energy functional (29) is not only valid in equilibrium but also describes systems that approach the equilibrium distribution. This is an important insight, as this functional framework will allow us to characterize the approach not only to equilibrium but also to stationary NEQ distributions.

B. When does a Langevin equation describe an equilibrium system?

In this section we will explore under which conditions the Langevin Eq. (13) describes an equilibrium system, which will put stringent conditions on the random correlation matrix C and on the friction matrix Γ. We in this paper define a system to be in equilibrium if the stationary state corresponds to the Boltzmann distribution, the relation to other definitions of equilibrium will be discussed in Sec. III D. We implement this condition by replacing the stationary covariance matrix E_{ij}° in the Lyapunov Eq. (21) by the equilibrium covariance matrix $E_{ij}^{\bullet}$ from Eq. (31), by which we obtain

$$2C_{ij}^{\bullet}/(k_BT) = A_{ik}H_{kj}^{-1} + A_{jk}H_{ki}^{-1}. \tag{36}$$

Inserting the expression for the Langevin matrix A from Eq. (12) and using the fact that the matrix U is antisymmetric [see Eq. (6)], we arrive at the Lyapunov-Boltzmann equation

$$2C_{ij}^{\bullet}/(k_BT) = \Gamma_{ik}H_{kj}^{-1} + \Gamma_{jk}H_{ki}^{-1}. \tag{37}$$

If for arbitrary Hamiltonian matrix H the friction matrix Γ and the random force correlation matrix C obey this equation, the Langevin equation given by Eq. (13) describes the dynamics of an equilibrium system. Conversely, if C and Γ do not satisfy Eq. (37) the Langevin equation describes a NEQ system. Two obvious NEQ scenarios come to mind: (1) a Newtonian Hamiltonian many-body system coupled to heat baths characterized by different temperatures (as will be discussed starting in Sec. IV B), and (2) a many-body system with off-diagonal friction terms that do not obey Eq. (37).

Let us give a simple example, namely, the harmonic oscillator with Hamiltonian $\mathcal{H} = Kx^2/2 + p^2/(2m)$, which is described by the Hamiltonian matrix

$$H = \begin{pmatrix} K & 0 \\ 0 & 1/m \end{pmatrix}. \tag{38}$$

We choose a diagonal momentum friction model described by the friction matrix

$$\Gamma = \begin{pmatrix} 0 & 0 \\ 0 & \gamma/m \end{pmatrix}, \tag{39}$$

where the friction force is proportional to the velocity of the particle and enters only the momentum degree of freedom. The matrix product appearing on the right side of the Lyapunov-Boltzmann Eq. (37) is given by

$$\Gamma H^{-1} = \begin{pmatrix} 0 & 0 \\ 0 & \gamma \end{pmatrix} \tag{40}$$

and thus the equilibrium random correlation matrix follows as

$$C^{\bullet} = \begin{pmatrix} 0 & 0 \\ 0 & k_BT\gamma \end{pmatrix}. \tag{41}$$

This agrees with the well-known result that the equilibrium Langevin equation of a massive particle involves a random field that acts on the momentum degree of freedom only and is proportional to the friction coefficient γ, where the proportionality constant k_BT defines the temperature of the reservoir. Conversely, any nontrivial deviation of the matrix C from Eq. (41), i.e., any deviation that cannot be captured by a modified temperature, indicates a NEQ system. For the simple example of a one-dimensional harmonic oscillator considered here, this could for example be the presence of an additional entry in the symmetric matrix C, e.g., additional entries in the off-diagonals or in the upper-left diagonal.

C. Dynamic approach to the stationary distribution

From the time-dependent analog of the Shannon entropy (23)

$$\mathcal{S}(t)/k_B = -\int d\vec{z}\rho(\vec{z},t)\ln\rho(\vec{z},t), \tag{42}$$

we obtain by differentiation

$$\dot{\mathcal{S}}(t)/k_B = -\int d\vec{z}\dot{\rho}(\vec{z},t)[\ln\rho(\vec{z},t)+1]. \tag{43}$$

The time derivative of the density distribution $\dot{\rho}(\vec{z},t)$ is determined by the Fokker-Planck equation

$$\dot{\rho}(\vec{z},t) = [\vec{\nabla}_k A_{km} z_m + \vec{\nabla}_k\vec{\nabla}_m C_{km}]\rho(\vec{z},t), \tag{44}$$

which follows via Kramers-Moyal expansion of the Langevin Eq. (13) [86]. With the Gaussian time-dependent distribution Eq. (24) we obtain from Eq. (44) the expression

$$\dot{\rho}(\vec{z},t) = \rho(\vec{z},t)\times\big[A_{kk} - A_{km}z_m E^{-1}_{kj}\Delta z_j + C_{ij}\big(E^{-1}_{ik}\Delta z_k E^{-1}_{jl}\Delta z_l - E^{-1}_{ij}\big)\big]. \tag{45}$$

Inserting this into the expression (43) and calculating all Gaussian expectation values, we obtain the final expression for the time derivative of the entropy as

$$\dot{\mathcal{S}}(t)/k_B = C_{km}E^{-1}_{km}(t) - A_{kk}, \tag{46}$$

which holds for equilibrium as well as for NEQ systems. From the Lyapunov Eq. (21) we can derive the expression

$$C_{ij}E^{\circ-1}_{ij} = A_{kk}, \tag{47}$$

inserting this into Eq. (46) we obtain the alternative expression

$$\dot{\mathcal{S}}(t)/k_B = C_{km}\big[E^{-1}_{km}(t) - E^{\circ-1}_{km}\big], \tag{48}$$

which demonstrates that the entropy change vanishes in the stationary state $E_{km}(t) = E^{\circ}_{km}$, as is expected. We will later come back to Eq. (48) as it allows us to write one of the constitutive dynamic equations for NEQ systems.

We next calculate the time derivative of the covariance matrix. From the expression (45) we immediately read off that

$$\partial\ln\rho(\vec{z},t)/\partial t = A_{kk} - A_{km}z_m E^{-1}_{kj}\Delta z_j + C_{ij}\big(E^{-1}_{ik}\Delta z_k E^{-1}_{jl}\Delta z_l - E^{-1}_{ij}\big), \tag{49}$$

which can be rewritten as

$$\partial\ln\rho(\vec{z},t)/\partial t = \Delta z_i\Delta z_j\big[C_{km}E^{-1}_{kj}E^{-1}_{mi} - A_{ki}E^{-1}_{kj}\big] + A_{kk} - C_{km}E^{-1}_{km} - A_{km}\langle z_m(t)\rangle E^{-1}_{kj}\Delta z_j. \tag{50}$$

On the other hand, using the definition of the Gaussian distribution (24) we find

$$\frac{\partial\ln\rho(\vec{z},t)}{\partial t} = \frac{\partial}{\partial t}\bigg\{-\frac{1}{2}\ln[(2\pi)^{2N}\det E(t)] - \frac{1}{2}[z_i - \langle z_i(t)\rangle]E^{-1}_{ij}(t)[z_j - \langle z_j(t)\rangle]\bigg\}, \tag{51}$$

which can be rewritten as

$$\frac{\partial\ln\rho(\vec{z},t)}{\partial t} = \langle\dot{z}_i(t)\rangle E^{-1}_{ij}(t)\Delta z_j + \frac{\dot{E}^{-1}_{kl}\partial}{\partial E^{-1}_{kl}}\bigg\{\frac{1}{2}\ln[(2\pi)^{-2N}\det E^{-1}(t)] - \frac{1}{2}\Delta z_i E^{-1}_{ij}(t)\Delta z_j\bigg\} \tag{52}$$

and finally yields

$$\frac{\partial\ln\rho(\vec{z},t)}{\partial t} = \langle\dot{z}_i(t)\rangle E^{-1}_{ij}\Delta z_j + \frac{\dot{E}^{-1}_{kl}}{2}\{E_{kl}(t) - \Delta z_k\Delta z_l\}. \tag{53}$$

Comparison of Eqs. (50) and (53) term by term yields

$$\dot{E}^{-1}_{ij}(t) = -2C_{km}E^{-1}_{kj}(t)E^{-1}_{mi}(t) + E^{-1}_{jk}(t)A_{ki} + E^{-1}_{ik}(t)A_{kj}, \tag{54}$$

where we have used Eq. (16). From the basic algebraic relation $\dot{E}_{il}(t) = -E_{ij}(t)\dot{E}^{-1}_{jk}(t)E^{-1}_{kl}(t)$ we finally obtain from Eq. (54) the temporal change of the covariance matrix as

$$\dot{E}_{ij}(t) = 2C_{ij} - A_{ik}E_{kj}(t) - A_{jk}E_{ki}(t), \tag{55}$$

which, using Eq. (21), can be rewritten as

$$\dot{E}_{ij}(t) = -A_{ik}\big[E_{kj}(t) - E^{\circ}_{kj}\big] - A_{jk}\big[E_{ki}(t) - E^{\circ}_{ki}\big]. \tag{56}$$

Note that this expression holds for equilibrium as well as for NEQ systems. As would be expected, the temporal change of the covariance matrix vanishes in the stationary state, i.e., when $E_{km}(t) = E^{\circ}_{km}$.

D. Conditions of detailed balance and vanishing probability current

Our definition of equilibrium employs the Boltzmann distribution (22) and not the condition of detailed balance, which is often used as the defining property of equilibrium [87]. The reason for using the Boltzmann condition is that it is very easy to implement, while the condition of detailed balance is for underdamped many-particle systems rather involved. In fact, for overdamped systems the detailed-balance condition becomes equivalent to the condition of vanishing probability current [5,8], which in literature is also called the potential condition. We will in this section formulate the condition for the probability current to vanish and then compare with the detailed-balance condition, for which the derivation is presented in Appendix D.

The Fokker-Planck Eq. (44) can be interpreted as a balance equation

$$\dot{\rho}(\vec{z}, t) = -\vec{\nabla}_k \mathcal{J}_k(\vec{z}, t) \tag{57}$$

with the probability current $\mathcal{J}(\vec{z}, t)$ being given as

$$\mathcal{J}_k(\vec{z}, t) = -[A_{km}z_m + \vec{\nabla}_m C_{km}]\rho(\vec{z}, t). \tag{58}$$

For the Gaussian distribution (24) we obtain [for simplicity we set $\langle z_i(t)\rangle = 0$ here] for the current

$$\mathcal{J}_k(\vec{z}, t) = -\big[A_{km} - C_{kj}E^{-1}_{mj}\big]z_m\rho(\vec{z}, t). \tag{59}$$

The probability current vanishes, $\mathcal{J}_k(\vec{z}, t) = 0$, for

$$A_{km}E_{mj} = C_{kj}. \tag{60}$$

From the equilibrium condition $E_{kj}(t) = E^{\bullet}_{kj} = k_BTH^{-1}_{kj}$, Eq. (31), and the explicit form of the matrix A in Eq. (12), we obtain the vanishing probability current condition

$$\Gamma_{km}H^{-1}_{mj} - U_{kj} = C_{kj}/(k_BT), \tag{61}$$

which in the general case is not satisfied in the equilibrium situation defined by the Lyapunov-Boltzmann Eq. (37), since C is a symmetric matrix, while U is antisymmetric and typically $\Gamma_{km}H^{-1}_{mj}$ is an asymmetric matrix. We conclude that for an underdamped system, the probability generally does not vanish, this is trivially illustrated by the fact that a harmonic oscillator performs orbits in phase space. In fact, current mathematical work is devoted to separating phase space trajectories of underdamped systems into periodic and diffusive parts [82,83]. In Appendix E we show that in the overdamped (i.e., massless) limit the probability current in equilibrium vanishes if the friction matrix Γ is symmetric. Since we did not impose that Γ is symmetric so far, we see that our definition of equilibrium, which is based on the stationary distribution being equal to the Boltzmann distribution, is for overdamped systems only equivalent to the vanishing probability current condition for a symmetric friction matrix. We conclude that the vanishing probability current condition is even in the overdamped limit a more restrictive criterion than our Boltzmann distribution criterion.

Furthermore, in Appendix F we show for the special case of two coupled particles, that the equilibrium fluctuation-dissipation theorem holds if our Boltzmann definition for equilibrium is satisfied and in particular also works for an asymmetric friction matrix. This suggests that equilibrium definitions based on the Boltzmann distribution and based on the fluctuation-dissipation theorem are equivalent and that the condition of vanishing probability current is more restrictive and requires the symmetry of the friction matrix.

Finally, in Appendix D we derive the condition of detailed balance for our underdamped Hamiltonian model and demonstrate that it is equivalent to the Boltzmann-distribution based criterion for equilibrium only if the friction matrix is symmetric. Clearly, an asymmetric friction matrix breaks the physical principle of equal actio and reactio, so for physical models where friction is produced, e.g., by hydrodynamic interactions, the friction matrix should be symmetric and our definition of equilibrium (based on the Boltzmann distribution) is fully equivalent to the condition of detailed balance. More abstract models with asymmetric friction matrices are conceivable, for such models the distribution is predicted to be of the Boltzmann type if the Lyapunov-Boltzmann Eq. (37) is satisfied, yet the condition of detailed balance is violated.

E. Time-dependent free energy: Extremal and stability properties

The time derivative of the free-energy expression Eq. (28) reads

$$\dot{\mathcal{F}}(t) = H_{ij}\langle \dot{z}_i(t)\rangle\langle z_j(t)\rangle + H_{ij}\dot{E}_{ij}(t)/2 - k_BT\dot{\mathcal{S}}(t) \tag{62}$$

and using Eqs. (16), (46), and (55) is explicitly given by

$$\dot{\mathcal{F}}(t) = -\langle z_j(t)\rangle H_{ji}A_{ik}\langle z_k(t)\rangle + H_{ij}C_{ij} - H_{ij}A_{ik}E_{kj}(t) \\ - k_BTC_{km}E^{-1}_{km}(t) + k_BTA_{kk}. \tag{63}$$

After some manipulation this expression can be rewritten as

$$\dot{\mathcal{F}}(t) = -\langle z_j(t)\rangle H_{ji}\Gamma_{ik}\langle z_k(t)\rangle - C_{km}\big[H_{ml} - k_BTE^{-1}_{ml}(t)\big] \\ \times \frac{E_{lj}(t)}{k_BT}\big[H_{jk} - k_BTE^{-1}_{jk}(t)\big]. \tag{64}$$

The quadratic form of Eq. (64) directly demonstrates that for the equilibrium distribution, defined by $E_{kj}(t) = E^{\bullet}_{kj} = k_BTH^{-1}_{kj}$ and $\langle z_k(t)\rangle = 0$, the free energy is stationary and does not change in time, i.e., $\dot{\mathcal{F}}(t) = 0$. This is somewhat trivial since this just reflects that the equilibrium distribution is a special case of a stationary distribution. More importantly, from the fact that H, C and E are symmetric matrices that are positive-definite or semi-positive-definite, we derive in Appendix C that the free energy does not increase

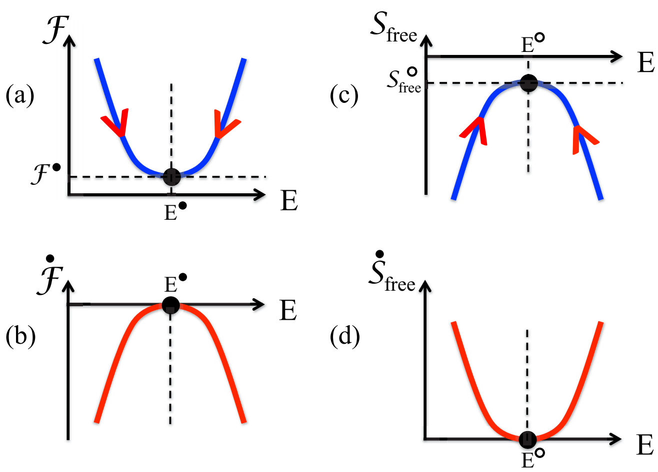

FIG. 1. Graphical illustration of extremal and stability properties of equilibrium and nonequilibrium (NEQ) systems. For an equilibrium system, the free energy in (a) takes its minimal value $\mathcal{F}^{\bullet}$ given by Eq. (32) at the equilibrium value of the covariance matrix $E^{\bullet}$, determined by Eq. (31). The free-energy production $\dot{\mathcal{F}}$ in (b) is negative and vanishes at the equilibrium state; see Eq. (65). This means that the system is stable with respect to perturbations and flows towards the equilibrium state, as indicated by red arrows in (a). For a NEQ system, the free entropy $\mathcal{S}_{\mathrm{free}}(t)$ in (c) takes its maximal value $\mathcal{S}^{\circ}_{\mathrm{free}}$ at the stationary value of the covariance matrix E°. The free entropy production $\dot{\mathcal{S}}_{\mathrm{free}}(t)$ in (d) is positive and vanishes in the stationary state. This means that the system is stable with respect to perturbations and flows towards the NEQ stationary state, as indicated by red arrows in (c).

in time,

$$\dot{\mathcal{F}}(t) \leqslant 0; \tag{65}$$

generally, this means that the equilibrium distribution is stable with respect to perturbations. For this we have used that the first term of Eq. (64) is not negative since the product of two positive semidefinite matrices is also positive semidefinite. Equation (65) corresponds to the second law of thermodynamics, here derived for many-body systems described by quadratic Hamiltonians and Langevin equations with friction terms and random fields. These results are graphically illustrated in Fig. 1. Figure 1(a) shows that the free energy $\mathcal{F}(t)$ is minimal in equilibrium defined by $E_{kj}(t) = E^{\bullet}_{kj}$ and $\langle z_k(t)\rangle = 0$, as demonstrated by Eq. (35). Figure 1(b) shows that the time derivative of the free energy $\dot{\mathcal{F}}(t)$ is negative except in equilibrium where it vanishes, as follows from Eq. (65). This indicates that a nonstationary distribution flows monotonically towards the equilibrium stationary distribution, as indicated by the red arrows in the top graph.

We finally derive a set of constitutive dynamic equations, which we will in the next section generalize for NEQ systems. For an equilibrium system the Lyapunov-Boltzmann condition (37) is satisfied and the stationary covariance matrix E°_{kj} is given by the equilibrium state, $E^{\circ}_{kj} = E^{\bullet}_{kj} = k_B T H^{-1}_{kj}$. Comparison of the expression for the entropy production (48) and the derivative of the free energy (30) yields

$$\dot{\mathcal{S}}(t) = -\frac{2}{k_B T} C_{ij} \frac{\partial \mathcal{F}(t)}{\partial E_{ij}}, \tag{66}$$

which is a simple relation between the entropy production $\dot{\mathcal{S}}$ and the free-energy derivative, the first constitutive dynamic relation. The other dynamic relations are obtained from the time derivative of the covariance matrix (56),

$$\dot{E}_{ij}(t) = -\frac{2}{k_B T}\left[A_{ik}E^{\bullet}_{kl}\frac{\partial \mathcal{F}(t)}{\partial E_{lm}}E_{jm} + A_{jk}E^{\bullet}_{kl}\frac{\partial \mathcal{F}(t)}{\partial E_{lm}}E_{mi}\right], \tag{67}$$

and by comparing Eq. (16) with the derivative of the free energy (29) with respect to $\langle z_i(t)\rangle$,

$$\langle \dot{z}_i(t)\rangle = -\frac{1}{k_B T} A_{ik} E^{\bullet}_{kl} \frac{\partial \mathcal{F}(t)}{\partial \langle z_l\rangle}. \tag{68}$$

Equations (66), (67), and (68) are the constitutive dynamic equations for equilibrium systems that relate temporal changes of all relevant quantities to derivatives of the free energy with respect to the state variables, i.e., to generalized thermodynamic forces.

IV. TRULY NONEQUILIBRIUM SYSTEMS

A. Extremal and stability properties of the free entropy

We remark that the word nonequilibrium (NEQ) typically refers to two very different situations: For a NEQ system, i.e., when the Lyapunov-Boltzmann condition (37) is not satisfied, the stationary distribution is a NEQ stationary distribution, which is characterized by a nonvanishing positive entropy production. For such a system an equilibrium distribution does not exist. An equilibrium system is one where the Lyapunov-Boltzmann condition (37) is satisfied, but even such a system is off equilibrium as long as it has not settled in its stationary equilibrium distribution.

Obviously, it is not possible to use the constitutive equilibrium dynamic equations (66), (67), and (68) in the NEQ case, because of the appearance of the equilibrium temperature T. The temperature can be trivially eliminated by introducing a modified thermodynamic potential, which is called the free entropy and which is, for the equilibrium scenario using Eqs. (27), (28), and (31), given by

$$\frac{\mathcal{S}_{\mathrm{free}}(t)}{k_B} \equiv -\frac{\mathcal{F}(t)}{k_B T} = \frac{\mathcal{S}(t)}{k_B} - \frac{E^{\bullet -1}_{ij}E_{ij}(t)}{2} - \frac{E^{\bullet -1}_{ij}\langle z_i(t)\rangle\langle z_j(t)\rangle}{2}. \tag{69}$$

As mentioned before, the equilibrium free entropy had been originally introduced by Massieu in 1869 [72,73] and the advantage of this functional was already recognized by Planck [74] and Schrödinger [75]. The free entropy concept is particularly useful in the current NEQ setting, since for the NEQ systems we will consider there is no unique temperature.

In fact, the free entropy concept allows for straightforward generalization to the NEQ case. For this we replace $E^{\bullet -1}_{ij}$ in Eq. (69) by $E^{\circ -1}_{ij}$, after which we obtain the NEQ version of the free entropy

$$\frac{\mathcal{S}_{\mathrm{free}}(t)}{k_B} = \frac{\mathcal{S}(t)}{k_B} - \frac{E^{\circ -1}_{ij}E_{ij}(t)}{2} - \frac{E^{\circ -1}_{ij}\langle z_i(t)\rangle\langle z_j(t)\rangle}{2}. \tag{70}$$

Note that the entropy production in the stationary NEQ state, which for suitably chosen friction and random strength matrices can be described as being due to heat fluxes in and out of heat reservoirs, is, as illustrated in Eq. (1), not included in the free entropy, but will be discussed in Sec. IV B. Clearly, for an equilibrium system, for which $E_{ij}^{\circ -1} = E_{ij}^{\bullet -1} = H_{ij}/(k_B T)$, the general expression Eq. (70) reduces to the equilibrium expression (69). Similar functionals, which in a NEQ stationary state are extremal, were derived previously [5–8,11,79–83]; in Appendix G we show that the free entropy expression (70) is equivalent to the Kullback-Leibler entropy [6,78]. Our model differs from previous approaches in that the definition of separate friction and random strength matrices will allow us to describe the coupling to multiple heat baths with different well-defined temperatures.

Using the free entropy, the constitutive dynamic equations (66), (67), and (68) can be rewritten as

$$\dot{\mathcal{S}}(t) = 2C_{ij}\frac{\partial \mathcal{S}_{\text{free}}(t)/k_B}{\partial E_{ij}}, \tag{71}$$

$$\dot{E}_{ij}(t) = 2A_{ik}E^{\circ}_{kl}\frac{\partial \mathcal{S}_{\text{free}}(t)/k_B}{\partial E_{lm}}E_{jm} + 2A_{jk}E^{\circ}_{kl}\frac{\partial \mathcal{S}_{\text{free}}(t)/k_B}{\partial E_{lm}}E_{mi}, \tag{72}$$

$$\langle \dot{z}_i(t)\rangle = A_{ik}E^{\circ}_{kl}\frac{\partial \mathcal{S}_{\text{free}}(t)/k_B}{\partial \langle z_l\rangle}, \tag{73}$$

where the temperature has obviously (and trivially) disappeared. When Eq. (70) is used in conjunction with the constitutive dynamic equations (71), (72), and (73), the exact dynamic evolution equations (derived for general NEQ systems) (48), (56), and (16) are reproduced, this independently confirms the validity of the expression (70).

We next show that the free entropy has very similar properties for NEQ systems as the free energy has for equilibrium systems, namely, $\mathcal{S}_{\text{free}}(t)$ is extremal and in fact maximal in the stationary NEQ state and the stationary state is stable in the sense that $\dot{\mathcal{S}}_{\text{free}}(t) \geqslant 0$. For this we basically repeat the steps leading to the minimal condition of the free energy (35). The extremum of $\mathcal{S}_{\text{free}}(t)$ is given by $\langle \vec{z}(t)\rangle = 0$ and determined by

$$\frac{\partial \mathcal{S}_{\text{free}}(t)/k_B}{\partial E_{ij}} = E^{-1}_{ji}(t)/2 - E^{\circ -1}_{ji}/2 = 0, \tag{74}$$

the solution of which yields the time-independent stationary distribution, $E_{ij}(t) = E^{\circ}_{ij}$. With the stationary covariance matrix E°_{ij} many observables can be calculated, for example, the internal energy follows according to Eq. (27). The stationary free entropy is given by

$$\mathcal{S}^{\circ}_{\text{free}}/k_B = (1/2)\ln[(2\pi)^{2N}\det E^{\circ}]. \tag{75}$$

The second derivative of the free entropy is given by

$$\frac{\partial^2 \mathcal{S}_{\text{free}}(t)/k_B}{\partial E_{kl}\partial E_{ij}} = -E^{-1}_{ik}(t)E^{-1}_{lj}(t)/2. \tag{76}$$

Around the stationary state $E_{ij} = E^{\circ}_{ij}$ the free entropy thus is to second order given by

$$\mathcal{S}_{\text{free}}(t)/k_B - \mathcal{S}^{\circ}_{\text{free}}/k_B \simeq -E^{\circ -1}_{ik}E^{\circ -1}_{lj}[E_{ij}(t) - E^{\circ}_{ij}k_BT] \times [E_{kl}(t) - E^{\circ}_{kl}k_BT]/2, \tag{77}$$

which is negative since $E^{\circ -1}$ is symmetric and negative definite (the general proof for this is given in Appendix C).

The free entropy production follows from Eq. (70) by taking a time derivative as

$$\frac{\dot{\mathcal{S}}_{\text{free}}(t)}{k_B} = \frac{\dot{\mathcal{S}}(t)}{k_B} - \frac{E^{\circ -1}_{ij}\dot{E}_{ij}(t)}{2} - E^{\circ -1}_{ij}\langle \dot{z}_i(t)\rangle\langle z_j(t)\rangle. \tag{78}$$

Using our previous results for $\dot{\mathcal{S}}(t)$, Eq. (48), for $\dot{E}_{ij}(t)$, Eq. (56), and for $\langle \dot{z}_i(t)\rangle$, Eq. (16), we arrive after a few intermediate steps at

$$\dot{\mathcal{S}}_{\text{free}}(t)/k_B = \langle z_j(t)\rangle E^{\circ -1}_{ji}A_{ik}\langle z_k(t)\rangle + C_{km}\big[E^{\circ -1}_{ml} - E^{-1}_{ml}(t)\big] \times E_{lj}(t)\big[E^{\circ -1}_{jk} - E^{-1}_{jk}(t)\big]. \tag{79}$$

The quadratic form of this expression shows that in the stationary state, defined by $E_{kj}(t) = E^{\circ}_{kj}$ and $\langle z_k(t)\rangle = 0$, the free entropy production of the system vanishes. More importantly, from the fact that E°, C, and E are symmetric matrices that are semi-positive-definite or positive-definite, it follows that the free entropy does not decrease in time, i.e.,

$$\dot{\mathcal{S}}_{\text{free}}(t) \geqslant 0; \tag{80}$$

generally, this means that the stationary NEQ distribution is stable with respect to perturbations; the proof is given in Appendix C. In writing Eq. (80) we have used that the first term of Eq. (79) is not negative since the product of the two positive-definite matrices A and E° is also positive-definite, where positive definiteness of the asymmetric matrix A is equivalent to demanding that all eigenvalues of the symmetric part of A are positive, which is more restrictive than the condition that the eigenvalues of A have all positive real components, as required for the existence of a stationary state in Sec. II B. This result is here derived for interacting many body systems described by quadratic Hamiltonians and is graphically illustrated in Fig. 1. Figure 1(c) shows that the free entropy $\mathcal{S}_{\text{free}}(t)$ is maximal at the stationary distribution defined by $E_{kj}(t) = E^{\circ}_{kj}$ and $\langle z_k(t)\rangle = 0$, which follows from Eq. (77). Figure 1(d) shows that the time derivative of the free entropy $\dot{\mathcal{S}}_{\text{free}}(t)$ is positive except at the stationary distribution where it vanishes, as follows from Eq. (79). This indicates that the system flows monotonically towards the stationary state, as indicated by the red arrows in Fig. 1(c).

The free entropy maximization principle applies to NEQ and equilibrium systems alike. On the one hand this allows us to treat NEQ and equilibrium systems within a unified framework and thereby eliminates a disturbing schism in the description of these systems. On the other hand, the usage of the free entropy, which does not include the stationary reservoir entropy production [which is related to the so-called house-keeping entropy [9,10,84] and increases linearly in time in a stationary NEQ state; see Eq. (1) and as discussed in the next section], brings a significant methodological advantage over early approaches to NEQ systems, according to which stationary NEQ states are defined by an extremum of the total entropy production (including the reservoirs) and stability criteria invoke the time derivative of the entropy production, i.e., the second time derivative of the total entropy [87]. The connection between the free entropy production (excluding the reservoirs) and the total entropy production (including the reservoirs) is discussed in the Conclusions and in Appendix H.

B. Stationary entropy production for particles coupled to different temperature reservoirs

The calculations so far were completely general and no restrictions on the type of the Hamiltonian matrix H, the friction matrix Γ, and the random correlation matrix C were imposed. To gain insight into the entropy production of the reservoirs, denoted by $\dot{\mathcal{S}}_{\rm res}(t)$, we need to restrict the discussion to Newtonian systems, meaning that in the Hamiltonian the spatial and momentum coordinates decouple and the kinetic energy is diagonal. This will allow us to ascribe well-defined temperatures to different heat baths, which will then be used to calculate the reservoir entropy production from the individual heat fluxes between the system and the heat baths. To conveniently use the symmetry of such Newtonian systems, we switch to particle indices, denoted by greek symbols, with which the Hamiltonian can be written as

$$\mathcal{H}(\vec{y}) = y_\alpha H_{\alpha\epsilon} y_\epsilon/2. \tag{81}$$

Here we introduced the particle state vector

$$y_\alpha(t) = (x_\alpha(t), p_\alpha(t))^T, \tag{82}$$

where $x_\alpha(t)$ and $p_\alpha(t)$ are the position and the momentum of particle α. The $N \times N$ entries of the Hamiltonian matrix $H_{\alpha\epsilon}$ each consist of 2×2 matrices. These submatrices are expanded in terms of the matrices

$$u = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad s = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad r = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}. \tag{83}$$

From the matrix products

$$us = ru = \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix}, \quad ur = su = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \tag{84}$$

it is easily seen that all 2×2 matrices can be expanded in terms of the four matrices s, r, us, and ur, which thus form a convenient complete set. We will make in the following repeated use of the properties:

$$s^2 = s, \quad r^2 = r, \quad u^2 = -\mathbb{1}, \quad rs = \mathbb{0} = sr. \tag{85}$$

The Langevin Eq. (13) can be written as

$$\dot{y}_\alpha(t) = -A_{\alpha\epsilon} y_\epsilon(t) + \Phi_{\alpha\epsilon} F_\epsilon(t), \tag{86}$$

where $A_{\alpha\epsilon}$ is given by

$$A_{\alpha\epsilon} = -U_{\alpha\gamma} H_{\gamma\epsilon} + \Gamma_{\alpha\epsilon} \tag{87}$$

and $U_{\alpha\gamma}$ can be written as $U_{\alpha\gamma} = u\delta_{\alpha\gamma}$. The Hamiltonian matrix is for Newtonian systems given by

$$H_{\alpha'\epsilon} = s h_{\alpha'\epsilon} + r\delta_{\alpha'\epsilon}/m_{\alpha'}, \tag{88}$$

where $h_{\alpha\epsilon}$ is a general $N \times N$ symmetric interaction matrix that acts only on positional degrees of freedom (and thus is multiplied by the matrix s) and the second term is the kinetic energy which is diagonal in the momentum degrees of freedom (and thus is multiplied by the matrix r) and $m_{\alpha'}$ is the mass of particle α'. Note that the primed index α' is not summed over. From the product properties of the 2×2 matrices the inverse Hamiltonian matrix follows as

$$H^{-1}_{\alpha'\epsilon} = s h^{-1}_{\alpha'\epsilon} + r\delta_{\alpha'\epsilon} m_{\alpha'}. \tag{89}$$

To allow for a clear definition of reservoir temperatures, we will in the remainder treat momentum-diagonal friction, for which the friction matrix Γ is reduced to the momenta entries as

$$\Gamma_{\alpha'\epsilon} = r\gamma_{\alpha'\epsilon}/m_{\alpha'} \tag{90}$$

and where the matrix $\gamma_{\alpha'\epsilon}$ is diagonal and given by $\gamma_{\alpha'\epsilon} = \delta_{\alpha'\epsilon}\gamma_{\alpha'}$, where $\gamma_{\alpha'}$ is the friction coefficient of particle α'. The Lyapunov-Boltzmann condition (37) in terms of particle indices reads

$$2C^\bullet_{\alpha\epsilon}/(k_B T) = \Gamma_{\alpha\gamma} H^{-1}_{\gamma\epsilon} + \Gamma_{\epsilon\gamma} H^{-1}_{\gamma\alpha}, \tag{91}$$

which, using Eqs. (89) and (90) and the matrix product properties (85), leads to

$$C^\bullet_{\alpha'\epsilon} = r\delta_{\alpha'\epsilon}\gamma_{\alpha'} k_B T, \qquad \Phi^\bullet_{\alpha'\epsilon} = r\phi^\bullet_{\alpha'\epsilon} = r\delta_{\alpha'\epsilon}\sqrt{\gamma_{\alpha'} k_B T}. \tag{92}$$

These expressions show that the random correlation and random strength matrices are diagonal in the particle indices and proportional to r and thus only couple momentum degrees to each other. The NEQ generalization of Eq. (92) for reservoirs with different temperatures reads

$$C_{\alpha'\epsilon} = r\gamma_{\alpha'}\delta_{\alpha'\epsilon}/\beta_{\alpha'}, \quad \Phi_{\alpha'\epsilon} = r\delta_{\alpha'\epsilon}\sqrt{\gamma_{\alpha'}/\beta_{\alpha'}}, \tag{93}$$

where $\beta_{\alpha'} = 1/(k_B T_{\alpha'})$ denote the different inverse thermal energies of reservoirs, each characterized by a temperature $T_{\alpha'}$. The expressions (93) are central to our paper as they define the NEQ model we are using to derive all following results.

To calculate an explicit expression for the reservoir entropy production we multiply the Langevin Eq. (86) for the momentum component $p_\alpha(t)$ by $p_\alpha(t)$ and use Eqs. (87) and (88) to obtain

$$\begin{aligned} p_{\alpha'}(t)\dot{p}_{\alpha'}(t) = d\big[p^2_{\alpha'}(t)\big]/(2dt) = -p_{\alpha'}(t) h_{\alpha'\epsilon} x_\epsilon(t) \\ - p_{\alpha'}(t)\gamma_{\alpha'} p_{\alpha'}(t)/m_{\alpha'} + p_{\alpha'}(t)\Phi_{\alpha'\alpha'} F_{\alpha'}(t), \end{aligned} \tag{94}$$

where we again note that the primed index α' is not summed over. From this expression the average heating rate of the system due to reservoir α', i.e., the work performed by the random force per unit time, the last term in Eq. (94), minus the friction work dissipated by the particle α' per unit time, the second-last term in Eq. (94), turns out to be

$$\begin{aligned} \dot{Q}_{\alpha'}(t) &= \Phi_{\alpha'\alpha'}\langle p_{\alpha'}(t) F_{\alpha'}(t)\rangle/m_{\alpha'} - \gamma_{\alpha'}\langle p_{\alpha'}(t) p_{\alpha'}(t)\rangle/m^2_{\alpha'} \\ &= \frac{1}{2m_{\alpha'}} \frac{d\langle p^2_{\alpha'}(t)\rangle}{dt} + h_{\alpha'\epsilon}\langle p_{\alpha'}(t) x_\epsilon(t)\rangle/m_{\alpha'}. \end{aligned} \tag{95}$$

Since the last line of Eq. (95) is nothing but the time derivative of the sum of the kinetic and potential energies of particle α', we see that the reservoir heating rate $\dot{Q}_{\alpha'}(t)$ balances the particle energy at each instance of time (a clear consequence of the quadratic Hamiltonian approximation). For the entropy production of all reservoirs $\dot{\mathcal{S}}_{\rm res}(t)$ we thus obtain from Eq. (95) the expression

$$\frac{\dot{\mathcal{S}}_{\rm res}(t)}{k_B} \equiv -\beta_\alpha \dot{Q}_\alpha(t) = -\frac{\beta_\alpha}{2m_\alpha}\frac{d\langle p^2_\alpha(t)\rangle}{dt} - \frac{\beta_\alpha h_{\alpha\epsilon}}{m_\alpha}\langle p_\alpha x_\epsilon\rangle. \tag{96}$$

We now replace momenta by velocities $v_{\alpha'}(t) = p_{\alpha'}(t)/m_{\alpha'}$. From the fact that in the stationary state

$$\frac{d}{dt}\langle x_\alpha x_\epsilon\rangle^\circ = \langle x_\alpha v_\epsilon\rangle^\circ + \langle v_\alpha x_\epsilon\rangle^\circ = 0 \tag{97}$$

and $d\langle v_\alpha^2\rangle^\circ/dt = 0$ and using that $h_{\alpha\epsilon}$ is symmetric, we obtain for the reservoir entropy production in the stationary state

$$\dot{\mathcal{S}}^\circ_{\text{res}}/k_B = -\beta_\alpha h_{\alpha\epsilon}\langle x_\epsilon v_\alpha\rangle^\circ = \tfrac{1}{2}h_{\alpha\epsilon}\langle x_\epsilon v_\alpha\rangle^\circ(\beta_\epsilon - \beta_\alpha). \tag{98}$$

This expression shows that one necessary condition for a nonzero stationary reservoir entropy production is that reservoirs have different temperatures. Other necessary conditions are a nonvanishing stationary position-velocity coupling $\langle x_\epsilon v_\alpha\rangle^\circ$, which for Newtonian Hamiltonian systems is only obtained off equilibrium, and a nonvanishing interaction strength $h_{\alpha\epsilon}$. To obtain explicit results for the stationary reservoir entropy production we need to calculate the position-velocity correlations $\langle x_\epsilon v_\alpha\rangle^\circ$, which do not vanish even in the overdamped massless limit for a NEQ system. For this a systematic perturbative scheme is introduced in the next section.

C. Perturbative solution of the Lyapunov equation

The solution of the Lyapunov equation (21) for N particles described by a state vector with $2N$ components consists of determining all $2N(2N+1)/2$ entries of the symmetric covariance matrix E. The calculation is cumbersome even for only $N = 2$ particles [19]. Here we introduce a systematic expansion of the Lyapunov equation with the particles mass as the perturbation parameter, which employs a projection of the matrix equations onto the complete set of 2×2 matrices (83) and (84) introduced in the previous section. This expansion is subtle, since the leading-order result for the covariance matrix in the limit $m_\alpha \to 0$ is not obtained by taking this limit upfront in the Langevin equation. In fact, the overdamped limit is commonly obtained by setting $m_\alpha = 0$ in the Langevin equation [86], which correctly describes the long-time particle dynamics but obviously misses the short-time ballistic particle dynamics and leads to divergent instantaneous particle velocities. Since we need position-velocity correlations in order to estimate the entropy production according to Eq. (98), it is advisable to keep velocities to leading order as $m_\alpha \to 0$. As it turns out, position-velocity correlations that result from the perturbation calculation stay finite even in the $m_\alpha \to 0$ limit.

To proceed, we expand the covariance matrix as

$$E_{\alpha\beta} = \langle x_\alpha x_\beta\rangle s + \langle p_\alpha p_\beta\rangle r + \langle x_\alpha p_\beta\rangle ut + \langle p_\alpha x_\beta\rangle us \tag{99}$$

and insert the expressions for A, Eq. (87), H, Eq. (88), Γ, Eq. (90), C, Eq. (93), and E, Eq. (99), into the Lyapunov Eq. (21). The resulting expression splits into an equation proportional to s for $\alpha \neq \beta$,

$$0 = \langle x_\alpha v_\beta\rangle^\circ + \langle x_\beta v_\alpha\rangle^\circ, \tag{100}$$

an equation proportional to s for $\alpha = \beta$,

$$0 = \langle x_\alpha v_\alpha\rangle^\circ, \tag{101}$$

an equation proportional to r for $\alpha' \neq \beta'$,

$$\begin{aligned} 0 = {} & (\gamma_{\alpha'} m_{\beta'} + \gamma_{\beta'} m_{\alpha'})\langle v_{\alpha'} v_{\beta'}\rangle^\circ \\ & + m_{\beta'} h_{\alpha'\gamma}\langle x_\gamma v_{\beta'}\rangle^\circ + m_{\alpha'} h_{\beta'\gamma}\langle x_\gamma v_{\alpha'}\rangle^\circ, \end{aligned} \tag{102}$$

an equation proportional to r for $\alpha' = \beta'$,

$$1/\beta_{\alpha'} = m_{\alpha'}\langle v_{\alpha'} v_{\alpha'}\rangle^\circ + \boxed{m_{\alpha'} h_{\alpha'\gamma}\langle x_\gamma v_{\beta'}\rangle^\circ/\gamma_{\alpha'}}, \tag{103}$$

an equation proportional to ur for $\alpha' \neq \beta'$,

$$\boxed{m_{\beta'}\langle v_{\alpha'} v_{\beta'}\rangle^\circ} = \gamma_{\beta'}\langle x_{\alpha'} v_{\beta'}\rangle^\circ + h_{\beta'\gamma}\langle x_\gamma x_{\alpha'}\rangle^\circ, \tag{104}$$

and an equation proportional to ur for $\alpha' = \beta'$,

$$m_{\alpha'}\langle v_{\alpha'} v_{\alpha'}\rangle^\circ = h_{\alpha'\gamma}\langle x_\gamma x_{\alpha'}\rangle^\circ, \tag{105}$$

where we have converted all momenta $p_{\alpha'}$ to velocities $v_{\alpha'} = p_{\alpha'}/m_{\alpha'}$. Note that there are also two equations proportional to us which however are equivalent to the equations proportional to ur. The limit $m_{\alpha'} \to 0$ must be taken with care. Equation (105) shows that $\langle v_{\alpha'} v_{\alpha'}\rangle^\circ \sim m^{-1}$, which reflects the equipartition theorem in the equilibrium case, while Eq. (102) suggests that $\langle v_\alpha v_\beta\rangle^\circ \sim \langle x_\alpha v_\beta\rangle^\circ$ for $\alpha \neq \beta$. Together with Eq. (104), this suggests that $\langle v_\alpha v_\beta\rangle^\circ \sim \langle x_\alpha v_\beta\rangle^\circ \sim m^0$ for $\alpha \neq \beta$ and thus that these terms do not necessarily vanish in the massless limit $m_\alpha \to 0$. This in turn means that the terms in the boxes in Eqs. (103) and (104) can be treated perturbatively in an expansion in powers of m_α and can, to leading order in the particles masses, be neglected. Corrections to the leading-order results for the covariances can be systematically calculated by inserting the leading-order results for the terms in the boxes and by solving the resulting equations term-by-term in powers of the particle masses. Such a calculation of next-leading-order terms would also allow assessment of the accuracy of the leading-order results, but is rather involved because for NEQ systems mixed position-velocity cross-correlations are present, as we will show explicitly for the simple case of two coupled particles in the next section. As a main result, we conclude that while the mean-squared velocities of particles diverge in the overdamped limit, the position-velocity correlations between different particles take nonzero and finite values for NEQ systems in the overdamped limit.

V. APPLICATIONS

A. Two particles coupled to different temperature reservoirs: Effective temperature concept and position and momentum localization

We now present explicit results for two particles that are described by the Newtonian Hamiltonian as defined generally in Eq. (88),

$$\mathcal{H} = h_{11}x_1^2/2 + h_{22}x_2^2/2 + h_{12}x_1x_2 + m_1v_1^2/2 + m_2v_2^2/2, \tag{106}$$

and which are characterized by the two diagonal friction coefficients γ_1 and γ_2 as defined in Eq. (90). The particles are coupled to two heat reservoirs characterized by inverse thermal energies β_1 and β_2 as defined in Eq. (93). This is a model system that has been considered by different researchers [69,88–90]. We here reproduce the complete covariance matrix from our previous calculation [19]. By straightforward solution of the set of linear equations (100)–(105) we obtain to leading order in the particle masses

$$\langle x_1x_1\rangle^\circ = h_{22}d\left[\frac{1}{\beta_1}+\Delta\frac{h_{12}}{h_{22}}\gamma_2\right], \tag{107}$$

$$\langle x_2x_2\rangle^\circ = h_{11}d\left[\frac{1}{\beta_2}-\Delta\frac{h_{12}}{h_{11}}\gamma_1\right], \tag{108}$$

$$\langle x_1x_2\rangle^\circ = -\frac{h_{12}d}{2}\left[\frac{1}{\beta_1}+\frac{1}{\beta_2}+\Delta\frac{h_{11}\gamma_2-h_{22}\gamma_1}{h_{12}}\right], \tag{109}$$

$$\langle v_1v_1\rangle^\circ = \frac{1}{m_1}\left[\frac{1}{\beta_1}+\Delta\frac{h_{12}m_1}{\gamma_1}\right], \tag{110}$$

$$\langle v_2v_2\rangle^\circ = \frac{1}{m_2}\left[\frac{1}{\beta_2}-\Delta\frac{h_{12}m_2}{\gamma_2}\right], \tag{111}$$

$$\langle v_1v_2\rangle^\circ = \Delta\frac{m_2h_{11}-m_1h_{22}}{\gamma_1m_2+\gamma_2m_1}, \tag{112}$$

$$\langle x_1v_2\rangle^\circ = -\langle v_1x_2\rangle^\circ = \Delta, \tag{113}$$

where $d=(h_{11}h_{22}-h_{12}^2)^{-1}$ is the inverse determinant of the interaction matrix and the parameter

$$\Delta = \frac{h_{12}}{\gamma_1h_{22}+\gamma_2h_{11}}\left(\frac{1}{\beta_2}-\frac{1}{\beta_1}\right) \tag{114}$$

is a measure of the departure from equilibrium. For $\Delta = 0$, i.e., in equilibrium, the covariance matrix elements are given by the inverse of the Hamiltonian matrix according to Eq. (31) and in particular the off-diagonal velocity coupling terms $\langle v_1v_2\rangle^\circ$ and the position-velocity correlations $\langle x_1v_2\rangle^\circ = -\langle v_1x_2\rangle^\circ$ vanish. Off equilibrium, that means for $\Delta\neq 0$, these covariances are nonzero and thus the symmetry of the covariance matrix changes abruptly. The expressions in the square brackets in Eqs. (107)–(111) could in principle be used to define inverse effective temperatures for the covariance elements that are nonzero in equilibrium: inspection of the terms in the square brackets shows that they are all different. For the covariances Eqs. (112) and (113) that are proportional to Δ, the effective inverse temperatures are also different and diverge as equilibrium is approached, i.e., as $\Delta\to 0$. While for one or two of the covariances an effective temperature can be defined [91], consideration of the entire covariance matrix shows that the ascription of an effective temperature to a particle is not possible. This suggests that an effective temperature picture, where one assigns effective temperatures to particles, does not describe the particle statistics correctly and in particular does not characterize well the transition from equilibrium, $\Delta = 0$, to NEQ, $\Delta\neq 0$. To rescue the effective temperature picture one would have to ascribe different temperatures to each covariance matrix element, which clearly is far from the usefulness of the temperature definition at equilibrium. When basing the effective temperature definition on the fluctuation-dissipation relation, as an additional effect a frequency dependence appears [19,42,43,92], which is not reflected in the effective temperatures one would obtain based on the covariance matrix elements.

As an additional illustration of NEQ effects, we calculate the mean-squared difference between the particle positions, which from Eqs. (107)–(109) is to order m^0 in the particle mass given by

$$\langle (x_1-x_2)^2\rangle^\circ = \frac{\beta_1+\beta_2}{2h\beta_1\beta_2}\left[1-\left(\frac{\beta_1-\beta_2}{\beta_1+\beta_2}\right)\left(\frac{\gamma_1-\gamma_2}{\gamma_1+\gamma_2}\right)\right], \tag{115}$$

where we have considered two particles whose center of mass is not confined, $h_{11}=h_{22}=-h_{12}=h>0$ (note that this limit must be taken with care since the interaction matrix h_{ij} is not invertible in this case). The first term is the equilibrium result which survives in the limit $\beta_1=\beta_2$. Note that when the two friction coefficients and the two temperatures are different from each other, NEQ effects modify the equilibrium result. In fact, in the limits $\beta_1/\beta_2\gg 1$ and $\gamma_1/\gamma_2\gg 1$ (or $\beta_2/\beta_1\gg 1$ and $\gamma_2/\gamma_1\gg 1$) the distance between the particles tends to zero, thus indicating positional colocalization of particles, which in equilibrium one would only obtain from strong attractive interactions between particles. Interestingly, since $\beta_\alpha=\gamma_\alpha/\phi_{\alpha\alpha}^2$, where $\phi_{\alpha\alpha}$ denotes the strength of the Gaussian white noise that enters the Langevin equation, we see that colocalization is automatically obtained when the friction coefficients of particles γ_α are modified while keeping the random strengths $\phi_{\alpha\alpha}$ fixed. This indicates a tendency of particles to phase separate in position space in NEQ, which is indeed obtained in mixtures of particles that are coupled to different heat baths [31–34].

A similar calculation for the mean-squared velocity difference based on Eqs. (110)–(112) gives

$$\langle (v_1-v_2)^2\rangle^\circ = \frac{\beta_1+\beta_2}{m\beta_1\beta_2}\left[1-\frac{mh_{12}^2}{h\gamma_1\gamma_2}\left(\frac{\beta_1-\beta_2}{\beta_1+\beta_2}\right)\left(\frac{\gamma_1-\gamma_2}{\gamma_1+\gamma_2}\right)\right] \tag{116}$$

to order m^0 and where we used the simplifications $m=m_1=m_2$ and $h_{11}=h_{22}=h$. Also here we see that in the limit $\beta_1/\beta_2\gg 1$ and $\gamma_1/\gamma_2\gg 1$ (or $\beta_2/\beta_1\gg 1$ and $\gamma_2/\gamma_1\gg 1$) the momentum difference between the particles goes down. This is indicative of colocalization in momentum space, meaning that different particles tend to move with the same velocity. For a Newtonian Hamiltonian Eq. (88) which is diagonal in momentum space and for which momenta do not couple to positions, particle velocities are uncorrelated to each other in equilibrium. We thus conclude that the momentum localization demonstrated in Eq. (116) is a NEQ phenomenon that has no equilibrium analog for Newtonian Hamiltonians.

B. Stationary entropy production for three coupled particles: Heat flux from cold to warm reservoir

We will now present an explicit solution for a system consisting of three coupled particles, as schematically visualized in Fig. 2(a). Since here we are interested only in the stationary entropy production, Eq. (98), we only need to calculate the stationary velocity-position cross terms $\langle x_\epsilon v_\alpha\rangle^\circ$. We consider the simplified Hamiltonian

$$\mathcal{H} = \frac{h_1}{2}(x_1-x_2)^2+\frac{h_3}{2}(x_2-x_3)^2+\frac{m_1}{2}v_1^2+\frac{m_2}{2}v_2^2+\frac{m_3}{2}v_3^2, \tag{117}$$

where only particles 1 and 2 and particles 2 and 3 are coupled via harmonic bonds. We also assume the friction coefficients

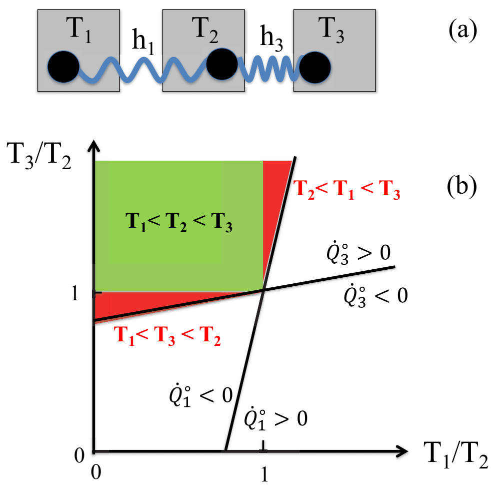

FIG. 2. (a) Schematic representation of three particles that interact elastically and that are coupled to heat baths with different temperatures. (b) Stationary heat flux diagram for symmetric elastic couplings $h_1 = h_3$, showing a colored region where heat flows from particle 3 to particle 2, $\dot{Q}_3^\circ > 0$, and heat flows from particle 2 to particle 1, $\dot{Q}_1^\circ < 0$. The green area denotes the temperature range $T_1 < T_2 < T_3$ and the red areas denote $T_2 < T_1 < T_3$ and $T_1 < T_3 < T_2$. The green area is the expected temperature range where heat flows from particle 3 coupled to the hot reservoir at T_3 to particle 2 coupled to the warm reservoir at T_2 and from particle 2 coupled to the warm reservoir at T_2 to particle 1 coupled to the cold reservoir at T_1. In the red areas heat flows between the particles against the temperature gradients of the heat reservoirs: In the upper red area heat flows from particle 2 coupled to the cold reservoir at T_2 to particle 1 coupled to the warm reservoir at T_1, in the lower red area heat flows from particle 3 coupled to the warm reservoir at T_3 to particle 2 coupled to the hot reservoir at T_2.

to be all the same: $\gamma_1 = \gamma_2 = \gamma_3 \equiv \gamma$. The solution strategy consists in eliminating $\langle v_{\alpha'}^2\rangle^\circ$ by inserting Eq. (105) into Eq. (103), which results in three equations, and solving the six equations defined by Eq. (104) by using Eqs. (100) and (101). These are nine equations for nine unknowns; it turns out that Eq. (102) is not needed. The results are to order m^0 in the particle mass given by

$$\langle x_2 v_3\rangle^\circ = \frac{1}{6\gamma(h_1+h_3)}\left[\frac{h_1}{\beta_1} + \frac{h_1+3h_3}{\beta_2} - \frac{2h_1+3h_3}{\beta_3}\right], \tag{118}$$

$$\langle x_1 v_2\rangle^\circ = \frac{1}{6\gamma(h_1+h_3)}\left[\frac{2h_3+3h_1}{\beta_1} - \frac{h_3}{\beta_3} - \frac{h_3+3h_1}{\beta_2}\right], \tag{119}$$

$$\langle x_1 v_3\rangle^\circ = \frac{1}{\gamma\Theta}\left[\frac{h_3+h_2}{\beta_1} + \frac{h_1-h_3}{\beta_2} - \frac{h_1+h_2}{\beta_3}\right], \tag{120}$$

where we defined for convenience

$$\Theta \equiv h_1 + 2h_2 + h_3 + \frac{h_2^2 - h_3^2}{h_1} + \frac{h_2^2 - h_1^2}{h_3}. \tag{121}$$

Inserting these results into the expression for the entropy production Eq. (98) we obtain

$$\begin{aligned}\dot{S}_{\rm res}^\circ/k_B &= \frac{1}{2}h_1\langle x_2 v_1\rangle^\circ(\beta_1-\beta_2) + \frac{1}{2}h_3\langle x_2 v_3\rangle^\circ(\beta_3-\beta_2)\\ &= \frac{1}{12\gamma(h_1+h_3)}\left[h_1h_3\left(\sqrt{\frac{\beta_1}{\beta_3}} - \sqrt{\frac{\beta_3}{\beta_1}}\right)^2\right.\\ &\quad + h_1(3h_1+h_3)\left(\sqrt{\frac{\beta_1}{\beta_2}} - \sqrt{\frac{\beta_2}{\beta_1}}\right)^2\\ &\quad \left. + h_3(3h_3+h_1)\left(\sqrt{\frac{\beta_3}{\beta_2}} - \sqrt{\frac{\beta_2}{\beta_3}}\right)^2\right].\end{aligned} \tag{122}$$

Obviously, the entropy production is positive and finite in the zero mass limit if the reservoir temperatures are different and if the coupling strengths h_1 and h_3 are finite and positive.

The limiting case of $h_3 = 0$ is insightful: in this case, particle 3 becomes decoupled and we are basically left with only two coupled particles. The expression (122) simplifies to

$$\dot{S}_{\rm res}^\circ/k_B = \frac{h_1}{4\gamma}\left(\sqrt{\frac{\beta_1}{\beta_2}} - \sqrt{\frac{\beta_2}{\beta_1}}\right)^2, \tag{123}$$

which describes the stationary entropy production of two reservoirs of inverse temperatures β_1 and β_2 that act on two particles that are subject to friction with coefficients γ and which are coupled by a harmonic spring of strength h_1. Obviously, also this entropy production is never negative and finite if the reservoir temperatures are different and if the coupling strength h_1 is finite.

The case of three particles that are coupled to heat reservoirs at three different temperatures allows demonstrating as an interesting NEQ effect the pumping of heat against a temperature gradient. Before we go into the analysis, we note that this does not constitute a violation of the second law of thermodynamics but rather is a NEQ entrainment effect. To proceed, we calculate the stationary heat flux from the heat reservoir at temperature T_1 to particle 1, which according to Eqs. (95) and (119) is given by

$$\begin{aligned}\dot{Q}_1^\circ &= -h_1\langle x_2 v_1\rangle^\circ = h_1\langle x_1 v_2\rangle^\circ\\ &= \frac{h_1}{6\gamma(h_1+h_3)}\left[\frac{2h_3+3h_1}{\beta_1} - \frac{h_3}{\beta_3} - \frac{h_3+3h_1}{\beta_2}\right].\end{aligned} \tag{124}$$

The stationary heat flux from the heat reservoir at temperature T_3 to particle 3 is according to Eqs. (95) and (118) given by

$$\begin{aligned}\dot{Q}_3^\circ &= -h_3\langle x_2 v_3\rangle^\circ\\ &= -\frac{h_3}{6\gamma(h_1+h_3)}\left[\frac{h_1}{\beta_1} + \frac{h_1+3h_3}{\beta_2} - \frac{2h_1+3h_3}{\beta_3}\right].\end{aligned} \tag{125}$$

Due to energy conservation $\dot{Q}_2^\circ = -\dot{Q}_1^\circ - \dot{Q}_3^\circ$ holds. Note that all stationary heat fluxes $\dot{Q}_1^\circ$, $\dot{Q}_2^\circ$, $\dot{Q}_3^\circ$ obviously vanish in equilibrium, i.e., when the temperatures of all heat reservoirs are the same. The flux from heat reservoir 1 according to Eq. (124) is negative, $\dot{Q}_1^\circ < 0$, i.e., heat flows into reservoir

1, for

$$\frac{T_3}{T_2} > -1 - 3h_1/h_3 + (2 + 3h_1/h_3)\frac{T_1}{T_2}, \tag{126}$$

which for equal elastic coupling strengths $h_1 = h_3$ simplifies to

$$\frac{T_3}{T_2} > -4 + 5\frac{T_1}{T_2}. \tag{127}$$

The flux from heat reservoir 3 according to Eq. (125) is positive, $\dot{Q}_3^\circ > 0$, i.e. heat flows out of reservoir 3, for

$$\frac{T_3}{T_2} > \frac{h_1 + 3h_3}{2h_1 + 3h_3} + \frac{h_1}{2h_1 + 3h_3}\frac{T_1}{T_2}, \tag{128}$$

which for equal elastic coupling strengths $h_1 = h_3$ simplifies to

$$\frac{T_3}{T_2} > 4/5 + \frac{T_1}{5T_2}. \tag{129}$$

As expected, the conditions (126) and (128) are simultaneously satisfied, i.e., heat flows into reservoir 1 and out of reservoir 3, for

$$T_1 < T_2 < T_3, \tag{130}$$

i.e., per unit time the heat amount $|\dot{Q}_3^\circ|$, flows from particle 3 (which is coupled to the hot heat reservoir) to particle 2 (which is coupled to the warm reservoir), and the heat amount $|\dot{Q}_1^\circ|$ flows from particle 2 to particle 1 (which is coupled to the cold reservoir).

More interestingly, the conditions (126) and (128) can also be simultaneously satisfied for the scenario

$$T_2 < T_1 < T_3. \tag{131}$$

In this case there is a small range of temperatures where heat flows from particle 3 (coupled to the hot heat reservoir) to particle 2 (coupled to the cold reservoir), which is expected, but at the same time a heat amount $|\dot{Q}_1^\circ|$ flows from particle 2 (coupled to the cold reservoir) to particle 1 (coupled to the warm reservoir). A similar scenario is provided for

$$T_1 < T_3 < T_2, \tag{132}$$

where for a small range of temperatures heat flows from particle 2 (coupled to the hot heat reservoir) to particle 1 (coupled to the cold reservoir), which is expected, but at the same time a heat amount $|\dot{Q}_3^\circ|$ flows from particle 3 (coupled to the warm reservoir) to particle 2 (coupled to the hot reservoir). This means we find situations where heat flows against the temperature gradient of the heat reservoirs that are coupled to the particles. This situation is indicated in Fig. 2(b) in a phase diagram for equal elastic coupling $h_1 = h_3$, in which case the simplified inequalities (127) and (129) hold. In the phase diagram the inequalities (127) and (129) are indicated by straight lines, and the region where Eq. (130) holds is indicated in green. The regions where the inequalities (127) and (129) and in addition Eq. (131) or Eq. (132) hold are indicated in red. Our discussion uses the fact that any heat $\dot{Q}_1^\circ$ that is transferred to or from reservoir 1 is transferred between particles 1 and 2 via elastic interactions, likewise, any heat $\dot{Q}_3^\circ$ that is transferred to or from reservoir 3 is transferred between particles 2 and 3. In other words, we do not only know the stationary heat fluxes from the reservoirs but also the stationary energy fluxes between the particles, which is due to the simple linear topology of the elastic particle interactions as indicated in Fig. 2(a).

As mentioned before, the finding of a heat flux against the reservoir temperature gradient does not violate the second law of thermodynamics. First, the heat flux from the particle coupled to the cold heat bath to the particle coupled to the warm heat bath is accompanied by an even larger heat flux from the particle coupled to the hot heat bath to the particle coupled to the warm heat bath, our result for the total entropy production (122) is strictly positive. Second, heat is not transferred directly between the heat reservoirs but only between particles that are coupled to heat reservoirs. In this connection it is crucial to remember [according to our previous discussion centered around the covariance elements (107)–(113)] that particles are not characterized by the temperature of the heat reservoir they are coupled to. Therefore, since the particles do not have well-defined effective temperatures, the second law of thermodynamics is not violated. One could be tempted to define effective temperatures based on the heat fluxes, but such a definition would be based solely on one entry of the covariance matrix, namely, the off-diagonal position velocity coupling $\langle x_\alpha v_\beta \rangle^\circ$, and not work for other applications.

C. Mapping on active particles

Similarly to recent calculations [70], we here show how active particle models can be described within the current framework of Markovian coupled particles. We want to describe a single active particle, for this we reduce the Hamiltonian (117) to two coupled massive particles and obtain

$$\mathcal{H} = \frac{h_1}{2}(x_1 - x_2)^2 + \frac{m_1}{2}v_1^2 + \frac{m_2}{2}v_2^2. \tag{133}$$

By assuming the friction coefficients to be the same, $\gamma_1 = \gamma_2 \equiv \gamma$, and choosing two different heat bath temperatures β_1 and β_2, we obtain from Eqs. (86), (87), (88), and (90) the coupled set of linear Langevin equations

$$m_1\, \dot{v}_1(t) = -\gamma v_1(t) - h_1[x_1(t) - x_2(t)] + \sqrt{\gamma/\beta_1}F_1(t),$$
$$m_2\, \dot{v}_2(t) = -\gamma v_2(t) - h_1[x_2(t) - x_1(t)] + \sqrt{\gamma/\beta_2}F_2(t).$$

The Langevin equation for the second particle is straightforwardly solved in the massless limit $m_2 = 0$ and gives

$$x_2(t) = \int_{-\infty}^{t} dt' e^{-(t-t')h_1/\gamma}\left[\frac{h_1}{\gamma}x_1(t') + (\beta_2\gamma)^{-1/2}F_2(t')\right]. \tag{134}$$

Inserting this solution into the Langevin equation for the first particle, we obtain the generalized Langevin equation

$$m_1\dot{v}_1(t) = -\int_{-\infty}^{\infty} dt'\Gamma(t-t')v_1(t') + F_R(t). \tag{135}$$

The memory function that appears in Eq. (135) is given by

$$\Gamma(t) = \theta(t)[2\gamma\delta(t) + h_1 e^{-th_1/\gamma}], \tag{136}$$

where $\theta(t)$ denotes the Theta function with the properties $\theta(t) = 1$ for $t > 0$ and $\theta(t) = 0$ for $t < 0$, which makes the memory function single-sided. The noise $F_R(t)$ in Eq. (135) is given by

$$F_R(t) = \sqrt{\gamma/\beta_1}F_1(t) + \frac{h_1}{\sqrt{\beta_2\gamma}}\int_{-\infty}^{t} dt' e^{-(t-t')h_1/\gamma}F_2(t') \tag{137}$$

and consists of the noise acting directly on the first particle and a term due to noise acting on the second particle. The latter term consists of a convolution integral because this noise is transmitted via the elastic bond of strength h_1. Defining the autocorrelation function of the random noise as $C_{\rm FF}(t) = \langle F_R(0)F_R(t)\rangle$, we obtain [19]

$$\beta_1 C_{\rm FF}(t) = 2\gamma\delta(t) + h_1(\beta_1/\beta_2)e^{-|t|h_1/\gamma}. \tag{138}$$

Comparing Eq. (136) and Eq. (138), we see that $\beta_1 C_{\rm FF}(t) = \Gamma(|t|)$, a consequence of the standard fluctuation-dissipation theorem [86], holds only for $\beta_1 = \beta_2$, i.e., if the two reservoir temperatures are the same. If $\beta_1 \neq \beta_2$, the memory kernel $\Gamma(t)$ and the random force correlation function $C_{\rm FF}(t)$ differ, which points to FDT violation and thus to the presence of an active NEQ process. The equation of motion (135) is similar to previously studied active particle models that were shown to exhibit NEQ phase transitions [93,94]. In fact, the active Ornstein-Uhlenbeck model is obtained by setting the exponential in the memory function (136) to zero while keeping the exponential term in the noise correlator (138). The entropy production of such active particle models has been intensely studied and debated [95–98].

The entropy production of the present active particle model, defined by Eq. (135) in conjunction with Eqs. (136) and (138), is given exactly by the simple expression (123); alternatively, it can be obtained directly from simulations by evaluating the position-velocity correlation functions in the general expression (98). The advantage of the present active particle model, which follows from a Newtonian Hamiltonian, is that the extremal and stability conditions in terms of the free entropy functional derived in this work are valid and not only describe the stationary NEQ distribution itself but also the approach to the stationary NEQ distribution. We hasten to add that not all active particle models can be mapped on our model, in particular models with non-Gaussian velocity distributions are not captured by our linear equations. In future work interacting active particles similar to the model derived here will be studied analytically as well as in simulations, it will be interesting to see whether the NEQ position and momentum localization effects demonstrated in Sec. V A also show up in those more complex systems.

VI. CONCLUSIONS

In this paper we consider the approach of Hamiltonian many-body systems that are coupled to multiple heat reservoirs with different temperatures to stationary NEQ distributions. Based on the exactly calculated approach of the covariance matrix $E(t)$ to its stationary NEQ form E°, we construct the functional $\mathcal{S}_{\rm free}(t)$ that yields the stationary covariance matrix E° at its extremum with respect to variations of $E(t)$. This function is called the free entropy, and it consists of the system distribution entropy $\mathcal{S}(t)$ and a term that accounts for interactions within the system, described by the Hamiltonian, and the frictional and noise coupling to the heat reservoirs.

Since in the stationary state the free entropy production by construction vanishes, as explained in Sec. IV A, the difference between the free entropy production and the total entropy production is the reservoir entropy production in the stationary state, $\dot{\mathcal{S}}_{\rm res}^\circ$, which is constant in time. The total entropy production $\dot{\mathcal{S}}_{\rm tot}(t)$, which is the sum of the system entropy production $\dot{\mathcal{S}}(t)$ and the reservoir entropy production $\dot{\mathcal{S}}_{\rm res}(t)$, follows from Eq. (1) by differentiation as

$$\dot{\mathcal{S}}_{\rm tot}(t) = \dot{\mathcal{S}}(t) + \dot{\mathcal{S}}_{\rm res}(t) = \dot{\mathcal{S}}_{\rm free}(t) + \dot{\mathcal{S}}_{\rm res}^\circ. \tag{139}$$

Using the result in Eq. (78) we obtain the explicit expression

$$\frac{\dot{\mathcal{S}}_{\rm tot}(t)}{k_B} = \frac{\dot{\mathcal{S}}(t)}{k_B} - \frac{E_{ij}^{\circ-1}\dot{E}_{ij}(t)}{2} - E_{ij}^{\circ-1}\langle\dot{z}_i(t)\rangle\langle z_j(t)\rangle + \dot{\mathcal{S}}_{\rm res}^\circ, \tag{140}$$

which consists of the system distribution entropy production $\dot{\mathcal{S}}(t)$, the stationary reservoir entropy production $\dot{\mathcal{S}}_{\rm res}^\circ$, and two terms that account for interactions within the system as well as the frictional and noise coupling to the environment. NEQ effects make these coupling terms differ dramatically from their equilibrium counterparts, in which case $E_{ij}^{\circ-1}$ becomes replaced by the Hamiltonian matrix $H_{ij}/(k_B T)$. In Appendix H we attempt to derive Eq. (139) explicitly by calculating the time-dependent reservoir entropy production $\dot{\mathcal{S}}_{\rm res}(t)$ from our explicit expressions for the time-dependent heat fluxes between the heat reservoirs and the system. We demonstrate that the heat fluxes are by themselves not sufficient to derive the reservoir entropy production $\dot{\mathcal{S}}_{\rm res}(t)$ if the reservoir temperatures are not the same, which implies that internal reservoir degrees of freedom contribute in a nonneglible manner to the reservoir entropy production for a NEQ system. The connection between the free entropy production $\dot{\mathcal{S}}_{\rm free}(t)$ (excluding the reservoirs) and the total entropy production $\dot{\mathcal{S}}_{\rm tot}(t)$ (including the reservoirs) will be reconsidered in future work using microscopic models for the heat reservoirs.

It is important to note that the reservoir entropy production in the stationary state, denoted as $\dot{\mathcal{S}}_{\rm res}^\circ$ and given explicitly in Eq. (98), does not depend on the time-dependent covariance matrix $E(t)$ but only on the stationary covariance matrix E°, so it is a constant with respect to variations in $E(t)$; thus, the total entropy production $\dot{\mathcal{S}}_{\rm tot}(t)$ exhibits the identical extremal and stability properties as the free entropy production $\dot{\mathcal{S}}_{\rm free}(t)$. It trivially follows from Eq. (140) that the total entropy $\mathcal{S}_{\rm tot}(t)$ of a NEQ system increases indefinitely with time, as shown explicitly in Eq. (1), and thus is not bounded, whereas the free entropy $\mathcal{S}_{\rm free}(t)$ is a well-defined and finite expression even for NEQ systems. Expression (140) will be useful for various applications whenever the total entropy production of different NEQ states of a system need to be compared. The functional Eq. (140) should also allow us to construct approximate methods for the description of NEQ nonlinear systems as well as for NEQ phase transitions.

One advantage of the current formulation is that in the limit when all heat reservoir temperatures become equal and the NEQ system transforms into an equilibrium system, the NEQ free entropy smoothly crosses over to the equilibrium free energy divided by $-T$, which displays the standard equilibrium extremal and stability properties of a canonical system. We thus have derived a unified framework to treat NEQ systems that are coupled to heat reservoirs at different temperatures on the same footing as equilibrium systems. We have in our work restricted ourselves to one specific class of NEQ models, namely, where NEQ is produced by stochastic forces with vanishing mean. Forces with a nonvanishing mean are straightforward to include and will be treated in future work.

Our approach rests on the harmonic approximation for the interaction Hamiltonian and for the friction and stochastic terms. It is not clear how to extend the present derivation to nonlinear systems, here variational and perturbative methods will most likely be helpful. Clearly, a harmonic model can always be obtained from a more complex, nonlinear model by a saddle-point expansion in terms of suitably defined deviatory coordinates. Our model should thus also apply to nonlinear systems as long as this saddle-point approximation is justified.

ACKNOWLEDGMENTS

We acknowledge funding from the Deutsche Forschungsgemeinschaft (DFG) via the SFB 1114, from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (Grant Agreement No. [835117]) and from the Infosys Foundation.

APPENDIX A: DERIVATION OF SHANNON ENTROPY

We start from the canonical partition function

$$\mathcal{Z} = \int d\vec{z} e^{-\beta\mathcal{H}(\vec{z})}, \tag{A1}$$

where $\beta = 1/(k_BT)$ denotes the inverse thermal energy. Using the thermodynamic definition of the free energy $\mathcal{F} = -k_BT \ln \mathcal{Z}$ we can write

$$\mathcal{F} = -k_BT \ln \int d\vec{z} e^{-\beta\mathcal{H}(\vec{z})}. \tag{A2}$$

From this we obtain, using the thermodynamic definition $\mathcal{S} = -\partial\mathcal{F}/\partial T$, for the entropy

$$\mathcal{S} = k_B \ln \mathcal{Z} + \mathcal{Z}^{-1}T^{-1} \int d\vec{z}\mathcal{H}(\vec{z})e^{-\beta\mathcal{H}(\vec{z})}. \tag{A3}$$

From the definition of the normalized equilibrium distribution (22), we obtain

$$\beta\mathcal{H}(\vec{z}) = -\ln \rho(\vec{z}) - \ln \mathcal{Z}, \tag{A4}$$

which is inserted into Eq. (A3) to give

$$\mathcal{S} = -k_B\mathcal{Z}^{-1} \int d\vec{z}e^{-\beta\mathcal{H}(\vec{z})} \ln \rho(\vec{z}). \tag{A5}$$

Using again the definition of the normalized equilibrium distribution (22), we finally obtain the Shannon expression for the entropy as

$$\mathcal{S}/k_B = -\int d\vec{z}\rho(\vec{z}) \ln \rho(\vec{z}). \tag{A6}$$

APPENDIX B: THE TIME-DEPENDENT PROBABILITY DISTRIBUTION IS GAUSSIAN

Here we show that the time-dependent distribution function is Gaussian and governed by the time-dependent inverse covariance matrix. The calculation holds for equilibrium as well as for NEQ systems. Given a solution $\vec{z}(t)$ of the Langevin Eq. (13), we construct the time-dependent probability distribution as

$$\rho(\vec{z}', t) = \vec{\delta}[\vec{z}' - \vec{z}(t)]. \tag{B1}$$

Using the Fourier representation of the δ function and the time-dependent solution (14) for given initial value, we obtain

$$\rho_F[\vec{z}', t, \vec{F}(\cdot)] = \int \frac{d\vec{\omega}}{(2\pi)^{2N}} \exp\left\{ -\imath\omega_i z_i' + \imath\omega_i\left[\langle z_i(t)\rangle + \int_0^t dt' e_{ij}^{-(t-t')A}\Phi_{jk}F_k(t')\right]\right\}, \tag{B2}$$

which is a functional of the random force trajectory $\vec{F}(t)$. The probability distribution is obtained by a path integral over all random force trajectories as

$$\rho(\vec{z}', t) = \int \mathcal{D}\vec{F}(\cdot)P[\vec{F}(\cdot)]\rho_F[\vec{z}', t, \vec{F}(\cdot)]. \tag{B3}$$

Here $P[\vec{F}(\cdot)]$ is the path integral weight, which for diagonal white noise is given by

$$P[\vec{F}(\cdot)] = \mathcal{N}_F^{-1} \exp\left\{-\frac{1}{4}\int dt\, dt' F_k(t)\delta(t-t')F_k(t')\right\} \tag{B4}$$

and leads to $\langle F_i(t)\rangle = 0$ and $\langle F_i(t)F_j(t')\rangle = 2\delta_{ij}\delta(t-t')$, where the normalization factor is given by $\mathcal{N}_F$. The path integral over the random noise in Eq. (B3) can be performed and leads to the expression

$$\rho(\vec{z}', t) = \int \frac{d\vec{\omega}}{(2\pi)^{2N}} \exp\{-\imath\omega_i z_i' + \imath\omega_i\langle z_i(t)\rangle - \omega_i E_{ij}(t)\omega_j\}, \tag{B5}$$

where the covariance matrix $E_{ij}(t)$ is given by Eq. (17). Performing the integral over $\vec{\omega}$ leads to

$$\rho(\vec{z}', t) = \mathcal{N}^{-1}(t) \exp\left\{-[z_i' - \langle z_i(t)\rangle]E_{ij}^{-1}(t)[z_j' - \langle z_j(t)\rangle]/2\right\}, \tag{B6}$$

with $\mathcal{N}$ given by Eq. (25), which is identical to the expression given in Eq. (24). Since the mean state vector $\langle z_i(t)\rangle$ in Eq. (B6) depends according to Eq. (15) on the initial state vector $z_i(0)$, the time-dependent probability distribution we derived here is in fact the conditional distribution and thus corresponds to the Green's function, this can be made explicit by rewriting Eq. (B6) as

$$\rho(\vec{z}', t|\vec{z}, 0) = \mathcal{N}^{-1}(t) \exp\left\{-\left[z_i' - e_{ik}^{-tA}z_k\right] \times E_{ij}^{-1}(t)\left[z_j' - e_{jl}^{-tA}z_l\right]/2\right\}. \tag{B7}$$

APPENDIX C: SEMIPOSITIVE DEFINITENESS OF MATRIX PRODUCT TRACE

We start from the part of the expression (64) for the free-energy production rate of an equilibrium system that involves the trace of a product of four matrices,

$$\dot{\mathcal{F}} = -C_{km}\big[H_{ml} - k_BTE_{ml}^{-1}\big]\frac{E_{lj}}{k_BT}\big[H_{jk} - k_BTE_{jk}^{-1}\big], \quad \text{(C1)}$$

where we have for simplicity omitted all time dependencies. An analogous expression also appears in the free entropy production in Eq. (79). The matrix trace expressions that appear in the free energy (34) and the free entropy (77) are special cases of the more general matrix product trace we consider here.

By defining the matrix $M_{ml} = H_{ml} - k_BTE_{ml}^{-1}$, the expression can be written more compactly as

$$k_BT\dot{\mathcal{F}} = -C_{km}M_{ml}E_{lj}M_{jk}, \quad \text{(C2)}$$

where all matrices C, M, E are symmetric and assumed to be nondefective, E is positive definite, C is semipositive definite and the definiteness of M (since it is the difference of two matrices) is not specified. We first diagonalize the matrix E by the similarity transformation

$$EP^EP^{E-1} = P^ED^EP^{E-1} = P^ED^EP^{E,T}, \quad \text{(C3)}$$

where in the last step we used that E is symmetric and P^E is thus an orthogonal matrix. The matrix $D^E_{i'j'} = \delta_{i'j}d^E_{i'}$ is diagonal with diagonal elements $d^E_{i'}$. We thus obtain

$$k_BT\dot{\mathcal{F}} = -C_{km}M_{ml}P_{li}D^E_{ij}P_{nj}M_{nk}. \quad \text{(C4)}$$

We next define $N_{mi} = M_{ml}P_{li}$ and obtain, by using that M is symmetric,

$$k_BT\dot{\mathcal{F}} = -C_{km}N_{mi}D^E_{ij}N_{kj}. \quad \text{(C5)}$$

We next diagonalize the matrix C by the similarity transformation

$$CP^CP^{C-1} = P^CD^CP^{C-1} = P^CD^CP^{C,T}, \quad \text{(C6)}$$

where $D^C_{i'j'} = \delta_{i'j}d^C_{i'}$ and obtain

$$k_BT\dot{\mathcal{F}} = -P^C_{kl}D^C_{lo}P^C_{mo}N_{mi}D^E_{ij}N_{kj}. \quad \text{(C7)}$$

We now define $R_{oi} = P^C_{mo}N_{mi}$, which is equivalent to $R_{lj} = N_{kj}P^C_{kl}$, and thus obtain

$$k_BT\dot{\mathcal{F}} = -D^C_{lo}R_{oi}D^E_{ij}R_{lj}. \quad \text{(C8)}$$

Now using that D^C and D^E are diagonal matrices we obtain

$$k_BT\dot{\mathcal{F}} = -d^C_lR_{li}d^E_iR_{li} = -d^C_ld^E_iR^2_{li} \leqslant 0, \quad \text{(C9)}$$

where the inequality follows since the matrix elements of R are real and all eigenvalues of the matrices C and E are not negative.

APPENDIX D: CONDITION OF DETAILED BALANCE

Detailed balance is satisfied if the probability for a transition from a state vector $\vec{z}$ at time t to a state vector $\vec{z}'$ at time $t+\tau$ is the same as the probability for a transition from $\vec{z}'$ at time t to $\vec{z}$ at time $t+\tau$, provided that all velocities are reversed. Detailed balance is satisfied for systems that are in equilibrium and is standardly used as the definition of equilibrium [87]. In this section we will demonstrate that the condition of detailed balance is equivalent to the condition based on the Boltzmann distribution, provided that the friction matrix is symmetric and that the stochastic field correlations satisfy certain symmetry relations. Our calculation is similar to classical derivations [5,8,99,100].

Using the joint probability distribution, the detailed balance condition can be written as [5,99,100]

$$\rho(\vec{z}',t;\vec{z},0) = \rho(W\vec{z},t;W\vec{z}',0), \quad \text{(D1)}$$

where W is the diagonal matrix that reverses all velocities and is given by

$$W = \begin{pmatrix} 1 & 0 & 0 & 0 & \\ 0 & -1 & 0 & 0 & \\ 0 & 0 & 1 & 0 & \\ 0 & 0 & 0 & -1 & \\ & & & & \ddots \end{pmatrix}. \quad \text{(D2)}$$

The detailed-balance condition can be brought into its standard form by using the conditional probability, resulting in

$$\rho(\vec{z}',t|\vec{z},0)\rho^\circ(\vec{z}) = \rho(W\vec{z},t|W\vec{z}',0)\rho^\circ(\vec{z}'), \quad \text{(D3)}$$

where

$$\rho^\circ(\vec{z}) = \rho(\vec{z},t\to\infty) = \mathcal{N}^{\circ-1}\exp\left\{-z_iE^{\circ-1}_{ij}z_j/2\right\} \quad \text{(D4)}$$

is the stationary distribution and accordingly the time argument is omitted. The stationary normalization constant follows from Eq. (25) as $\mathcal{N}^\circ = \sqrt{(2\pi)^{2N}\det E^\circ}$. The conditional probability distribution $\rho(\vec{z}',t|\vec{z},0)$, which is equivalent to the Green's function of the stochastic problem defined by the Langevin or the Fokker-Planck equation, is explicitly given in Eq. (B7) and is valid for equilibrium as well as for nonequilibrium systems. With the definition

$$\vec{z}' = \vec{z} + \vec{\delta}, \quad \text{(D5)}$$

the covariance matrix elements according to the two sides of the detailed-balance condition (D3) follow explicitly as

$$\langle\delta_o\delta_p\rangle^{\text{ls}} = E_{op}(t) + \big(\delta_{ok} - e^{-tA}_{ok}\big)\big(\delta_{pl} - e^{-tA}_{pl}\big)E^\circ_{kl} \quad \text{(D6)}$$

and

$$\langle\delta_o\delta_p\rangle^{\text{rs}} = W_{oi}E_{ij}(t)W_{jp} + \big(\delta_{ok} - W_{ol}e^{-tA}_{lm}W_{mk}\big) \times \big(\delta_{pn} - W_{pq}e^{-tA}_{qr}W_{rn}\big)E^\circ_{kn}. \quad \text{(D7)}$$

Since a Gaussian distribution is completely specified by its covariance matrix, detailed balance is satisfied if $\langle\delta_o\delta_p\rangle^{\text{ls}} = \langle\delta_o\delta_p\rangle^{\text{rs}}$ holds. Similar to previous treatments [5,99,100], we expand the detailed-balance condition in time. From Eq. (17) we obtain

$$E_{ij}(t) = 2tC_{ij} - t^2(A_{ik}C_{kj} + A_{jk}C_{ki}) + O(t^3). \quad \text{(D8)}$$

To second order in t we thus obtain

$$\langle\delta_o\delta_p\rangle^{\text{ls}} = 2tC_{op} + t^2\big(A_{ok}E^\circ_{kl}A_{pl} - A_{ok}C_{kp} - A_{pk}C_{ko}\big) \quad \text{(D9)}$$

and

$$\langle\delta_o\delta_p\rangle^{\text{rs}} = 2tW_{oi}C_{ij}W_{jp} + t^2W_{ol}\big(A_{lm}W_{mn}E^\circ_{nt}W_{ts}A_{rs} - A_{lk}C_{kr} - A_{rk}C_{kl}\big)W_{rp}. \quad \text{(D10)}$$

From enforcing the detailed-balance condition term by term in powers of t, we obtain the two conditions

$$C_{op} = W_{oi}C_{ij}W_{jp} \quad \text{(D11)}$$

and

$$A_{ok}E^{\circ}_{kl}A_{pl} - A_{ok}C_{kp} - A_{pk}C_{ko} = W_{ol}\big(A_{lm}W_{mn}E^{\circ}_{nt}W_{ts}A_{rs} - A_{lk}C_{kr} - A_{rk}C_{kl}\big)W_{rp}, \quad \text{(D12)}$$

which are consistent with previous results [5,99,100]. To evaluate the conditions (D11) and (D12), it is useful to express all matrices as products of particle matrices and 2×2 submatrices as introduced in Sec. IV B. With this, the matrix W in Eq. (D2) can be written as

$$W_{\alpha\beta} = \delta_{\alpha\beta}(s - r), \quad \text{(D13)}$$

where s and r are 2×2 matrices defined in Eq. (83) and where greek indices number particles. To test the consequences of condition (D11), we expand the random correlation matrix in the complete set of 2×2 submatrices according to

$$C_{\alpha\beta} = sC^{\mathrm{s}}_{\alpha\beta} + rC^{\mathrm{r}}_{\alpha\beta}r + utC^{\mathrm{ut}}_{\alpha\beta} + usC^{\mathrm{us}}_{\alpha\beta}. \quad \text{(D14)}$$

Inserting Eqs. (D13) and (D14) into Eq. (D11), we obtain by using the matrix product properties Eq. (85) that $C^{\mathrm{ut}}_{\alpha\beta} = 0 = C^{\mathrm{us}}_{\alpha\beta}$, i.e., the random correlation matrix cannot contain components that couple momenta and positions. We will later see that this condition is indeed satisfied for a large class of Hamiltonian models.

To evaluate the consequences of the condition (D12), we need to specify the symmetry of the Hamiltonian model. As in Eq. (88), the Hamiltonian matrix shall describe a Newtonian system where positions and momenta are decoupled and furthermore momentum contributions are diagonal in the particles indices,

$$H_{\alpha'\epsilon} = sh_{\alpha'\epsilon} + r\delta_{\alpha'\epsilon}/m_{\alpha'}. \quad \text{(D15)}$$

The friction matrix is copied from Eq. (90) and acts only on momentum degrees of freedom,

$$\Gamma_{\alpha\epsilon'} = r\gamma_{\alpha\epsilon'}/m_{\epsilon'}, \quad \text{(D16)}$$

in the following we explicitly allow the matrix $\gamma_{\alpha\epsilon'}$ also to be asymmetric. Using the inverse Hamiltonian from Eq. (89) and Eq. (D16), the Lyapunov-Boltzmann condition (37) yields the equilibrium random correlation matrix

$$2C_{\alpha\epsilon}/(k_BT) = (\gamma_{\alpha\epsilon} + \gamma_{\epsilon\alpha})r, \quad \text{(D17)}$$

which acts only on momentum degrees of freedom. We conclude that the first detailed-balance condition (D11) is automatically satisfied for a Newtonian system that is described by the Boltzmann distribution and thus satisfies the Lyapunov-Boltzmann condition (37).

We next probe condition (D12). According to the definition used in the main text, in equilibrium $E^{\circ}_{kl} = E^{\bullet}_{kl} = k_BTH^{-1}_{kl}$ holds and thus the stationary covariance matrix E°_{kl} does not couple momentum and position coordinates, therefore condition (D12) simplifies to

$$A_{ok}E^{\bullet}_{kl}A^T_{lp} - A_{ok}C_{kp} - C^T_{ok}A^T_{kp} = W_{ol}\big(A_{lm}E^{\bullet}_{ms}A^T_{sr} - A_{lk}C_{kr} - C^T_{lk}A^T_{kr}\big)W_{rp}, \quad \text{(D18)}$$

where the superindex T denotes the transpose of the matrix including the 2×2 submatrices. Equation (D18) holds if the left side is a matrix that does not couple momenta and positions. From the expression for A, Eq. (87), and $U_{\alpha\gamma} = u\delta_{\alpha\gamma}$ we obtain

$$A_{\alpha'\epsilon'} = -ush_{\alpha'\epsilon'} - ur\delta_{\alpha'\epsilon'}/m_{\alpha'} + r\gamma_{\alpha'\epsilon'}/m_{\epsilon'}. \quad \text{(D19)}$$

Replacing $E^{\bullet}_{kl}$ by the explicit result for H^{-1}_{kl} from Eq. (89), and using the results for C and A from Eqs. (D17) and (D19) we finally obtain for the left-hand side of Eq. (D18)

$$\begin{aligned} &rh_{\alpha'\beta'} + s\delta_{\alpha'\beta'}m^{-1}_{\beta'} \\ &+ r[(\gamma_{\epsilon\alpha'} - \gamma_{\alpha'\epsilon})\gamma_{\beta'\epsilon} + (\gamma_{\epsilon\beta'} - \gamma_{\beta'\epsilon})\gamma_{\alpha'\epsilon}]m^{-1}_{\epsilon}/2 \\ &+ ru(\gamma_{\alpha'\beta'} - \gamma_{\beta'\alpha'})m^{-1}_{\beta'}/2 + ur(\gamma_{\alpha'\beta'} - \gamma_{\beta'\alpha'})m^{-1}_{\alpha'}/2. \end{aligned} \quad \text{(D20)}$$

For an asymmetric friction matrix $\gamma_{\alpha\beta}$, this expression contains terms that couple positions and momenta, these are the two terms proportional to ru and ur, and therefore the second detailed-balance condition (D12) is not satisfied, even if the system obeys the Boltzmann distribution. If the friction matrix $\gamma_{\alpha\beta}$ is symmetric, the terms proportional to ru and ur disappear and the detailed-balance condition is satisfied. We conclude that for a symmetric friction matrix the condition of detailed balance and the definition of equilibrium we use in the main text, based on the Boltzmann distribution, are equivalent. Since none of the effects we study in this paper depend on the presence of asymmetries in the friction matrix, it is permissible and convenient to use the Boltzmann definition for equilibrium, which we primarily do since it is much easier to implement. The presence of an asymmetric friction matrix means that the principle of equal actio and reactio is broken on the level of the Langevin equation, this therefore constitutes a distinct NEQ scenario.

APPENDIX E: CONDITION OF VANISHING PROBABILITY CURRENT FOR OVERDAMPED PARTICLE DYNAMICS

In Sec. III D we show that the probability current is generally nonzero for underdamped particles. Here we consider overdamped particle motion and show that the probability current vanishes only for a symmetric friction matrix. To proceed, we rewrite the equation of motion expressed in terms of particle coordinates $y_\alpha(t) = (x_\alpha(t), p_\alpha(t))^T$, Eq. (86), as a second-order differential equation as

$$m_{\alpha'}\ddot{x}_{\alpha'}(t) = -\gamma_{\alpha'\beta}\dot{x}_\beta(t) - h_{\alpha'\beta}x_\beta(t) + \phi_{\alpha'\beta}F_\beta(t), \quad \text{(E1)}$$

where the Hamiltonian matrix $h_{\alpha'\beta}$ only acts on positions and the friction matrix $\gamma_{\alpha'\beta}$ only acts on momenta. $\phi_{\alpha'\beta}$ is the $N\times N$ random coupling strength matrix. Setting all masses m_α to zero, we obtain

$$\dot{x_\beta}(t) = -\gamma^{-1}_{\alpha\beta}h_{\beta\gamma}x_\gamma(t) + \gamma^{-1}_{\alpha\beta}\phi_{\beta\gamma}F_\gamma(t), \quad \text{(E2)}$$

which we can rewrite in a form similar to Eq. (13) as

$$\dot{x}_\alpha(t) = -A^{\mathrm{od}}_{\alpha\beta}x_\beta(t) + \Phi^{\mathrm{od}}_{\alpha\beta}F_\beta(t), \quad \text{(E3)}$$

where the overdamped versions of A and Φ are $N\times N$ matrices given by $A^{\mathrm{od}}_{\alpha\gamma} = \gamma^{-1}_{\alpha\beta}h_{\beta\gamma}$ and $\Phi^{\mathrm{od}}_{\alpha\gamma} = \gamma^{-1}_{\alpha\beta}\phi_{\beta\gamma}$. The overdamped version of Eq. (18) is $C^{\mathrm{od}}_{kl} = \Phi^{\mathrm{od}}_{km}\Phi^{\mathrm{od}}_{lm}$. The overdamped version of the Lyapunov-Boltzmann Eq. (37) turns out to be

$$2C^{\mathrm{od}}_{\alpha\beta}/(k_BT) = \gamma^{-1}_{\alpha\beta} + \gamma^{-1}_{\beta\alpha} \quad \text{(E4)}$$

and constitutes the relation between the random strength matrix and the friction matrix for overdamped systems. The overdamped version of the condition of vanishing probability current follows from Eq. (60) as

$$C^{\text{od}}_{\alpha\beta}/(k_BT) = \gamma^{-1}_{\alpha\beta} \tag{E5}$$

and, since $C^{\text{od}}_{\alpha\beta}$ is symmetric by construction, can be satisfied only if the friction matrix $\gamma_{\alpha\beta}$ is symmetric. Thus, if $\gamma_{\alpha\beta}$ is not symmetric, the probability current is nonzero even if the distribution is given by the Boltzmann distribution. We conclude that the condition of vanishing probability current and the equilibrium definition we use in this paper, namely, that the stationary distribution is given by the Boltzmann distribution [see Eqs. (22) and (31)] are not equivalent even in the overdamped limit. In fact, the condition of vanishing probability current is more restrictive since it precludes asymmetric friction matrices, which is not necessary to obtain Boltzmann distributions. Thus, there is a class of overdamped particle models that exhibit stationary Boltzmann distributions yet that exhibit nonzero probability currents in the stationary state.

APPENDIX F: FLUCTUATION-DISSIPATION THEOREM

Here we demonstrate for a specific two-particle model with an asymmetric friction matrix, that the fluctuation-dissipation theorem can be satisfied even when the system, as demonstrated in Appendix D, does not obey detailed balance. To proceed, we consider a system of two massive particles that are coupled by a general friction matrix. Following the notation of Eq. (E1) we write

$$m_1\,\ddot{x}_1(t) = -\gamma_1\dot{x}_1(t) - \gamma_{12}\dot{x}_2(t) + \phi_1F_1(t) + \phi_{12}F_2(t), \tag{F1}$$

$$m_2\,\ddot{x}_2(t) = -\gamma_2\dot{x}_2(t) - \gamma_{21}\dot{x}_1(t) + \phi_2F_2(t) + \phi_{21}F_1(t). \tag{F2}$$

Note that for ease of calculation we omit any positional couplings between the particles. Similar to the active-particle model in Sec. V C, where the two particle positions are coupled, the Langevin equation for the second particle can be solved and gives

$$\dot{x}_2(t) = \int_{-\infty}^{t} dt' e^{-(t-t')\gamma_2/m_2}\Bigg[-\frac{\gamma_{21}}{m_2}\dot{x}_1(t') + \frac{\phi_2}{m_2}F_2(t') + \frac{\phi_{21}}{m_2}F_1(t')\Bigg]. \tag{F3}$$

Inserting this solution into the Langevin equation for the first particle, we obtain the generalized Langevin equation

$$m_1\ddot{x}_1(t) = -\int_{-\infty}^{\infty} dt'\Gamma(t-t')\dot{x}_1(t') + F_R(t). \tag{F4}$$

The memory function that appears in Eq. (F4) is given by

$$\Gamma(t) = \theta(t)\bigg[2\gamma_1\delta(t) - \frac{\gamma_{12}\gamma_{21}}{m_2}e^{-t\gamma_2/m_2}\bigg], \tag{F5}$$

where $\theta(t)$ denotes the Theta function with the properties $\theta(t) = 1$ for $t > 0$ and $\theta(t) = 0$ for $t < 0$, which makes the memory function single-sided. The noise $F_R(t)$ in Eq. (F4) is given by

$$F_R(t) = \phi_1F_1(t) + \phi_{12}F_2(t) - \int_{-\infty}^{t} dt' e^{-(t-t')\gamma_2/m_2} \times \bigg[\frac{\gamma_{12}\phi_2}{m_2}F_2(t') + \frac{\gamma_{12}\phi_{21}}{m_2}F_1(t')\bigg]. \tag{F6}$$

The autocorrelation function of the random noise $C_{\text{FF}}(t) = \langle F_R(0)F_R(t)\rangle$ follows as

$$C_{\text{FF}}(t) = 2\big(\phi_1^2 + \phi_{12}^2\big)\delta(t) - \big[2\gamma_{12}(\phi_2\phi_{12} + \phi_1\phi_{21}) - \gamma_{12}^2\big(\phi_2^2 + \phi_{21}^2\big)/\gamma_2\big]e^{-|t|\gamma_2/m_2}. \tag{F7}$$

Defining the Langevin equation in analogy to Eq. (86) in terms of the particle momenta $p_{\alpha'}(t) = m_{\alpha'}v_{\alpha'}(t)$ as $\dot{p}_\alpha(t) = -A_{\alpha\beta}p_\beta(t) + \phi_{\alpha\beta}F_\beta(t)$ where $A_{\alpha\beta'} = \gamma_{\alpha\beta'}/m_{\beta'}$ and $H_{\alpha\beta'} = \delta_{\alpha\beta'}/m_{\beta'}$, we obtain from Eq. (36) the Boltzmann-Lyapunov equation

$$2C_{\alpha\beta} = 2\phi_{km}\phi_{lm} = k_BT(\gamma_{\alpha\beta} + \gamma_{\beta\alpha}), \tag{F8}$$

or, explicitly,

$$2\begin{pmatrix}\phi_1^2 + \phi_{12}^2 & \phi_1\phi_{21} + \phi_2\phi_{12}\\ \phi_1\phi_{21} + \phi_2\phi_{12} & \phi_2^2 + \phi_{21}^2\end{pmatrix} = k_BT\begin{pmatrix}2\gamma_1 & \gamma_{12} + \gamma_{21}\\ \gamma_{12} + \gamma_{21} & 2\gamma_2\end{pmatrix}. \tag{F9}$$

Provided Eq. (F9) holds, it is easy to see that Eqs. (F5) and Eq. (F7) satisfy the fluctuation-dissipation theorem [86]

$$C_{\text{FF}}(t) = k_BT\Gamma(|t|), \tag{F10}$$

even if the friction matrix $\gamma_{\alpha\beta}$ is asymmetric. In contrast and as shown in Appendix D, the more restrictive detailed-balance condition is satisfied only if the friction matrix $\gamma_{\alpha\beta}$ is symmetric. This suggests that the definition of equilibrium we use in this paper, namely, that the stationary distribution corresponds to the Boltzmann distribution, coincides with the fluctuation-dissipation theorem even for asymmetric $\gamma_{\alpha\beta}$. On the other hand, the detailed-balance condition is more restrictive and can be satisfied only if the friction matrix is symmetric.

APPENDIX G: EQUIVALENCE OF FREE ENTROPY AND KULLBACK-LEIBLER ENTROPY

The multidimensional expression for the Kullback-Leibler entropy reads [6,78]

$$\mathcal{S}_{\text{KL}} = -\int d\vec{z}\rho(\vec{z})\ln\bigg[\frac{\rho^\circ(\vec{z})}{\rho(\vec{z})}\bigg], \tag{G1}$$

where the Gaussian probability distribution $\rho(\vec{z})$ and the Gaussian stationary probability distribution $\rho^\circ(\vec{z})$ are given by

$$\rho(\vec{z}) = \mathcal{N}^{-1}\exp\big(-z_iE^{-1}_{ij}z_j/2\big), \tag{G2}$$

$$\rho^\circ(\vec{z}) = \mathcal{N}^{\circ-1}\exp\big(-z_iE^{\circ-1}_{ij}z_j/2\big) \tag{G3}$$

with the normalization constants $\mathcal{N} = \sqrt{(2\pi)^{2N}\det E}$ and $\mathcal{N}^\circ = \sqrt{(2\pi)^{2N}\det E^\circ}$. Note that we dropped all time dependencies for simplicity. Using the inequality

$$-\ln\left[\frac{\rho^\circ(\vec{z})}{\rho(\vec{z})}\right] \geqslant 1 - \frac{\rho^\circ(\vec{z})}{\rho(\vec{z})} \tag{G4}$$

and the fact that $\rho(\vec{z})$ and $\rho^\circ(\vec{z})$ are normalized, it immediately follows that

$$\mathcal{S}_{\rm KL} \geqslant 0 \tag{G5}$$

and that $\mathcal{S}_{\rm KL} = 0$ for $\rho(\vec{z}) = \rho^\circ(\vec{z})$. A short calculation shows that also $\delta\mathcal{S}_{\rm KL}/\delta\rho(\vec{z}) = 0$ for $\rho(\vec{z}) = \rho^\circ(\vec{z})$, thus, the Kullback-Leibler is minimal for $\rho(\vec{z}) = \rho^\circ(\vec{z})$. Performing the Gaussian integrals in Eq. (G1), one immediately obtains

$$\mathcal{S}_{\rm KL} = E^{\circ-1}_{ij} E_{ij}/2 - N - \ln[(2\pi)^{2N} \det E]/2 + \ln[(2\pi)^{2N} \det E^\circ]/2. \tag{G6}$$

Comparison of the Kullback-Leibler entropy (G6) with the nonequilibrium free entropy expression (70) shows that the two expressions are identical except the sign and terms that do not depend on covariance matrix elements E_{ij}.

APPENDIX H: NONSTATIONARY ENTROPY PRODUCTION OF HEAT RESERVOIRS

We start from the expression for the time-dependent entropy production of all reservoirs [Eq. (96)], which is based on the result for the heat fluxes between the reservoirs and the system in Eq. (95). We recall that the calculation of the reservoir heat fluxes is based on the general Newtonian Hamiltonian as defined by Eq. (88) with diagonal momentum friction as defined by Eq. (90) and the NEQ model defined by the diagonal noise correlation matrix (93).

Using the result for the time derivative of the covariance matrix (55) together with the explicit expression for the A matrix (87) and the expansion of the covariance matrix E (99), we obtain for the time derivative of the squared particle momenta

$$\frac{d\langle p^2_{\alpha'}(t)\rangle}{dt} = 2C_{\alpha'\alpha'} - \frac{2\gamma_{\alpha'}}{m_{\alpha'}}\langle p^2_{\alpha'}(t)\rangle - 2h_{\alpha'\epsilon}\langle p_{\alpha'}x_\epsilon\rangle. \tag{H1}$$

Inserting this into the total reservoir entropy production (96) we obtain

$$\frac{\dot{\mathcal{S}}_{\rm res}(t)}{k_B} = \frac{\beta_\alpha\gamma_\alpha}{m_\alpha}\left[\frac{\langle p^2_\alpha(t)\rangle}{m_\alpha} - \frac{1}{\beta_\alpha}\right], \tag{H2}$$

where the nonprimed index α is summed over. Clearly, in an equilibrium stationary state, where the kinetic energy of each particle obeys $\langle p^2_\alpha(t)\rangle/(m_\alpha) = 1/\beta_\alpha$ and $\beta_\alpha = \beta$, the entropy production vanishes. In a stationary state we can replace $\langle p^2_\alpha(t)\rangle$ by the expression (103) and thereby recover Eq. (98), so our calculation thus far is consistent. We will now test whether the stationary NEQ distribution constitutes an extremum of the total entropy production, which according to Eq. (139) is given by

$$\dot{\mathcal{S}}_{\rm tot}(t) = \dot{\mathcal{S}}(t) + \dot{\mathcal{S}}_{\rm res}(t), \tag{H3}$$

where $\dot{\mathcal{S}}(t)$ denotes the Shannon entropy production of the system distribution (46) and $\dot{\mathcal{S}}_{\rm res}(t)$ is given by Eq. (H2). As an explicit calculation shows, the derivative of the reservoir entropy production (H2) with respect to the covariance matrix gives

$$\frac{\partial\dot{\mathcal{S}}_{\rm res}(t)}{\partial E_{ij}} = \Gamma_{ik}C^{-1}_{km}\Gamma_{mj}, \tag{H4}$$

which using Eqs. (90) and (93) is a diagonal matrix with momentum entries $\beta_{\alpha'}\gamma_{\alpha'}/m^2_{\alpha'}$. For the system entropy production we obtain from Eq. (46)

$$\frac{\partial\dot{\mathcal{S}}(t)}{\partial E_{ij}} = -E^{-1}_{ik}(t)C_{km}E^{-1}_{mj}(t) \tag{H5}$$

and thus for the total entropy

$$\frac{\partial\dot{\mathcal{S}}_{\rm tot}(t)}{\partial E_{ij}} = \Gamma_{ik}C^{-1}_{km}\Gamma_{mj} - E^{-1}_{ik}(t)C_{km}E^{-1}_{mj}(t). \tag{H6}$$

Multiplying by C_{li} we obtain

$$C_{li}\frac{\partial\dot{\mathcal{S}}_{\rm tot}(t)}{\partial E_{ij}} = \Gamma_{lk}\Gamma_{kj} - C_{li}E^{-1}_{ik}(t)C^{-1}_{km}E^{-1}_{mj}(t), \tag{H7}$$

which vanishes when

$$\Gamma_{lk}\Gamma_{kj} = C_{li}E^{-1}_{ik}(t)C^{-1}_{km}E^{-1}_{mj}(t) \tag{H8}$$

holds. Taking the square root and multiplying by E we obtain

$$\Gamma_{lk}E_{ki}(t) = C_{li} \tag{H9}$$

as the equation which determines the extremum of the total entropy production. It turns out that that Eq. (H9) is satisfied by the equilibrium stationary distribution $E^\bullet_{ki}$, therefore for an equilibrium system the total entropy production as defined in Eq. (H3) is extremal in the equilibrium state.

It is easy to verify that for the stationary NEQ distribution E°_{ki} with nonzero matrix elements $\langle p_i p_j\rangle^\circ$ and $\langle x_i p_j\rangle^\circ$ for $i \neq j$, and Γ and C given by Eqs. (90) and (93), Eq. (H9) is not satisfied. This implies that the total entropy production (H3), which is the sum of the nonstationary reservoir entropy production (96) or (H2) (calculated from the heat fluxes between reservoirs and the system) and the system distribution entropy production (46), does not yield the exactly calculated stationary distribution E°_{ki} at its extremum and thus differs from the total entropy production constructed from the sum of the free entropy and the stationary reservoir entropy production according to Eq. (139).

We conclude that nonstationary entropy contributions that presumably stem from internal reservoir degrees of freedom render the expressions (96) and (H2) incomplete. These internal reservoir degrees could be included by using microscopic models for the heat reservoirs.

[1] L. Onsager, Reciprocal relations in irreversible processes. I, Phys. Rev. **37**, 405 (1931).

[2] I. Prigogine, *Etude Thermodynamique des Phénomènes Irréversibles* (Desoer, Liège, 1947).

[3] I. Prigogine and P. Mazur, Sur l'extension de la thermodynamique aux phénomènes irreversibles liés aux degrés de liberté internes, Physica **19**, 241 (1953).
[4] J. L. Lebowitz, Stationary nonequilibrium Gibbsian ensembles, Phys. Rev. **114**, 1192 (1959).
[5] R. Graham and H. Haken, Generalized thermodynamic potential for Markoff systems in detailed balance and far from thermal equilibrium, Z. Phys. **243**, 289 (1971).
[6] F. Schlögl, On stability of steady states, Z. Phys. **243**, 303 (1971).
[7] I. Procaccia and R. D. Levine, Potential work: A statistical-mechanical approach for systems in disequilibrium, J. Chem. Phys. **65**, 3357 (1976).
[8] J. Schnakenberg, Network theory of microscopic and macroscopic behavior of master equation system, Rev. Mod. Phys. **48**, 571 (1976).
[9] Y. Oono and M. Paniconi, Steady state thermodynamics, Prog. Theor. Phys. Suppl. **130**, 29 (1998).
[10] T. Hatano and S. I. Sasa, Steady-State Thermodynamics of Langevin Systems, Phys. Rev. Lett. **86**, 3463 (2001).
[11] M. Esposito, U. Harbola, and S. Mukamel, Entropy fluctuation theorems in driven open systems: Application to electron counting statistics, Phys. Rev. E **76**, 031132 (2007).
[12] S. Ramaswamy, The mechanics and statics of active matter, Annu. Rev. Condens. Matter Phys. **1**, 323 (2010).
[13] T. Chou, K. Mallick, and R. K. P. Zia, Non-equilibrium statistical mechanics: From a paradigmatic model to biological transport, Rep. Prog. Phys. **74**, 116601 (2011).
[14] U. Seifert, Stochastic thermodynamics, fluctuation theorems and molecular machines, Rep. Prog. Phys. **75**, 126001 (2012).
[15] 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).
[16] F. S. Gnesotto, F. Mura, J. Gladrow, and C. P. Broedersz, Broken detailed balance and non-equilibrium dynamics in living systems: A review, Rep. Prog. Phys. **81**, 066601 (2018).
[17] 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).
[18] P. Illien, R. Golestanian, and A. Sen, 'Fuelled' motion: Phoretic motility and collective behaviour of active colloids, Chem. Soc. Rev. **46**, 5508 (2017).
[19] R. R. Netz, Fluctuation-dissipation relation and stationary distribution of an exactly solvable many-particle model for active biomatter far from equilibrium, J. Chem. Phys. **148**, 185101 (2018).
[20] S. Katz, J. L. Lebowitz, and H. Spohn, Phase transitions in stationary nonequilibrium states of model lattice systems, Phys. Rev. B **28**, 1655 (1983).
[21] J. Krug, Boundary-Induced Phase Transitions in Driven Diffusive Systems, Phys. Rev. Lett. **67**, 1882 (1991).
[22] B. Schmittmann and R. K. P. Zia, Driven diffusive systems. An introduction and recent developments, Phys. Rep. **301**, 45 (1998).
[23] J. Dzubiella, G. P. Hoffmann, and H. Löwen, Lane formation in colloidal mixtures driven by an external field, Phys. Rev. E **65**, 021402 (2002).
[24] R. R. Netz, Conduction and diffusion in two-dimensional electrolytes, Europhys. Lett. **63**, 616 (2003).
[25] N. Kumar, H. Soni, S. Ramaswamy, and A. K. Sood, Flocking at a distance in active granular matter, Nat. Commun. **5**, 4688 (2014).
[26] E. Bertin, M. Droz, and G. Grégoire, Boltzmann and hydrodynamic description for self-propelled particles, Phys. Rev. E **74**, 022101 (2006).
[27] I. Theurkauff, C. Cottin-Bizonne, J. Palacci, C. Ybert, and L. Bocquet, Dynamic Clustering in Active Colloidal Suspensions with Chemical Signaling, Phys. Rev. Lett. **108**, 268303 (2012).
[28] R. Golestanian, Collective Behavior of Thermally Active Colloids, Phys. Rev. Lett. **108**, 038303 (2012).
[29] T. Speck, J. Bialke, A. M. Menzel, and H. Löwen, Effective Cahn-Hilliard Equation for the Phase Separation of Active Brownian Particles, Phys. Rev. Lett. **112**, 218304 (2014).
[30] F. Ginot, I. Theurkauff, D. Levis, C. Ybert, L. Bocquet, L. Berthier, and C. Cottin-Bizonne, Nonequilibrium Equation of State in Suspensions of Active Colloids, Phys. Rev. X **5**, 011004 (2015).
[31] A. Y. Grosberg and J. F. Joanny, Nonequilibrium statistical mechanics of mixtures of particles in contact with different thermostats, Phys. Rev. E **92**, 032118 (2015).
[32] S. N. Weber, C. A. Weber, and E. Frey, Binary Mixtures of Particles with Different Diffusivities Demix, Phys. Rev. Lett. **116**, 058301 (2016).
[33] H. Tanaka, A. A. Lee, and M. P. Brenner, Hot particles attract in a cold bath, Phys. Rev. Fluids **2**, 043103 (2017).
[34] J. Smrek and K. Kremer, Small Activity Differences Drive Phase Separation in Active-Passive Polymer Mixtures, Phys. Rev. Lett. **118**, 098002 (2017).
[35] S. S. N. Chari, C. Dasgupta, and P. K. Maiti, Scalar activity induced phase separation and liquid-solid transition in Lennard-Jones system, Soft Matter **15**, 7275 (2019).
[36] A. Sokolov, I. S. Aranson, J. O. Kessler, and R. E. Goldstein, Concentration Dependence of the Collective Dynamics of Swimming Bacteria, Phys. Rev. Lett. **98**, 158102 (2007).
[37] R. Suzuki and A. R. Bausch, The emergence and transient behaviour of collective motion in active filament systems, Nat. Commun. **8**, 41 (2017).
[38] D. Helbing, Traffic and related self-driven many-particle systems, Rev. Mod. Phys. **73**, 1067 (2001).
[39] C. W. Wolgemuth, T. R. Powers, and R. E. Goldstein, Twirling and Whirling: Viscous Dynamics of Rotating Elastic Filaments, Phys. Rev. Lett. **84**, 1623 (2000).
[40] H. Wada and R. R. Netz, Plectoneme creation reduces the rotational friction of a polymer, Europhys. Lett. **87**, 38001 (2009).
[41] G. S. Agarwal, Fluctuation-dissipation theorems for systems in non-thermal equilibrium and applications, Z. Phys. **252**, 25 (1972).
[42] P. C. Hohenberg and B. I. Shraiman, Chaotic behavior of an extended system, Physica D (Amsterdam) **37**, 109 (1989).
[43] L. F. Cugliandolo, J. Kurchan, and L. Peliti, Energy flow, partial equilibration, and effective temperatures in systems with slow dynamics, Phys. Rev. E **55**, 3898 (1997).
[44] T. Speck and U. Seifert, Restoring a fluctuation-dissipation theorem in a nonequilibrium steady state, Europhys. Lett. **74**, 391 (2006).
[45] N. Xu and C. S. O'Hern, Effective Temperature in Athermal Systems Sheared at Fixed Normal Load, Phys. Rev. Lett. **94**, 055701 (2005).

[46] P. Ilg and J. L. Barrat, From single-particle to collective effective temperatures in an active fluid of self-propelled particles, Europhys. Lett. **111**, 26001 (2007).

[47] M. Krüger and M. Fuchs, Fluctuation Dissipation Relations in Stationary States of Interacting Brownian Particles Under Shear, Phys. Rev. Lett. **102**, 135701 (2009).

[48] D. Levis and L. Berthier, Driven activation vs. thermal activation, Europhys. Lett. **79**, 60006 (2015).

[49] T. Harada and S. I. Sasa, Equality Connecting Energy Dissipation with a Violation of the Fluctuation-Response Relation, Phys. Rev. Lett. **95**, 130602 (2005).

[50] F. Zamponi, F. Bonetto, L. F. Cugliandolo, and J. Kurchan, A fluctuation theorem for non-equilibrium relaxational systems driven by external forces, J. Stat. Mech. (2005) P09013.

[51] J. Prost, J.-F. Joanny, and J. M. R. Parrondo, Generalized Fluctuation-Dissipation Theorem for Steady-State Systems, Phys. Rev. Lett. **103**, 090601 (2009).

[52] M. Baiesi, C. Maes, and B. Wynants, Fluctuations and Response of Nonequilibrium States, Phys. Rev. Lett. **103**, 010602 (2009).

[53] U. Seifert and T. Speck, Fluctuation-dissipation theorem in nonequilibrium steady states, Europhys. Lett. **89**, 10007 (2010).

[54] L. Willareth, I. M. Sokolov, Y. Roichman, and B. Lindner, Generalized fluctuation-dissipation theorem as a test of the Markovianity of a system, Europhys. Lett. **118**, 20001 (2017).

[55] S. Jabbari-Farouji, D. Mizuno, D. Derks, G. H. Wegdam, F. C. MacKintosh, C. F. Schmidt, and D. Bonn, Effective temperatures from the fluctuation-dissipation measurements in soft glassy materials, Europhys. Lett. **84**, 20006 (2008).

[56] J. R. Gomez-Solano, A. Petrosyan, S. Ciliberto, R. Chetrite, and K. Gawedzki, Experimental Verification of a Modified Fluctuation-Dissipation Relation for a Micron-Sized Particle in a Nonequilibrium Steady State, Phys. Rev. Lett. **103**, 040601 (2009).

[57] J. Mehl, V. Blickle, U. Seifert, and C. Bechinger, Experimental accessibility of generalized fluctuation-dissipation relations for nonequilibrium steady states, Phys. Rev. E **82**, 032401 (2010).

[58] P. Martin, A. J. Hudspeth, and F. Jülicher, Comparison of a hair bundle's spontaneous oscillations with its response to mechanical stimulation reveals the underlying active process, Proc. Natl. Acad. Sci. U. S. A. **98**, 14380 (2001).

[59] L. Dinis, P. Martin, J. Barral, J. Prost, and J.-F. Joanny, Fluctuation-Response Theorem for the Active Noisy Oscillator of the Hair-Cell Bundle, Phys. Rev. Lett. **109**, 160602 (2012).

[60] P. Bohec, F. F. Gallet, C. Maes, S. Safaverdi, P. Visco, and F. van Wijland, Probing active forces via a fluctuation-dissipation relation: Application to living cells, Europhys. Lett. **102**, 50005 (2013).

[61] D. Mizuno, C. Tardin, C. F. Schmidt, and F. C. MacKintosh, Nonequilibrium mechanics of active cytoskeletal networks, Science **315**, 370 (2007).

[62] J. L. Lebowitz and H. Spohn, A Gallavotti–Cohen-type symmetry in the large deviation functional for stochastic dynamics, J. Stat. Phys. **95**, 333 (1999).

[63] G. E. Crooks, Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences, Phys. Rev. E **60**, 2721 (1999).

[64] C. Jarzynski, Hamiltonian derivation of a detailed fluctuation theorem, J. Stat. Phys. **98**, 77 (2000).

[65] U. Seifert, Entropy Production Along a Stochastic Trajectory and an Integral Fluctuation Theorem, Phys. Rev. Lett. **95**, 040602 (2005).

[66] M. Doi, J. Zhou, Y. Di, and X. Xu, Application of the Onsager-Machlup integral in solving dynamic equations in nonequilibrium systems, Phys. Rev. E **99**, 063303 (2019).

[67] Z. Rieder, J. L. Lebowitz, and E. Lieb, Properties of a harmonic crystal in a stationary nonequilibrium state, J. Math. Phys. **8**, 1073 (1967).

[68] R. Zwanzig, Nonlinear generalized Langevin equations, J. Stat. Phys. **9**, 215 (1973).

[69] R. Filliger and P. Reimann, Brownian Gyrator: A Minimal Heat Engine on the Nanoscale, Phys. Rev. Lett. **99**, 230602 (2007).

[70] L. P. Dadhichi, A. Maitra, and S. Ramaswamy, Origins and diagnostics of the nonequilibrium character of active systems, J. Stat. Mech. (2018) 123201.

[71] J. Li, J. M. Horowitz, T. R. Gingrich, and N. Fakhri, Quantifying dissipation using fluctuating currents, Nat. Commun. **15**, 1666 (2019).

[72] H. B. Callen, *Thermodynamics and an Introduction to Thermostatistics* (John Wiley and Sons, New York, 1985).

[73] R. Balian, FranMassieu and the thermodynamic potentials, C. R. Phys. **18**, 526 (2017).

[74] M. Planck, *Treatise on Thermodynamics* (Dover Publications, New York, 1945).

[75] E. Schrödinger, *Statistical Thermodynamics* (Cambridge University Press, Cambridge, 1964).

[76] P. Biane and R. Speicher, Free diffusions, free entropy and free Fisher information, Ann. Inst. Henri Poincaré: Probab. Stat. **37**, 581 (2001).

[77] D. Voiculescu, Free entropy, Bull. London Math. Soc. **34**, 257 (2002).

[78] S. Kullback and R. A. Leibler, On information and sufficiency, Ann. Math. Stat. **22**, 79 (1951).

[79] C. Kwon, P. Ao, and D. J. Thouless, Structure of stochastic dynamics near fixed points, Proc. Natl. Acad. Sci. U. S. A. **102**, 13029 (2005).

[80] H. Ge, Extended forms of the second law for general time-dependent stochastic processes, Phys. Rev. E **80**, 021137 (2009).

[81] H. Ge and H. Qian, Physical origins of entropy production, free energy dissipation, and their mathematical representations, Phys. Rev. E **81**, 051133 (2010).

[82] H. Qian, A decomposition of irreversible diffusion processes without detailed balance, J. Math. Phys. **54**, 053302 (2013).

[83] H. Qian, Thermodynamics of the general diffusion process: Equilibrium supercurrent and nonequilibrium driven circulation with dissipation, Eur. Phys. J. Special Topics **224**, 781 (2015).

[84] T. Speck and U. Seifert, Integral fluctuation theorem for the housekeeping heat, J. Phys. A **38**, L581 (2005).

[85] R. Zwanzig, Memory effects in irreversible thermodynamics, Phys. Rev. **124**, 983 (1961).

[86] H. Risken, *The Fokker-Planck Equation* (Springer, Berlin, 1984).

[87] S. R. de Groot and P. Mazur, *Non-Equilibrium Thermodynamics* (North-Holland Publishing Co., Amsterdam, 1962).

[88] A. Crisanti, A. Puglisi, and D. Villamaina, Nonequilibrium and information: The role of cross correlations, Phys. Rev. E **85**, 061127 (2012).

[89] V. Dotsenko, A. Maciolek, O. Vasilyev, and G. Oshanin, Two-temperature Langevin dynamics in a parabolic potential, Phys. Rev. E **87**, 062130 (2013).

[90] A. Berut, A. Imparato, A. Petrosyan, and S. Ciliberto, Theoretical description of effective heat transfer between two viscously coupled beads, Phys. Rev. E **94**, 052148 (2016).

[91] A. Y. Grosberg and J. F. Joanny, Dissipation in a system driven by two different thermostats, Polym. Sci., Series C **60**, 118 (2018).

[92] A. Puglisi, A. Sarracino, and A. Vulpiani, Temperature in and out of equilibrium: A review of concepts, tools and attempts, Phys. Rep. **709**, 1 (2017).

[93] Y. Fily and M. C. Marchetti, Athermal Phase Separation of Self-Propelled Particles with No Alignment, Phys. Rev. Lett. **108**, 235702 (2012).

[94] T. F. F. Farage, P. Krinninger, and J. M. Brader, Effective interactions in active Brownian suspensions, Phys. Rev. E **91**, 042310 (2015).

[95] E. Fodor, C. Nardini, M. E. Cates, J. Tailleur, P. Visco, and F. van Wijland, How Far from Equilibrium is Active Matter? Phys. Rev. Lett. **117**, 038103 (2016).

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

[97] S. Shankar and M. C. Marchetti, Hidden entropy production and work fluctuations in an ideal active gas, Phys. Rev. E **98**, 020604(R) (2018).

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

[99] N. G. Van Kampen, Derivation of the phenomenological equations from the master equation I, Physica **23**, 707 (1957).

[100] N. G. Van Kampen, Derivation of the phenomenological equations from the master equation II, Physica **23**, 816 (1957).