PAPER

Fluctuation–dissipation relations in the absence of detailed balance: formalism and applications to active matter

Journal of Statistical Mechanics: Theory and Experiment
An IOP and SISSA journal

PAPER: Classical statistical mechanics, equilibrium and non-equilibrium

Fluctuation–dissipation relations in the absence of detailed balance: formalism and applications to active matter

Sara Dal Cengio[1], Demian Levis[1,2] and Ignacio Pagonabarraga[1,2,3]

[1] Departamento de Física de la Matèria Condensada, Universitat de Barcelona, Martí i Franquès 1, E08028 Barcelona, Spain
[2] UBICS University of Barcelona Institute of Complex Systems, Martí i Franquès 1, E08028 Barcelona, Spain
[3] Centre Européen de Calcul Atomique et Moléculaire (CECAM), École Polytechnique Fédérale de Lausanne (EPFL), Batochimie, Avenue Forel 2, 1015 Lausanne, Switzerland
E-mail: sara01.dalcengio@gmail.com

Received 16 July 2020
Accepted for publication 19 February 2021
Published 6 April 2021

Online at stacks.iop.org/JSTAT/2021/043201
https://doi.org/10.1088/1742-5468/abee22

Abstract. We present a comprehensive study about the relationship between the way detailed balance is broken in non-equilibrium systems and the resulting violations of the fluctuation–dissipation theorem. Starting from stochastic dynamics with both odd and even variables under time-reversal, we derive an explicit expression for the time-reversal operator, i.e. the Markovian operator which generates the time-reversed trajectories. We then exploit the relation between entropy production and the breakdown of detailed balance to establish general constraints on the non-equilibrium steady-states (NESS), which relate the non-equilibrium character of the dynamics with symmetry properties of the NESS distribution. This provides a direct route to derive extended fluctuation–dissipation relations, expressing the linear response function in terms of NESS correlations. Such framework provides a unified way to understand the

*Author to whom any correspondence should be addressed.

1742-5468/21/043201+30$33.00

departure from equilibrium of active systems and its linear response. We then consider two paradigmatic models of interacting self-propelled particles, namely active Brownian particles and active Ornstein–Uhlenbeck particles. We analyze the non-equilibrium character of these systems (also within a Markov and a Chapman–Enskog approximation) and derive extended fluctuation–dissipation relations for them, clarifying which features of these active model systems are genuinely non-equilibrium.

Keywords: active matter, stationary states, transport properties, stochastic particle dynamics

Contents

J. Stat. Mech. (2021) 043201

1. Introduction

The fluctuation–dissipation theorem (FDT) relates the correlations of spontaneous fluctuations, to the fluctuations induced by external stimuli [1]. In practice, it allows to probe the response to external fields by analyzing the corresponding time-dependent equilibrium fluctuations, either in experiments or in simulations. For instance, it allows to infer transport or mechanical properties of soft materials from light scattering without ever perturbing them [2, 3]. The FDT plays a particularly important role in statistical mechanics as it is among the very rare general results in non-equilibrium, although near to, conditions. It is valid for any equilibrium system (both in the classical and quantum realm) gently driven out-of-equilibrium by a small perturbation. Accordingly to the FDT, the response of an observable A at time t to a perturbation h, applied at time s, and causing the change in the energy of the system $E \to E - h(s)B$, is determined by an equilibrium correlation function as

$$\left.\frac{\delta\langle A\rangle_t}{\delta h(s)}\right|_{h\to 0} = R_A(t,s) = \beta\frac{\partial}{\partial s}\langle A(t)B(s)\rangle_{\rm eq}, \quad t > s \tag{1}$$

where R_A is the response function of the observable A reacting to a perturbation conjugated to B and $\langle A(t)B(s)\rangle_{\rm eq}$ is the equilibrium correlation function between these two latter observables at temperature $\beta^{-1} = k_{\rm B}T$[4].

In its general formulation above, the FDT was first derived in the context of Hamiltonian mechanics [4, 5]. It has later been extended to stochastic descriptions [6] which rely on the hypothesis of scale separation between the system of interest and the bath, giving rise to dissipative and noisy terms entering in the equations of motion. As a result, the latter are no longer invariant under time-reversal. Nevertheless a footprint of reversibility holds at the stochastic description level under the name of detailed balance (DB) [7, 8]. As long as DB is guaranteed, the FDT holds, both in thermal and athermal states [9, 10].

For systems breaking DB, relentlessly evolving far-from-equilibrium, the FDT is no longer justified. The question of whether a similar relation as equation (1) can be derived in this case, has been the focus of a great deal of research efforts over the last decades. In particular, several extended fluctuation–dissipation relations (FDR) have been derived, using different approaches, for systems in non-equilibrium steady-states (NESS). However, contrary to equilibrium states, no universal relation such as equation (1) exists for NESS. The establishment of a general extended FDR with the features of the equilibrium FDT, remains a central challenge towards the construction of a general framework to deal with non-equilibrium systems. In the context of stochastic dynamics, extended FDR for NESS have mostly focused on overdamped descriptions [11–17]. We refer to [18–21] for recent reviews on the topic.

Among the variety of non-equilibrium systems, living matter constitutes a particularly interesting class. From a physics viewpoint, it can be considered as active matter: systems composed of interacting units—be it a cell, a molecular motor, an auto-catalytic colloid—capable of extracting energy from their environment to perform some task and,

[4] We consider, without loss of generality, observables with zero mean.

J. Stat. Mech. (2021) 043201

typically (as in the cases we consider here), self-propel. In contrast with passive systems relaxing towards NESS, which are driven out-of-equilibrium by external global means (usually through their boundaries), active systems break DB at the level of each of its constituents, defining a fundamentally different class of non-equilibrium systems [22, 23].

A renewed interest in the characterization of NESS comes indeed from active matter physics. The possibility to extend equilibrium-like concepts to characterize active matter, in particular their NESS, has been the focus of intense efforts over the past decade. Most of our general understanding of such fundamental aspects of active matter has been gained through the detailed investigation of simple models of self-propelled particles, such as the active Brownian particles (ABP) [24–27] and active Ornstein–Uhlenbeck particles (AOUP) [28–33] models that we consider here. Quantities such as the pressure or chemical potential have been defined for model active systems and exploited to characterize their phase behavior and the nature of the (non-equilibrium) phase transitions they exhibit [27, 34–45].

Activity results on transport phenomena which are impossible in equilibrium passive systems, as recently observed in experiments involving biological microorganisms [46–49] as well as artificial phoretic motors [50–53]. A key step towards the fundamental understanding of active materials is to characterize transport coefficients and establish extended Green–Kubo expressions resulting from the FDR. Attempts of this sort have been limited to specific cases or regimes, mostly considering the small activity limit. In [54, 55], activity is treated as the perturbation on an otherwise equilibrium state, while in [31] an FDR is obtained in a small activity regime for which the dynamics of the system fulfills DB. In both cases, the reference state that is perturbed is not a genuine NESS: in the first case, it is an equilibrium state with Boltzmann statistics, while in the latter an effective equilibrium state with a generalized potential. The fundamental difficulties arising from the violation of DB are therefore bypassed. The linear response beyond such small activity limit has been analyzed for a single active particle in [56]. In [57], response functions were obtained beyond such limit regimes, although they are not written in terms of NESS time-correlation functions, as one wills for establishing FDR, but as weighted averages (in the spirit of Malliavin weight sampling [58, 59]). Another strategy consists in systematically quantifying the violations of the FDT through an effective temperature [28, 60–68]. While this approach provides useful insights into the dynamics of NESS, it does not carry the same piece of information as an FDR, i.e. a generic way to assess the response function of an active system in terms of the steady-state fluctuations of measurable observables.

Although the non-equilibrium nature of active systems is intrinsically different from the one of passive driven systems, as for the construction of a linear response theory, the fundamental difficulty to be tackled in both cases is the breakdown of DB. Thus, before restricting our analysis to specific models of active particles, we first consider a generic out-of-equilibrium Markovian process and establish the consequences of DB violation. We consider here a dynamics with both even and odd variables under time-reversal (the prototypical example being the underdamped Langevin process) and derive a general expression for its time-reversal operator. This allows us to relate explicitly the breakdown of DB to the presence of irreversible fluxes while properly including the parity

J. Stat. Mech. (2021) 043201

under time-reversal in the picture. Further, by using the definition of entropy production we establish novel general constraints on the NESS to be fulfilled by *any* Markovian dynamics. The framework and results obtained are particularly relevant for systems with both even and odd variables under time-reversal and stand for a relation between the nature of the non-equilibrium fluxes and the symmetry (under time-reversal) of the NESS distribution. We finally turn to discuss the validity/violation of the FDT in archetypical models of active particles: ABP and AOUP. For ABP we consider the non-interacting limit and an effective equilibrium regime resulting from a Markovian approximation, extending the results and discussion in [69] in connection to the general formalism introduced here. For AOUP we derive a genuine, although approximated, non-equilibrium FDR unveiling the interplay between activity and interactions beyond an effective equilibrium regime. We discuss in detail the specificities of AOUP as compared with ABP as well as the different approximation schemes used in the literature to deal with many-body effects, providing a thorough account of the non-equilibrium aspects of both model systems.

The paper is organized as follows. In section 2 we establish the general framework and notation used throughout the paper. Here standard concepts and tools of stochastic modeling and equilibrium dynamics are recalled in order to make the text entirely self-consistent. A reader familiar with the formalism of stochastic processes may directly move to section 3. In the latter section, we derive an explicit expression for the generator of the time-reversed dynamics in the presence of both even and odd variables under time-reversal. By making the connection with entropy production, we thus establish general constraints a non-equilibrium stationary distribution must fulfill and link them to previously derived generalized FDR. Notably the framework here presented holds both for underdamped and overdamped dynamics breaking DB. Finally section 4 is dedicated to the application of these results to simple models of self-propelled particles (ABP and AOUP). Section 5 contains our conclusions and final remarks.

2. General aspects and definitions of stochastic dynamics

2.1. Fokker–Planck equation

Our starting point is a generic system with N dynamic variables $\mathbf{\Gamma} \equiv \{\Gamma_i\}_{i=1}^N$ defined on a manifold $\mathcal{M} \subset \mathbb{R}^N$. We introduce a probability distribution Ψ which assigns $\Psi(\mathbf{\Gamma}, t)$ to any point $\mathbf{\Gamma} \in \mathcal{M}$ at a time t. The time evolution of $\Psi(\mathbf{\Gamma}, t)$ is described by a generator $\Omega_0(\mathbf{\Gamma})$:

$$\partial_t \Psi(\mathbf{\Gamma}, t) = \Omega_0(\mathbf{\Gamma})\Psi(\mathbf{\Gamma}, t) \tag{2}$$

together with an appropriate initial condition $\Psi(\mathbf{\Gamma}, 0)$. Formal integration of equation (2) leads to $\Psi(\mathbf{\Gamma}, t) = \mathrm{e}^{\Omega_0 t}\Psi(\mathbf{\Gamma}, 0)$. We denote Ψ_0 a steady-state solution:

$$\Omega_0 \Psi_0 = 0. \tag{3}$$

Up to here, we did not need to specify the nature of the dynamics. We shall now focus on stochastic dynamics (although the following formalism could be extended to, say,

J. Stat. Mech. (2021) 043201

Hamiltonian dynamics). Under the hypothesis of Markovianity [70], the generator in equation (2) has the so-called Fokker–Planck form:

$$\Omega_0(\mathbf{\Gamma}) = \sum_i \left(-\partial_i \mathcal{A}_i(\mathbf{\Gamma}) + \sum_j \partial_i \partial_j \mathcal{B}_{ij}(\mathbf{\Gamma}) \right) \tag{4}$$

where $\partial_i \equiv \partial/\partial\Gamma_i$, $\mathcal{A} \equiv \{\mathcal{A}_i\}_{i=1}^N$ is the drift vector and $\mathcal{B} \equiv \{\mathcal{B}_{ij}\}_{i,j=1}^N$ is the $N \times N$ diffusion matrix. In the following, unless explicitly stated otherwise, we will take $\mathcal{B}$ to be invertible and diagonal with constant entries, such that $\mathcal{B}_{ij} \equiv D_i \delta_{ij}$.

The dynamics is fully specified by the knowledge of $\Psi(\mathbf{\Gamma}, t)$ or, equivalently, by the knowledge of the conditional probability density $P(\mathbf{\Gamma}, t|\mathbf{\Gamma}_0, t_0)$ defined as the probability to be in $\mathbf{\Gamma}$ at time t given the configuration $\mathbf{\Gamma}_0$ at time t_0. By replacing $\Psi(\mathbf{\Gamma}, t)$ with $P(\mathbf{\Gamma}, t|\mathbf{\Gamma}_0, t_0)$ in equation (2) one gets

$$\partial_t P(\mathbf{\Gamma}, t|\mathbf{\Gamma}_0, t_0) = \Omega_0(\mathbf{\Gamma}) P(\mathbf{\Gamma}, t|\mathbf{\Gamma}_0, t_0) \tag{5}$$

which is often called *forward* equation to distinguish it from the *backward* equation:

$$\partial_{t_0} P(\mathbf{\Gamma}, t|\mathbf{\Gamma}_0, t_0) = -\Omega_0^\dagger(\mathbf{\Gamma}_0) P(\mathbf{\Gamma}, t|\mathbf{\Gamma}_0, t_0) \tag{6}$$

where $\Omega_0^\dagger(\mathbf{\Gamma}) = \sum_i \mathcal{A}_i(\mathbf{\Gamma})\partial_i + D_i\partial_i^2$ is the adjoint operator of Ω_0. Equation (6) will be our starting point to characterize the departure from equilibrium in systems breaking DB (see section 3.1).

2.2. Symmetry aspects under time-reversal

We shall distinguish the dynamic variables $\{\Gamma_i\}_{i=1}^N$ according to their parity under time-reversal

$$\mathcal{T} : \mathbf{\Gamma} \in \mathcal{M} \mapsto \boldsymbol{\varepsilon}\mathbf{\Gamma} \equiv \{\varepsilon_i \Gamma_i\} \in \mathcal{M}, \quad \varepsilon_i = \pm 1. \tag{7}$$

Variables Γ_i for which $\varepsilon_i = 1$ are said even under time-reversal and variables for which $\varepsilon_i = -1$ are said to be odd[5].

The Fokker–Planck equation stands for the conservation of the probability density and can thus be written in terms of a probability flux $\boldsymbol{J} \equiv \{J_i\}_{i=1}^N$ as

$$\partial_t \Psi(\mathbf{\Gamma}, t) = \Omega_0(\mathbf{\Gamma})\Psi(\mathbf{\Gamma}, t) = -\nabla \cdot \boldsymbol{J}(\mathbf{\Gamma}, t) \tag{8}$$

$$J_i(\mathbf{\Gamma}, t) \equiv \mathcal{A}_i(\mathbf{\Gamma})\Psi(\mathbf{\Gamma}, t) - D_i\partial_i\Psi(\mathbf{\Gamma}, t) \tag{9}$$

where $\nabla \equiv \{\partial_i\}$. Since we allow $\{\Gamma_i\}$ to be either even or odd under time-reversal, we can decompose the drift vector in a *reversible* and an *irreversible* part, $\mathcal{A} = \mathcal{A}^{\text{rev}} + \mathcal{A}^{\text{irr}}$, defined as

$$\mathcal{A}_i^{\text{rev}}(\mathbf{\Gamma}) \equiv \frac{1}{2}\left[\mathcal{A}_i(\mathbf{\Gamma}) - \varepsilon_i \mathcal{A}_i(\boldsymbol{\varepsilon}\mathbf{\Gamma})\right] \tag{10}$$

$$\mathcal{A}_i^{\text{irr}}(\mathbf{\Gamma}) \equiv \frac{1}{2}\left[\mathcal{A}_i(\mathbf{\Gamma}) + \varepsilon_i \mathcal{A}_i(\boldsymbol{\varepsilon}\mathbf{\Gamma})\right] \tag{11}$$

[5] For instance, if one has in mind the dynamics of a particle in phase space, $\mathbf{\Gamma} = (r, p)$ and $\mathcal{T}:(r, p) \mapsto (r, -p)$.

J. Stat. Mech. (2021) 043201

which, under time-reversal transform as

$$\mathcal{A}_i^{\rm rev}(\boldsymbol{\varepsilon}\boldsymbol{\Gamma}) = -\varepsilon_i \mathcal{A}_i^{\rm rev}(\boldsymbol{\Gamma}), \qquad \mathcal{A}_i^{\rm irr}(\boldsymbol{\varepsilon}\boldsymbol{\Gamma}) = \varepsilon_i \mathcal{A}_i^{\rm irr}(\boldsymbol{\Gamma}). \tag{12}$$

We thus identify two distinct contributions to the total probability flux $J_i(\boldsymbol{\Gamma},t) = J_i^{\rm rev}(\boldsymbol{\Gamma},t) + J_i^{\rm irr}(\boldsymbol{\Gamma},t)$, where

$$J_i^{\rm rev}(\boldsymbol{\Gamma},t) = \mathcal{A}_i^{\rm rev}(\boldsymbol{\Gamma})\Psi(\boldsymbol{\Gamma},t) \tag{13}$$

$$J_i^{\rm irr}(\boldsymbol{\Gamma},t) = \mathcal{A}_i^{\rm irr}(\boldsymbol{\Gamma})\Psi(\boldsymbol{\Gamma},t) - D_i\partial_i\Psi(\boldsymbol{\Gamma},t). \tag{14}$$

We denote the steady-state flux $\boldsymbol{J}_0$ and define the *steady-state velocity* as

$$\boldsymbol{\mathcal{V}}(\boldsymbol{\Gamma}) \equiv \boldsymbol{J}_0(\boldsymbol{\Gamma})/\Psi_0(\boldsymbol{\Gamma}) = \{\mathcal{V}_i^{\rm rev} + \mathcal{V}_i^{\rm irr}\}_{i=1}^N. \tag{15}$$

(In the following, we also report its time-dependence $\boldsymbol{\mathcal{V}}(\boldsymbol{\Gamma},t) \equiv \boldsymbol{J}(\boldsymbol{\Gamma},t)/\Psi(\boldsymbol{\Gamma},t)$). The decomposition of the probability flux into two contributions with different symmetry under time-reversal will play a central role in the following treatment [71–73].

It is worth at this stage to make a few remarks. By definition (see equation (10)), in the absence of odd variables under time-reversal, as for the case of overdamped Brownian dynamics, $\mathcal{A}^{\rm rev}$ is identically zero. In this case the reversible flux vanishes and the total probability flux is identified with the irreversible contribution, i.e. $\boldsymbol{J}(\boldsymbol{\Gamma},t) = \boldsymbol{J}^{\rm irr}(\boldsymbol{\Gamma},t)$. Finally, in the absence of dissipation (and diffusion) the generator Ω_0 would be identified with the Liouville operator. In that case, the irreversible part of the flux would vanish and the motion would be purely reversible, as expected from time-reversal symmetry of the 'microscopic' Hamilton equations of motion.

2.3. Observables

Physical observables are represented by real functions acting on $\mathcal{M}$, such that $A:\boldsymbol{\Gamma}\in\mathcal{M}\mapsto A(\boldsymbol{\Gamma})\in\mathbb{R}$. Their steady-state average is defined as

$$\langle A\rangle_0 = \int_{\mathcal{M}} {\rm d}\boldsymbol{\Gamma} A(\boldsymbol{\Gamma})\Psi_0(\boldsymbol{\Gamma}) \tag{16}$$

and the ensemble average at time t is defined as

$$\langle A\rangle_t = \int_{\mathcal{M}} {\rm d}\boldsymbol{\Gamma} A(\boldsymbol{\Gamma})\Psi(\boldsymbol{\Gamma},t) = \int_{\mathcal{M}} {\rm d}\boldsymbol{\Gamma} A(\boldsymbol{\Gamma}){\rm e}^{\Omega_0 t}\Psi(\boldsymbol{\Gamma},0). \tag{17}$$

The time evolution may be given to the observables (instead of the probability distribution) using the adjoint of the Fokker–Planck operator

$$\langle A\rangle_t = \int_{\mathcal{M}} {\rm d}\boldsymbol{\Gamma}{\rm e}^{\Omega_0^\dagger t}A(\boldsymbol{\Gamma})\Psi(\boldsymbol{\Gamma},0) \equiv \int_{\mathcal{M}} {\rm d}\boldsymbol{\Gamma} A(t)\Psi(\boldsymbol{\Gamma},0). \tag{18}$$

Both expressions of $\langle A\rangle_t$ are fully equivalent and are respectively referred to as the Schrödinger and Heisenberg representation, by analogy with quantum mechanics [74].

J. Stat. Mech. (2021) 043201

Finally, we introduce the linear response function $R_A(t,s)$ encoding the change of an observable A, due to an infinitesimal perturbation h which results in a change of generator $\Omega_0 \to \Omega = \Omega_0 + \Omega_{\rm ext}(h)$. It is defined as

$$\langle A\rangle_t - \langle A\rangle_0 = \int_0^t \mathrm{d}s R_A(t,s)h(s) + O(h^2), \quad t > s, \tag{19}$$

where we assumed the perturbation to be switched on at $t = 0$ and the system to be originally at steady state. Then, writing the response function R_A as given by the equilibrium FDT equation (1), considering a constant perturbation h and integrating, we obtain

$$\langle A\rangle_t - \langle A\rangle_0 = h\beta[\langle A(t)B(t)\rangle_{\rm eq} - \langle A(t)B(0)\rangle_{\rm eq}]. \tag{20}$$

In its integrated version, the equilibrium FDT reduces to a simple linear relation between the integrated response and its conjugated correlation function.

2.4. Detailed balance

Whether we interpret a stochastic dynamics as deriving from an underlying microscopic description fulfilling the laws of classical or quantum mechanics or not, DB constitutes the key symmetry of equilibrium. A system is said to satisfy DB if, at stationarity, any microscopic process is balanced by the reversed one. It can thus be formally written as

$$P(\mathbf{\Gamma}_f, t_f|\mathbf{\Gamma}_i, t_i)\Psi_0(\mathbf{\Gamma}_i) = P(\boldsymbol{\varepsilon}\,\mathbf{\Gamma}_i, t_f|\boldsymbol{\varepsilon}\mathbf{\Gamma}_f, t_i)\Psi_0(\boldsymbol{\varepsilon}\mathbf{\Gamma}_f) \tag{21}$$

for any pair of states $(\mathbf{\Gamma}_i, \mathbf{\Gamma}_f)$ and at any times (t_i, t_f). By setting $t_i = t_f$ in equation (21) we get:

$$\Psi_0(\mathbf{\Gamma}) = \Psi_0(\boldsymbol{\varepsilon}\mathbf{\Gamma}), \quad \forall\ \mathbf{\Gamma} \in \mathcal{M}. \tag{22}$$

It follows that the mean value of any *current-like* observable, i.e. $A(\boldsymbol{\varepsilon}\mathbf{\Gamma}) = -A(\mathbf{\Gamma})$ must be zero if DB is fulfilled[6]. Rather than being a symmetry at the level of single trajectories, DB is formulated in equation (22) as a symmetry property of the steady-state distribution $\Psi_0(\mathbf{\Gamma})$. Actually, a necessary and sufficient condition for DB to hold is the absence of irreversible fluxes in steady conditions [75]:

$$\mathrm{DB} \Leftrightarrow \boldsymbol{J}_0^{\rm irr} = 0 \tag{23}$$

where $\boldsymbol{J}_0^{\rm irr} = \{\mathcal{A}_i^{\rm irr}(\mathbf{\Gamma})\Psi_0(\mathbf{\Gamma}) - D_i\partial_i\Psi_0(\mathbf{\Gamma})\}_{i=1}^N$. This means that reversible steady-state fluxes are not constrained by DB. As such, in the presence of odd (momentum-like) variables as in the underdamped description, reversible steady-state fluxes can be present in a system fulfilling DB (although physical currents must have all zero ensemble averages, see equation (22)) [76, 77].

The absence of irreversible fluxes, equation (23), imposes the so-called

[6] Note that here we are referring to physical currents and not to the probability current J_i in equation (8). Indeed, equation (22) implies $\langle A\rangle_0 = \int \mathrm{d}\mathbf{\Gamma} A(\boldsymbol{\varepsilon}\mathbf{\Gamma})\Psi_0(\boldsymbol{\varepsilon}\mathbf{\Gamma}) = -\int \mathrm{d}\mathbf{\Gamma} A(\mathbf{\Gamma})\Psi_0(\mathbf{\Gamma}) = -\langle A\rangle_0$ if $A(\boldsymbol{\varepsilon}\mathbf{\Gamma}) = -A(\mathbf{\Gamma})$.

J. Stat. Mech. (2021) 043201

'thermodynamic curvature' [78] (the curl in N dimensions) of the irreversible drift to vanish:

$$D_i^{-1}\partial_j \mathcal{A}_i^{\mathrm{irr}}(\mathbf{\Gamma}) - D_j^{-1}\partial_i \mathcal{A}_j^{\mathrm{irr}}(\mathbf{\Gamma}) = 0. \tag{24}$$

Note that the latter expression provides a remarkable advantage over the definition in equation (23) as it expresses DB as a purely geometrical property of the drift and diffusion terms. In order to attest DB in equation (23) one needs to solve for $\Psi_0(\mathbf{\Gamma})$ explicitly, which in turn is typically prohibitive when DB does not hold. Whenever the thermodynamic curvature vanishes, one can derive a steady solution $\Psi_0(\mathbf{\Gamma})$ by direct integration. No such a procedure exists if DB is broken, and no prescribed functional form of the multivariate Fokker–Planck equation can be derived in general.

2.5. The fluctuation–dissipation theorem

We focus now on the linear response of a system initially prepared in a steady-state. At $t = 0$ we apply an infinitesimal perturbation (generally time dependent) to the drift vector $\mathcal{A} \to \mathcal{A} + \delta\mathcal{A}(t)$. The evolution equation (2) now reads:

$$\partial_t \Psi(\mathbf{\Gamma}, t) = \Omega\Psi(\mathbf{\Gamma}, t) = \left[\Omega_0(\mathbf{\Gamma}) + \Omega_{\mathrm{ext}}(\mathbf{\Gamma}, t)\right]\Psi(\mathbf{\Gamma}, t) \tag{25}$$

where

$$\Omega_{\mathrm{ext}}(\mathbf{\Gamma}, t)\Psi(\mathbf{\Gamma}, t) = -\nabla\cdot\left[\delta\mathcal{A}(\mathbf{\Gamma}, t)\Psi(\mathbf{\Gamma}, t)\right] \tag{26}$$

accounts for the perturbation. Equation (25) is equivalent to the integral equation [1]:

$$\Psi(\mathbf{\Gamma}, t) = \mathrm{e}^{\Omega_0(\mathbf{\Gamma})t}\Psi(\mathbf{\Gamma}, 0) + \int_0^t \mathrm{d}s\, \mathrm{e}^{\Omega_0(\mathbf{\Gamma})(t-s)}\Omega_{\mathrm{ext}}(\mathbf{\Gamma}, s)\Psi(\mathbf{\Gamma}, s), \tag{27}$$

as it is seen by differentiating both side of equation (27) by t. In the case of a constant perturbation, and to first order in $\delta\mathcal{A}$, equation (27) reduces to:

$$\Psi(\mathbf{\Gamma}, t) = \Psi_0(\mathbf{\Gamma}) + \int_0^t \mathrm{d}t'\, \mathrm{e}^{\Omega_0(\mathbf{\Gamma})t'}\Omega_{\mathrm{ext}}(\mathbf{\Gamma})\Psi_0(\mathbf{\Gamma}). \tag{28}$$

For an observable A, we then find the so-called Agarwal FDR [9]:

$$\langle A\rangle_t - \langle A\rangle_0 = \int_0^t \mathrm{d}s\langle A(s)B(0)\rangle_0, \qquad B(0) \equiv \frac{\Omega_{\mathrm{ext}}(\mathbf{\Gamma})\Psi_0(\mathbf{\Gamma})}{\Psi_0(\mathbf{\Gamma})} \tag{29}$$

where B is the observable conjugated to the perturbation $\delta\mathcal{A}$. Alternatively, using

J. Stat. Mech. (2021) 043201

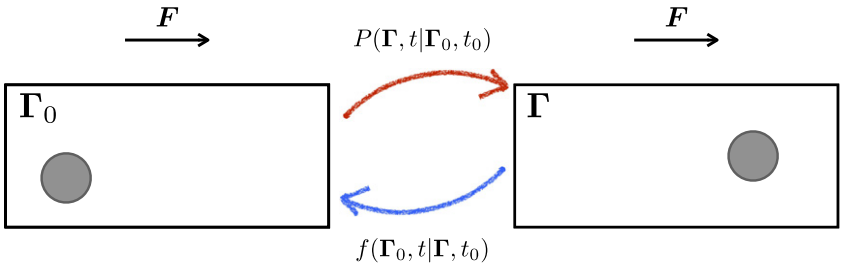

Figure 1. Illustration of the conditional probability f introduced to quantify violations of DB. A force $\boldsymbol{F}$ is applied to a colloidal particle, which favors motion towards the right, and thus $f(\mathbf{\Gamma}_0, t|\mathbf{\Gamma}, t_0) < P(\mathbf{\Gamma}, t|\mathbf{\Gamma}_0 t_0)$. In the absence of external drive, DB is satisfied and there is no preference to move right or left, thus $f(\mathbf{\Gamma}_0, t|\mathbf{\Gamma}, t_0) = P(\mathbf{\Gamma}, t|\mathbf{\Gamma}_0 t_0)$.

equation (26) into equation (29) the integrated response can be expressed as:

$$\langle A\rangle_t - \langle A\rangle_0 = -\int_0^t \mathrm{d}s \langle A(s)\left[\nabla\cdot\delta\mathcal{A} + \delta\mathcal{A}\cdot\nabla\,\log\,\Psi_0\right](0)\rangle_0 \tag{30}$$

which resembles the FDR as originally appeared in the context of dynamical systems [79]. At this level, no equilibrium hypothesis has been made and, as such, equations (29) and (30) remain valid for any NESS provided it is described by a Markovian evolution (equations (2)–(4)). For an equilibrium system fulfilling DB, i.e. $\Psi_0(\mathbf{\Gamma}) = \Psi_{\rm eq}(\mathbf{\Gamma})$, the conjugated observable B can be computed explicitly.

As an illustration, we consider a one-dimensional overdamped Brownian particle in contact with a thermal bath at temperature T that we perturb by applying, at $t = 0$, a constant external force h. In this case $\Omega_0 = \mu\partial_x[k_{\rm B}T\partial_x - F]$ with $F \equiv -U'(x)$ the conservative force acting on the particle, μ the particle mobility, and $\Omega_{\rm ext} = -\mu h\partial_x$. From equation (29) we immediately obtain

$$\langle A\rangle_t - \langle A\rangle_0 = -\mu\beta h\int_0^t \mathrm{d}s\langle A(s)F(0)\rangle_0. \tag{31}$$

If we now replace $\mu F = \dot{x} - \sqrt{2\mu k_{\rm B}T}\xi$ and use the fact that $R_A(t) = \sqrt{\beta\mu/2}\langle A(t)\xi(0)\rangle$ [80] we find the following familiar form of the FDT for Brownian suspensions

$$\langle A\rangle_t - \langle A\rangle_0 = \beta h\int_0^t \mathrm{d}s\langle A(s)\dot{x}(0)\rangle_0. \tag{32}$$

From this relation, it is straightforward to derive *Green–Kubo expressions* for transport coefficients, such as the mobility and diffusivity. For instance, by choosing $A \equiv \dot{x}$ we find the well-known expressions

$$\mu = \lim_{t\to\infty}\frac{\langle\dot{x}\rangle_t}{h} = \beta\int_0^\infty \mathrm{d}s\langle\dot{x}(s)\dot{x}(0)\rangle_0, \qquad D = \int_0^\infty \mathrm{d}s\langle\dot{x}(s)\dot{x}(0)\rangle_0. \tag{33}$$

J. Stat. Mech. (2021) 043201

3. Quantifying the violations of detailed balance

3.1. The time-reversal operator

In order to quantify the breakdown of DB let introduce the conditional probability to go from $\mathbf{\Gamma}$ to $\mathbf{\Gamma}_0$ forward in time (see figure 1):

$$f(\mathbf{\Gamma}_0, t|\mathbf{\Gamma}, t_0) \equiv \frac{\Psi_0(\varepsilon\mathbf{\Gamma})}{\Psi_0(\mathbf{\Gamma}_0)} P(\varepsilon\mathbf{\Gamma}_0, t|\varepsilon\mathbf{\Gamma}, t_0). \tag{34}$$

Whenever DB holds $f(\mathbf{\Gamma}_0, t|\mathbf{\Gamma}, t_0) = P(\mathbf{\Gamma}, t|\mathbf{\Gamma}_0 t_0)$. When DB is violated, such identification is no longer valid. Nevertheless, it is still interesting to look at the evolution equation for $f(\mathbf{\Gamma}_0, t|\mathbf{\Gamma}, t_0)$ as it allows to quantify the breakdown of DB. Following the approach of [81] we start by writing the backward evolution equation, equation (6), for $P(\varepsilon\mathbf{\Gamma}_0, t|\varepsilon\mathbf{\Gamma}, t_0)$:

$$\partial_t P\left(\varepsilon\mathbf{\Gamma}_0, t|\varepsilon\mathbf{\Gamma}, t_0\right) = \sum_i \varepsilon_i \mathcal{A}_i\left(\varepsilon\mathbf{\Gamma}\right) \partial_i P\left(\varepsilon\mathbf{\Gamma}_0, t|\varepsilon\mathbf{\Gamma}, t_0\right) + D_i \partial_i^2 P\left(\varepsilon\mathbf{\Gamma}_0, t|\varepsilon\mathbf{\Gamma}, t_0\right). \tag{35}$$

By multiplying both sides of the latter equation by $\Psi_0(\varepsilon\mathbf{\Gamma})/\Psi_0(\mathbf{\Gamma}_0)$ and using the definition of $f(\mathbf{\Gamma}_0, t|\mathbf{\Gamma}, t_0)$, we obtain

$$\begin{aligned}\partial_t f(\mathbf{\Gamma}_0, t|\mathbf{\Gamma}, t_0) = &\sum_i \varepsilon_i \partial_i \left(\mathcal{A}_i(\varepsilon\mathbf{\Gamma}) \times f\right) - \varepsilon_i f \times \frac{\partial_i(\mathcal{A}_i(\varepsilon\mathbf{\Gamma})\Psi_0(\varepsilon\mathbf{\Gamma}))}{\Psi_0(\varepsilon\Gamma)} \\ &+ D_i \left\{\partial_i^2 f - 2\partial_i \left(f \times \frac{\partial_i \Psi_0(\varepsilon\mathbf{\Gamma})}{\Psi_0(\varepsilon\mathbf{\Gamma})}\right) + f \times \frac{\partial_i^2 \Psi_0(\varepsilon\mathbf{\Gamma})}{\Psi_0(\varepsilon\mathbf{\Gamma})}\right\}\end{aligned} \tag{36}$$

where, for simplicity, we wrote $f(\mathbf{\Gamma}_0, t|\mathbf{\Gamma}, t_0) \equiv f$. Let now split the drift vector into its reversible and irreversible parts to rewrite the first term on the *rhs* of equation (36) as

$$\begin{aligned}\sum_i \varepsilon_i \partial_i (\mathcal{A}_i(\varepsilon\mathbf{\Gamma}) \times f) &= \sum_i \partial_i ([-\mathcal{A}_i^{\mathrm{rev}}(\mathbf{\Gamma}) + \mathcal{A}_i^{\mathrm{irr}}(\mathbf{\Gamma})] \times f) \\ &= \sum_i \left(-\partial_i [\mathcal{A}_i(\mathbf{\Gamma}) \times f] + 2\partial_i \left[\mathcal{A}_i^{\mathrm{irr}}(\mathbf{\Gamma}) \times f\right]\right).\end{aligned} \tag{37}$$

By substituting equation (37) back into equation (36) and using the property $\Omega_0(\varepsilon\mathbf{\Gamma})\Psi_0(\varepsilon\mathbf{\Gamma}) = 0$ we finally get

$$\partial_t f = \Omega_0(\mathbf{\Gamma}) f + \sum_i \partial_i \left[f \times \left(2\mathcal{A}_i^{\mathrm{irr}}(\mathbf{\Gamma}) - 2D_i \partial_i \log \Psi_0(\varepsilon\mathbf{\Gamma})\right]\right). \tag{38}$$

When DB holds, the Fokker–Planck equation is recovered: since the initial condition for $f(\mathbf{\Gamma}_0, t|\mathbf{\Gamma}, t_0)$ and $P(\mathbf{\Gamma}, t|\mathbf{\Gamma}_0, t_0)$ is the same by definition, it follows that $f = P(\mathbf{\Gamma}, t|\mathbf{\Gamma}_0, t_0)$, as expected. We introduce the *time-reversal operator*, $\bar{\Omega}_0$, defined as the Markov generator of $f(\mathbf{\Gamma}_0, t|\mathbf{\Gamma}, t_0)$:

$$\partial_t f(\mathbf{\Gamma}_0, t|\mathbf{\Gamma}, t_0) \equiv \bar{\Omega}_0(\mathbf{\Gamma}) f(\mathbf{\Gamma}_0, t|\mathbf{\Gamma}, t_0). \tag{39}$$

J. Stat. Mech. (2021) 043201

By comparing equation (38) with equation (39) and after some manipulations, we find an explicit expression for $\bar{\Omega}_0$ in terms of the Fokker–Planck operator:

$$\bar{\Omega}_0(\mathbf{\Gamma}) f(\mathbf{\Gamma}_0, t|\mathbf{\Gamma}, t_0) = \Psi_0(\boldsymbol{\varepsilon}\mathbf{\Gamma})\Omega_0^\dagger(\boldsymbol{\varepsilon}\mathbf{\Gamma}) \left[f(\mathbf{\Gamma}_0, t|\mathbf{\Gamma}, t_0)\Psi_0(\boldsymbol{\varepsilon}\mathbf{\Gamma})^{-1} \right]. \tag{40}$$

Notably, its adjoint

$$\bar{\Omega}_0^\dagger(\mathbf{\Gamma}) = \Psi_0(\boldsymbol{\varepsilon}\mathbf{\Gamma})^{-1}\Omega_0(\boldsymbol{\varepsilon}\mathbf{\Gamma})\Psi_0(\boldsymbol{\varepsilon}\mathbf{\Gamma}) \tag{41}$$

has a precise dynamical meaning: it evolves the observable A along time-reversed paths, such that

$$\langle A \rangle_{(-t)} \equiv \int \mathrm{d}\mathbf{\Gamma} \Psi_0(\mathbf{\Gamma}) \mathrm{e}^{\bar{\Omega}_0^\dagger t} A(\mathbf{\Gamma}), \quad t > 0 \tag{42}$$

or, for the two-time correlation function

$$\langle B(0)A(-t) \rangle_0 \equiv \int \mathrm{d}\mathbf{\Gamma} \Psi_0(\mathbf{\Gamma}) B(\mathbf{\Gamma}) \mathrm{e}^{\bar{\Omega}_0^\dagger t} A(\mathbf{\Gamma}), \quad t > 0. \tag{43}$$

As such, it generalizes the standard formulation [19] to the intricate case of variables with a different parity under time-reversal. Equations (40) and (41) are consistent with the expressions presented in [77] and avoid any ambiguity concerning the parity transformation under time-reversal as they were directly derived from the dynamical equation (35). From equation (41), we find the following expression quantifying the breakdown of DB in terms of irreversible probability fluxes:

$$\frac{\Omega_0^\dagger(\mathbf{\Gamma}) - \bar{\Omega}_0^\dagger(\mathbf{\Gamma})}{2} = \frac{\boldsymbol{\varepsilon}\boldsymbol{J}_0^{\mathrm{irr}}(\boldsymbol{\varepsilon}\mathbf{\Gamma})}{\Psi_0(\boldsymbol{\varepsilon}\mathbf{\Gamma})} \cdot \nabla \equiv \boldsymbol{\varepsilon}\boldsymbol{\mathcal{V}}^{\mathrm{irr}}(\boldsymbol{\varepsilon}\mathbf{\Gamma}) \cdot \nabla. \tag{44}$$

It immediately follows that DB holds if and only if the time-reversal operator equals the Fokker–Planck operator

$$\mathrm{DB} \quad \Leftrightarrow \quad \Omega_0(\mathbf{\Gamma}) = \bar{\Omega}_0(\mathbf{\Gamma}). \tag{45}$$

From equation (44) the explicit expression for the time-reversed dynamics follows:

$$\bar{\Omega}_0(\mathbf{\Gamma}) = -\sum_i \partial_i \left[\left(\mathcal{A}_i^{\mathrm{rev}}(\mathbf{\Gamma}) - \mathcal{A}_i^{\mathrm{irr}}(\mathbf{\Gamma}) + 2\partial_i \log \Psi(\boldsymbol{\varepsilon}\mathbf{\Gamma}) \right) + D_i \partial_i \right]. \tag{46}$$

By comparing it with the original dynamics in equation (4) we see that, in order to generate time-reversed trajectories with the same probability as the original trajectories, one needs not just to reverse the irreversible drift $\mathcal{A}_i^{\mathrm{irr}}$ but also to introduce a new term depending on Ψ_0 evaluated at the $\{\boldsymbol{\varepsilon}\mathbf{\Gamma}\}$. This highlights the importance of keeping track of the parity transformation.

Notably the reversible drift is not modified in the time-reversed dynamics. Since DB only constraints irreversible fluxes to vanish, it is thus expected that its breakdown only concerns their presence. The expression equation (44) derived here is general. It quantifies the breakdown of DB in systems with both odd and even variables, such as collections of interacting particles with inertia [82] or noisy RLC circuits [83].

J. Stat. Mech. (2021) 043201

3.2. General constraints on non-equilibrium steady-states

It is useful to notice that the irreversible steady-state velocity quantifying the non-equilibrium character of the dynamics, or breakdown of DB, is directly related to the house-keeping entropy production, i.e. the entropy being produced in a NESS. For a stochastic dynamics described by the Fokker–Planck equation (2), the house-keeping entropy production rate is expressed as [84]:

$$\langle \dot{S}_{\rm hk} \rangle_t = k_{\rm B} \sum_i \int {\rm d}\mathbf{\Gamma} D_i^{-1} \Psi(\mathbf{\Gamma}, t) \left(\varepsilon \mathcal{V}_i^{\rm irr}(\varepsilon \mathbf{\Gamma}) \right)^2 \geqslant 0 \tag{47}$$

where we recognize inside the integral the steady-state velocity quantifying the violations of DB in equation (44). The house-keeping term constitutes the entropy production necessary to sustain a NESS and it is positive as soon as DB is broken. It may be compared with the standard formula for the total entropy production [84–86]:

$$\langle \dot{S}_{\rm tot} \rangle_t = k_{\rm B} \sum_i \int {\rm d}\mathbf{\Gamma} D_i^{-1} \Psi(\mathbf{\Gamma}, t) \left(\mathcal{V}_i^{\rm irr}(\mathbf{\Gamma}, t) \right)^2 \geqslant 0. \tag{48}$$

embedding contributions coming from the breakdown of DB as well as from any time-dependent processes (the so-called excess entropy production). By definition, in stationary conditions, $\dot{S}_{\rm hk}$ must reduce to the total entropy change $\dot{S}_{\rm tot}$, since no time-dependent contribution is present. While this is obvious for even variables, in the presence of odd variables this observation provides a *general constraint on the steady-state velocity* such that:

$$\left[\mathcal{V}_i^{\rm irr}(\varepsilon \mathbf{\Gamma}) \right]^2 = \left[\mathcal{V}_i^{\rm irr}(\mathbf{\Gamma}) \right]^2. \tag{49}$$

As already mentioned, Ψ_0 does not necessary have a given parity under time-reversal. Nevertheless (without lack of generality) we may write it as:

$$\Psi_0 = \mathcal{N} \exp(-\Phi) = \mathcal{N} \exp\left[-(\Phi_+ + \Phi_-) \right] \tag{50}$$

where the generalized potential Φ was decomposed into an even and odd contribution: $\Phi(\varepsilon \mathbf{\Gamma})_\pm = \pm \Phi(\mathbf{\Gamma})$[7]. Equation (49) then imposes:

$$\left[\mathcal{A}_i^{\rm irr} + D_i \partial_i \Phi_+ \right] D_i \partial_i \Phi_- = 0 \tag{51}$$

which, in order to be verified, requires one of the two following conditions:

$$\partial_i \Phi_+(\mathbf{\Gamma}) = -D_i^{-1} \mathcal{A}_i^{\rm irr}(\mathbf{\Gamma}) \quad \text{or} \quad \partial_i \Phi_- = 0. \tag{52}$$

Note that if both constraints are fulfilled *simultaneously* then the system obeys DB (see equation (23)). Equation (49) 'relaxes' one of the two constraints, allowing to distinguish two different ways of breaking DB, namely

$$\Phi_-(\mathbf{\Gamma}) = 0 \quad \text{but} \quad \partial_i \Phi_+(\mathbf{\Gamma}) \neq -D_i^{-1} \mathcal{A}_i^{\rm irr}(\mathbf{\Gamma}) \tag{53}$$

[7] The generalized potential is assumed to be a differential function of the dynamic variables.

J. Stat. Mech. (2021) 043201

or

$$\partial_i \Phi_+(\mathbf{\Gamma}) = -D_i^{-1}\mathcal{A}_i^{\rm irr}(\mathbf{\Gamma}) \quad \text{but} \quad \Phi_-(\mathbf{\Gamma}) \neq 0. \tag{54}$$

Any non-equilibrium system with only even variables falls into the first category. On the contrary, systems with odd variables, such as the driven underdamped Brownian particle considered below, or the AOUP that we treat in detail in section 4.2, fall into the second category. For systems with odd variables, it follows that the irreversible flux is directly related (only) to the odd part of the NESS distribution, such that:

$$\mathcal{V}_i^{\rm irr}(\mathbf{\Gamma}) = D_i \partial_i \Phi_-(\mathbf{\Gamma}). \tag{55}$$

Equations (49), (52)–(54) constrain any NESS by relating explicitly the nature of the irreversible fluxes, responsible for the breakdown of DB, with the symmetry of the NESS distribution Ψ_0 under time-reversal.

In order to illustrate these results, we now consider a paradigmatic non-equilibrium model: a one-dimensional particle in a periodic potential U, coupled to a thermal bath and driven out-of-equilibrium by a non-conservative force F[8]. The Langevin equation governing the dynamics of such system is (see figure 1)

$$\dot{x}(t) = p(t) \qquad \dot{p}(t) = -U'(x) + F - \gamma p(t) + \sqrt{2\gamma k_{\rm B} T}\xi(t), \tag{56}$$

where $\gamma = 1/\mu$ is the friction coefficient. The direct application of equation (54) in this simple case yields

$$\Phi_+(x,p) = \beta\frac{p^2}{2} + \Lambda(x) \tag{57}$$

$$\mathcal{V}_p^{\rm irr}(x,p) = \gamma k_{\rm B} T \frac{\partial \Phi_-}{\partial p} \tag{58}$$

where $\Lambda(x)$ is any derivable function of x. Thus, the even part of the generalized potential must be of the latter form, meaning that even terms p^{2n} with $n > 1$ are strictly forbidden in the NESS distribution. As expected, the dependence of $\Psi(x,p)$ on the positions, which are associated to reversible fluxes, is not constrained at all. Instead, irreversible fluxes are given by (gradients of) the odd part of the generalized potential. In the limit $U \to 0$ [84], the NESS distribution of this simple model can be computed analytically and consists in a 'tilted' Gaussian distribution in p and a uniform distribution in x, such as:

$$\Psi_0(x,p) = \mathcal{N} \exp\left[-\frac{(p - \langle p\rangle)^2}{2k_{\rm B}T}\right] \tag{59}$$

where $\langle p\rangle = F/\gamma$ is the mean velocity of the particle. Combining equation (59) with equation (58) we find $\Phi_- = -pF/\gamma k_{\rm B}T$ and consequently:

$$\mathcal{V}_p^{\rm irr} = -F \tag{60}$$

$$\langle \dot{S}_{\rm hk}\rangle_0 = \frac{F^2}{\gamma T} > 0 \tag{61}$$

[8] Periodicity is needed in order to guarantee the existence of a steady-state.

J. Stat. Mech. (2021) 043201

which reveals the non-equilibrium character of this system and the nature of its steady-state velocity: it is proportional to the particle net current, resulting in a positive housekeeping entropy production.

3.3. Extended fluctuation–dissipation relations

We now consider a system initially prepared in a NESS with probability distribution Ψ_0. At time $t = 0$, we apply a perturbation such that $\mathcal{A} \to \mathcal{A} + \delta\mathcal{A}$. Although the Agarwal FDR equation (29) remains valid (and exact), the lack of knowledge on Ψ_0 does not allow us to derive an explicit expression of the response function in terms of NESS correlations. To move further, an option would be to provide a reliable scheme to approximate Ψ_0. We follow this strategy in the next section when dealing with active particles. However, at a more fundamental level, the Agarwal expression can be used to derive alternative FDRs that explicitly relate the non-equilibrium response of the system to the symmetry properties of the NESS. For instance, alternative FDRs can be written in a way that explicitly relates the non-equilibrium response of the system to the symmetry properties of Ψ_0.

We start by rewriting the Agarwal FDR in terms of the decomposition in equation (50) as:

$$\langle A\rangle_t - \langle A\rangle_0 = \sum_i \left\{ -\int_0^t \mathrm{d}s \langle A(s)(\partial_i \delta\mathcal{A}_i)(0)\rangle_0 + \int_0^t \mathrm{d}s \langle A(s)(\delta\mathcal{A}_i \partial_i \Phi_+)(0)\rangle_0 + \int_0^t \mathrm{d}s \langle A(s)(\delta\mathcal{A}_i \partial_i \Phi_-)(0)\rangle_0 \right\}. \tag{62}$$

We then make use of the relations, derived in the previous section, between the generalized potential and the steady-state velocity to write

$$\langle A\rangle_t - \langle A\rangle_0 = \sum_i \left\{ -\int_0^t \mathrm{d}s \langle A(s)(\partial_i \delta\mathcal{A}_i)(0)\rangle_0 + D_i^{-1} \int_0^t \mathrm{d}s \langle A(s)(\delta\mathcal{A}_i \mathcal{A}_i^{\mathrm{irr}})(0)\rangle_0 + D_i^{-1} \int_0^t \mathrm{d}s \langle A(s)(\delta\mathcal{A}_i \mathcal{V}_i^{\mathrm{irr}})(0)\rangle_0 \right\} \tag{63}$$

which is valid both for overdamped and underdamped systems in the presence of any type of external perturbation $\delta\mathcal{A}$. As such, it generalizes the FDR appearing in [12, 19, 87], derived for an overdamped particle subjected to a conservative perturbation $\delta\mathcal{A} \equiv -\mu\partial_x U$, for which it reduces to:

$$\langle A\rangle_t - \langle A\rangle_0 = D_x^{-1}\mu \int_0^t \mathrm{d}s \left(\left\langle A(s)\dot{U}(0) \right\rangle_0 - \left\langle A(s)\partial_x U(0)\mathcal{V}_x^{\mathrm{irr}}(0) \right\rangle_0 \right). \tag{64}$$

If we consider the perturbation to be a constant force of amplitude h, such that $U \equiv hx$, equation (64) gives the non-equilibrium extension of the FDT as derived by Speck and Seifert in [12]. The term comprising the irreversible steady-state velocity in

J. Stat. Mech. (2021) 043201

equations (63) and (64) is responsible for the non-equilibrium character of the dynamics. Thus we notice, by comparing equations (62) and (63), that in the presence of odd variables full knowledge on Ψ_0 is not required to determine the response, but only on its odd part under time-reversal.

4. Application to active particles

We consider in this section ABP and AOUP. We first discuss the non-equilibrium nature of their dynamics to then derive extended FDR. First, we show that ABP generically breaks DB and how, under a Markovian approximation, DB can be restored resulting in an effective FDT (where inter-particle interactions are replaced by an effective force). Secondly, we recall the results of Fodor *et al* [31] and Bonilla [33], showing that an equilibrium regime, fulfilling DB, exists for AOUP. Taking advantage of the Chapman–Enskog expansion in [33], we derive an extended FDR for AOUP taking into account terms that, as we show, explicitly break DB.

4.1. Active Brownian particles

4.1.1. The model. We consider now N overdamped ABP moving in the 2D plane. They self-propel with a constant velocity v_0, along their orientation $\boldsymbol{n}_i = (\cos(\theta_i),\ \sin(\theta_i))$ and obey the following set of coupled Langevin equations

$$\dot{\boldsymbol{r}}_i(t) = \mu_0 \boldsymbol{F}_i + v_0 \boldsymbol{n}_i(t) + \boldsymbol{\xi}_i(t), \qquad \dot{\theta}_i(t) = \nu_i(t) \tag{65}$$

where $\boldsymbol{F}_i = -\partial U/\partial \boldsymbol{r}_i$ accounts for all (inter-particle or external) potential forces, μ_0 is the single particle mobility and $\boldsymbol{\xi}$ and ν are zero-mean Gaussian noises verifying $\langle \boldsymbol{\xi}_i(t)\boldsymbol{\xi}_j(t')\rangle = 2\mu_0 k_\mathrm{B} T \delta_{ij}\delta(t-t')\mathbf{1}$ and $\langle \nu_i(t)\nu_j(t')\rangle = 2D_\theta \delta_{ij}\delta(t-t')$, $\mathbf{1}$ being the 2×2 identity matrix. It follows that

$$\langle \boldsymbol{n}_i(t)\cdot \boldsymbol{n}_i(0)\rangle = \mathrm{e}^{-D_\theta t}, \tag{66}$$

defining a persistence time $\tau = 1/D_\theta$. The generator, or Fokker–Planck operator, corresponding to this Langevin dynamics is

$$\Omega_0(\boldsymbol{\Gamma}) = \sum_i \frac{\partial}{\partial \boldsymbol{r}_i}\left(\mu_0 k_\mathrm{B} T \frac{\partial}{\partial \boldsymbol{r}_i} - \mu_0 \boldsymbol{F}_i - v_0 \boldsymbol{n}_i\right) + D_\theta \sum_i \frac{\partial^2}{\partial \theta_i^2} \tag{67}$$

where $\boldsymbol{\Gamma} \equiv \{\boldsymbol{r}_i, \theta_i\}$: all dynamical variables in ABP are considered even under time-reversal.

4.1.2. Non-equilibrium character and non-interacting regime. For ABP, DB is fulfilled if and only if (see equation (23))

$$\begin{cases} \left(\mu_0 \boldsymbol{F}_i + v_0 \boldsymbol{n}_i\right)\Psi_0(\boldsymbol{\Gamma}) = \mu_0 k_\mathrm{B} T \dfrac{\partial}{\partial \boldsymbol{r}_i}\Psi_0(\boldsymbol{\Gamma}) \\ \dfrac{\partial}{\partial \theta_i}\Psi_0(\boldsymbol{\Gamma}) = 0. \end{cases} \tag{68}$$

J. Stat. Mech. (2021) 043201

These two equations cannot simultaneously hold due to the self-propulsion term, and therefore, ABP generically break DB. Since the forces $\boldsymbol{F}_i$ are conservative, the first condition can be integrated to get $\log \Psi_0 \sim -\beta[U - \sum_i v_0 \boldsymbol{n}_i \cdot \boldsymbol{r}_i/\mu_0]$. However, the second condition imposes Ψ_0 to be a function of positions only, which is inconsistent with the first condition because of the term in $\boldsymbol{n}_i \cdot \boldsymbol{r}_i$. In the passive case, $v_0 \to 0$, DB is recovered together with the standard Boltzmann distribution.

An illustrative example for which we can explicitly compute the steady-state velocity is a free ABP. The Fokker–Plank generator reads $\Omega_0 = \left[\partial_{\boldsymbol{r}}(\mu_0 k_{\mathrm{B}} T \partial_{\boldsymbol{r}} - v_0 \boldsymbol{n}) + D_\theta \frac{\partial^2}{\partial \theta^2}\right]$ and a stationary solution of the Fokker–Planck equation is $\Psi_0(\boldsymbol{r}, \theta) = \rho_0/2\pi$ [88]. The steady-state velocity corresponding to this homogeneous NESS is

$$\boldsymbol{\mathcal{V}}^{\mathrm{irr}}(t) = (\mathcal{V}_{\boldsymbol{r}}^{\mathrm{irr}}, \mathcal{V}_\theta^{\mathrm{irr}}) = (v_0 \boldsymbol{n}(t), 0). \tag{69}$$

Since the system is overdamped, there are no reversible fluxes. In order to apply the extended FDR, we consider a constant force perturbation h applied along the x-axis. By choosing $A = \dot{x}$ and $U \equiv hx$ in equation (64) we obtain

$$\frac{\langle \dot{x} \rangle_t}{h} = \beta \int_0^t \mathrm{d}s \langle \dot{x}(s) \dot{x}(0) \rangle_0 - \beta v_0 \int_0^t \mathrm{d}s \langle \dot{x}(s) \cos\theta(0) \rangle_0 \tag{70}$$

resulting in the following extended Stokes–Einstein relation in the long time regime

$$D/\mu_0 = k_{\mathrm{B}} T + \frac{v_0^2}{2 D_\theta \mu_0}. \tag{71}$$

Therefore, non-interacting ABP fulfill the Stokes–Einstein relation with an effective temperature

$$T_{\mathrm{eff}}/T = 1 + \frac{v_0^2}{2 D_\theta \mu_0 k_{\mathrm{B}} T}. \tag{72}$$

It is worth noting here that, although ABP generically break DB in a fundamental way (and do not allow for a zero current steady-state solution), in the non-interacting limit they admit a NESS which fulfills the Stokes–Einstein relation.

4.1.3. Interacting regime: an effective Markovian description. As evidenced by the extended FDR equation (63), the response of a non-equilibrium system is not completely determined by NESS correlations of measurable observables, but also depends on the specific form of its steady-state velocity. In order to go beyond the non-interacting limit considered above and establish explicit FDR for interacting ABP one can approximate the dynamics by an effective equilibrium one that fulfills DB. Such kind of approximation has been used for ABP and also AOUP and comes under different names, the most usual ones being unified colored noise and Fox approximation [30, 89–91]. In both cases, the steady-state distribution of the new dynamics does not correspond to the equilibrium Boltzmann distribution in terms of the energy function of the original dynamics, but a 'Boltzmann-like' distribution in terms of an effective energy function. Despite the non-Boltzmann character of the steady-state distribution resulting from these approaches, all the difficulties associated with the absence of DB are lifted and one can readily derive an FDR by direct application of the general results presented in the previous section.

J. Stat. Mech. (2021) 043201

To be more specific, we turn now into the analysis of interacting ABP within the Fox approximation [92, 93], as we previously presented in [69]. The starting point is to integrate out the angular variables appearing in the ABP dynamics. The equations of motion (65) are then approximated by

$$\dot{\boldsymbol{r}}_i(t) = \mu_0 \boldsymbol{F}_i + \boldsymbol{\eta}_i(t) \tag{73}$$

where the noise $\boldsymbol{\eta}_i$ is approximately Gaussian with zero mean and variance $\langle \boldsymbol{\eta}_i(t)\boldsymbol{\eta}_j(s)\rangle = (2\mu_0 k_{\mathrm{B}}T\delta(t-s) + v_0^2 \mathrm{e}^{-|t-s|/\tau}/2)\delta_{ij}\mathbf{1}$. The ABP dynamics in the reduced configuration space $\tilde{\boldsymbol{\Gamma}} \equiv \{\boldsymbol{r}_i\}$ is approximated by an effective Fokker–Planck dynamics generated by the operator:

$$\Omega_0^M(\tilde{\boldsymbol{\Gamma}}) = \sum_\alpha \partial_\alpha \left(\sum_\beta \partial_\beta \mathcal{D}_{\beta\alpha}(\tilde{\boldsymbol{\Gamma}}) - \mu_0 F_\alpha(\tilde{\boldsymbol{\Gamma}}) \right) \tag{74}$$

where $\mathcal{D} \equiv \{\mathcal{D}_{\beta\alpha}\}$ is an effective non-diagonal $2N \times 2N$ diffusivity tensor and Greek indices run over the spatial coordinates and the particle labels. To first order in $\tau\mu_0\partial_\alpha F_\beta$, it reads

$$\mathcal{D}_{\alpha\beta}(\tilde{\boldsymbol{\Gamma}}) = \mu_0 k_{\mathrm{B}}T\delta_{\alpha\beta} + \frac{v_0^2\tau}{2}\left(\delta_{\alpha\beta} + \tau\mu_0\partial_\alpha F_\beta\right). \tag{75}$$

Note that the Fox approximation is meaningful only when $|\tau\mu_0\partial_\alpha F_\beta| < 1$.

In view of the discussion in section 2.4, we now show that this effective dynamics fulfills DB by making use of equation (24). The first step is to write the condition of DB equation (23) for the stationary probability density $\Psi_0(\tilde{\boldsymbol{\Gamma}})$:

$$\sum_\beta \mathcal{D}_{\beta\alpha}(\tilde{\boldsymbol{\Gamma}})\partial_\beta \Psi_0(\tilde{\boldsymbol{\Gamma}}) = \Psi_0(\tilde{\boldsymbol{\Gamma}}) \left(\mu_0 F_\alpha - \sum_\beta \partial_\beta \mathcal{D}_{\beta\alpha}(\tilde{\boldsymbol{\Gamma}}) \right). \tag{76}$$

We then multiply both sides by $\mathcal{D}^{-1}_{\alpha\gamma}$ and sum over the index α to get:

$$\partial_\gamma \log \Psi_0(\tilde{\boldsymbol{\Gamma}}) = \sum_\alpha \mathcal{D}^{-1}_{\alpha\gamma}(\tilde{\boldsymbol{\Gamma}}) \left[\mu_0 F_\alpha - \sum_\beta \partial_\beta \mathcal{D}_{\beta\alpha}(\tilde{\boldsymbol{\Gamma}}) \right] \equiv \beta F_\gamma^{\mathrm{eff}}(\tilde{\boldsymbol{\Gamma}}). \tag{77}$$

In order for DB to hold, the 'thermodynamic curvature' of $\boldsymbol{F}^{\mathrm{eff}}(\tilde{\boldsymbol{\Gamma}})$ must vanish, i.e.

$$\partial_\beta F_\gamma^{\mathrm{eff}} = \partial_\gamma F_\beta^{\mathrm{eff}} \quad \forall\ (\beta, \gamma). \tag{78}$$

We follow the approach of [91] and substitute equation (75) into equation (77). After some manipulations, we re-express the effective force as:

$$\beta F_\gamma^{\mathrm{eff}}(\tilde{\boldsymbol{\Gamma}}) = \partial_\gamma \left[-\frac{\mu_0}{D_a} U(\tilde{\boldsymbol{\Gamma}}) - \left(\frac{\mu_0 v_0 \tau}{2D_a}\right)^2 \sum_\alpha (F_\alpha(\tilde{\boldsymbol{\Gamma}}))^2 - \log\left(\det\, D_a^{-1}\mathcal{D}(\tilde{\boldsymbol{\Gamma}})\right) \right] \tag{79}$$

where $D_a \equiv \mu_0 k_{\mathrm{B}}T + v_0^2\tau/2$. It is now straightforward to verify equation (78). As a result, the system fulfills DB and therefore the FDT. In particular, in the presence of an

J. Stat. Mech. (2021) 043201

infinitesimal constant force h applied on the γ-degree of freedom the integrated response results:

$$\langle A\rangle_t - \langle A\rangle_0 = -\mu_0\beta h\int_0^t \mathrm{d}s\langle F_\gamma^{\mathrm{eff}}(0)A(s)\rangle_0 \tag{80}$$

which is the analogous of the equilibrium FDT, equation (31), with the effective force replacing the actual one [69]. Equation (78) in turn implies that the effective force derives from an effective (non-pairwise) potential, such that $\Psi_0 \propto \exp\left[-\mu_0\Phi_{\mathrm{eff}}^{\mathrm{Fox}}/D_a\right]$ with:

$$\Phi_{\mathrm{eff}}^{\mathrm{Fox}} = U + \frac{\mu_0 v_0^2\tau^2}{4D_a}\sum_\alpha(\partial_\alpha U)^2 + \frac{D_a}{\mu_0}\log\left(\det\, D_a^{-1}\mathcal{D}\right). \tag{81}$$

Note that in the limit of small thermal noise, i.e. $D_a \to v_0^2\tau/2$ the effective potential in equation (81) reduces to the one derived in [30] for the adiabatic unified colored noise approximation.

Before leaving this section, it is worth mentioning that a diagonal-Laplacian approximation for $\mathcal{D}(\tilde{\mathbf{\Gamma}})$ can be introduced [91, 94–96] by noticing that the effective diffusivity $\mathcal{D}(\tilde{\mathbf{\Gamma}})$ only enters in the expression for the effective force, equation (79), through its determinant. Within this approximation the effective force $\boldsymbol{F}_i^{\mathrm{eff}}$ on particle i reads:

$$\boldsymbol{F}_i^{\mathrm{eff}}(\tilde{\mathbf{\Gamma}}) = k_{\mathrm{B}}T(\mu_0\boldsymbol{F}_i - \partial_i\mathcal{D}_i)/\mathcal{D}_i \tag{82}$$

with

$$\mathcal{D}_i(\tilde{\mathbf{\Gamma}}) = D_a\left(\frac{1}{1-\tau\mu_0\partial_i\cdot\boldsymbol{F}_i}\right) \tag{83}$$

further simplifying the analysis of ABP within the Fox approximation. Despite fulfilling DB, this latter model has proved useful to capture some aspects of the out-of equilibrium behavior of ABP such as the emergence of attractive interactions out of purely steric ones [94] and the leading correction to the equilibrium Stokes–Einstein relation [69]. Those are reflected in the non-trivial shape of the effective potential in equation (81) which depends on the many-body diffusivity matrix, in contrast to the standard Boltzmann weight.

4.2. Active Ornstein–Uhlenbeck particles

4.2.1. The model. We consider in this section a similar model of self-propelled particles, now governed by the following set of two-dimensional overdamped Langevin equations

$$\dot{\boldsymbol{r}}_i(t) = \mu_0\boldsymbol{F}_i + \boldsymbol{v}_i \tag{84}$$

$$\dot{\boldsymbol{v}}_i(t) = -\frac{\boldsymbol{v}_i}{\tau} + \sqrt{\frac{2D_0}{\tau^2}}\boldsymbol{\eta}_i(t) \tag{85}$$

J. Stat. Mech. (2021) 043201

where $\boldsymbol{F}_i \equiv -\partial U/\partial \boldsymbol{r}_i$ is the conservative force acting on particle i (whose origin can be interactions with other particles or an external potential) and $\boldsymbol{v}_i$ is the fluctuating self-propulsion velocity which is described by an Ornstein–Uhlenbeck process with characteristic persistence time τ. Self-propulsion introduces persistence in the spatio-temporal dynamics of the active particles via the autocorrelation function of the self-propulsion velocity $\langle \boldsymbol{v}_i(t)\boldsymbol{v}_j(t')\rangle = D_0/\tau \exp(-|t-t'|/\tau)\delta_{ij}\mathbf{1}$ and reduces to passive (equilibrium) Brownian motion in the limit $\tau \to 0$, for which $\langle \boldsymbol{v}_i(t)\boldsymbol{v}_j(t')\rangle \to 2D_0\delta(t-t')\delta_{ij}\mathbf{1}$. Although a standard thermal noise could be added into the Langevin equation of AOUP, such contribution is assumed to be small with respect to the active noise $\boldsymbol{v}$ and might be considered redundant, as it is not needed to recover equilibrium. These AOUP can be thought of as an approximate treatment of the ABP dynamics. Indeed, the reduced ABP dynamics obtained from the integration of the angular variables equation (73) can be identified, in the absence of translational noise ($T = 0$), to the AOUP dynamics by setting $v_0^2/2$ (in ABP) to D_0/τ (in AOUP).

Although originally thought of as an overdamped process, equation (84) involves velocity variables $\boldsymbol{v}$ that can be considered as being odd under time-reversal [97]. Following this interpretation, one can introduce the velocities $\boldsymbol{p}_i \equiv \dot{\boldsymbol{r}}_i$ and rewrite equation (84) as an effective underdamped Langevin process [31]

$$\dot{\boldsymbol{r}}_i(t) = \boldsymbol{p}_i \tag{86}$$

$$\dot{\boldsymbol{p}}_i(t) = \mu_0 \left(\sum_j \boldsymbol{p}_j \cdot \partial_j\right) \boldsymbol{F}_i - \frac{\boldsymbol{p}_i}{\tau} + \mu_0 \frac{\boldsymbol{F}_i}{\tau} + \sqrt{\frac{2D_0}{\tau^2}}\boldsymbol{\eta}_i(t) \tag{87}$$

where $\partial_i \equiv \partial/\partial \boldsymbol{r}_i$. The corresponding generator reads

$$\Omega_0(\boldsymbol{\Gamma}) = \sum_i \left[-\boldsymbol{p}_i \cdot \partial_i - \partial_{\boldsymbol{p}_i} \cdot \left(\mu_0 \left(\sum_j \boldsymbol{p}_j \cdot \partial_j\right) \boldsymbol{F}_i - \frac{\boldsymbol{p}_i}{\tau} + \mu_0 \frac{\boldsymbol{F}_i}{\tau} - \frac{D_0}{\tau^2}\partial_{\boldsymbol{p}_i}\right)\right] \tag{88}$$

where $\partial_{\boldsymbol{p}_i} \equiv \partial/\partial \boldsymbol{p}_i$ and $\boldsymbol{\Gamma} = \{\boldsymbol{r}_i, \boldsymbol{p}_i\}$. We conclude the section with a remark. The thermodynamic consistency of the original AOUP dynamics equations (84) and (85) as well as the effective underdamped systems of equations (86) and (87) has been questioned [98] due to the absence of a thermal noise associated with a heat bath and a well-defined temperature. Nevertheless, as we will show below, the mapping to equations (86) and (87) is convenient as it allows for a perturbative approach to the stationary distribution $\Psi_0(\boldsymbol{\Gamma})$ in the limit of small relaxation time τ.

4.2.2. Effective equilibrium regime. For the AOUP model, the DB condition equation (23) reduces to:

$$\frac{1}{\tau^2}\partial_{\boldsymbol{p}_i} \log \Psi_0(\boldsymbol{\Gamma}) = \beta \left(\sum_j \boldsymbol{p}_j \cdot \partial_j\right) \boldsymbol{F}_i - \frac{\boldsymbol{p}_i}{D_0\tau}. \tag{89}$$

J. Stat. Mech. (2021) 043201

As before, since $\boldsymbol{F}_i$ is conservative, the formal solution of equation (89) can be expressed up to a function $\Lambda(\{\boldsymbol{r}_i\})$ that only depends on space variables such as

$$\Psi_0 = \exp\left[-\Lambda(\{\boldsymbol{r}_i\}) - \frac{\beta\tau^2}{2}\left(\sum_i \boldsymbol{p}_i\cdot\partial_i\right)^2 U - \sum_i \frac{\tau}{D_0}\frac{\boldsymbol{p}_i^2}{2}\right] \tag{90}$$

where $\beta \equiv \mu_0/D_0$. In order for DB to hold, the system must fulfill the following equation

$$\sum_i \left[\partial_i\Lambda + \frac{\beta\tau^2}{2}\left(\sum_j \boldsymbol{p}_j\cdot\partial_j\right)^2\partial_i U - \frac{\beta^2 D_0\tau}{2}\partial_i\sum_j|\partial_j U|^2 - \beta\partial_i U\right]\Psi_0 = 0. \tag{91}$$

This equation does not have a solution because of the second term comprising a p-dependence. Interestingly, for a potential with vanishing third derivatives the latter term vanishes and an exact 'equilibrium' solution exists [31, 33]:

$$\Psi_{\mathrm{eq}} = \mathcal{N}\exp\left[-\beta U - \frac{\tau}{2}\sum_i\left(\frac{\boldsymbol{p}_i^2}{D_0} + \beta^2 D_0|\partial_i U|^2\right) - \frac{\beta\tau^2}{2}\left(\sum_i \boldsymbol{p}_i\cdot\partial_i\right)^2 U\right]. \tag{92}$$

We wrote equilibrium in quotes because, contrarily to the standard Boltzmann measure, the probability of a given configuration is not solely given by $\mathrm{e}^{-\beta U}$, but by a more complicated function, also involving $\boldsymbol{p}$. This form is not *a priori* obvious from the mere inspection of the generator of the microscopic dynamics. (The same remark holds for ABP within the Fox approximation discussed earlier, as the effective potential equation (81) can hardly be guessed from the original Fokker–Planck equation.) However, in this equilibrium-like regime, AOUP fulfill DB, there are no irreversible fluxes, and the FDT holds. An equilibrium solution also exists in the case of non-interacting particles $U = 0$ for which AOUP are formally equivalent to an ideal gas of underdamped particles. In all other cases, for a generic U, the model breaks DB and therefore falls out-of-equilibrium.

4.2.3. Non-equilibrium regime: Chapman–Enskog expansion. An approximated stationary distribution Ψ_0 for AOUP, beyond its equilibrium-like regime, has recently been derived via the Chapman–Enskog expansion by Bonilla [33]. Our aim being to study the impact of activity on the response of an interacting system, we briefly present the Chapman–Enskog results as appeared in [33] and use them to establish extended FDR for AOUP.

The Chapman–Enskog expansion constitutes a standard perturbative approach to derive the Navier–Stokes equation from the Boltzmann equation [99–101]. It is based on the notions of local equilibrium and time scale separation. The latter is accounted for by the introduction of a small parameter $\epsilon = \ell/L$ defined as the ratio between a microscopic and a macroscopic characteristic length. In kinetic theory, ℓ is typically the *mean free path* between collisional events and L the size of the system. Likewise, we may associate to the AOUP two different scales: a microscopic one associated to the persistence time τ and diffusive length $\sqrt{D_0\tau}$ (characterizing the local persistence due to activity), and a mesoscopic one associated to the inter-particle interactions, with a 'slow'

J. Stat. Mech. (2021) 043201

characteristic time τ_0 and a large characteristic length L. We introduce the ratio parameter $\epsilon \equiv \sqrt{D_0\tau}/L \equiv \tau/\tau_0$ and rescale the AOUP equations of motion equations (86) and (87) according to $t \equiv t/\tau_0$, $\boldsymbol{r} \equiv \boldsymbol{r}/L$ and $\boldsymbol{F} \equiv \boldsymbol{F}L/\beta^{-1}$. The Fokker–Planck equation of AOUP can thus be written in the following non-dimensional form

$$\sum_i \partial_{\boldsymbol{p}_i} \cdot (\boldsymbol{p}_i + \partial_{\boldsymbol{p}_i})\,\Psi = \epsilon\partial_t\Psi + \epsilon\sum_i \left[\boldsymbol{p}_i \cdot \partial_i + \boldsymbol{F}_i \cdot \partial_{\boldsymbol{p}_i} + \epsilon\partial_{\boldsymbol{p}_i} \cdot \left(\sum_j \boldsymbol{p}_j \cdot \partial_j\right)\boldsymbol{F}_i\right]\Psi. \tag{93}$$

From here, the idea is to carry on a perturbative expansion in ϵ. For $\epsilon = 0$, a solution of equation (93) is:

$$\Psi^{(\epsilon=0)}(\boldsymbol{\Gamma}, t) = \frac{\mathrm{e}^{-\sum_i \boldsymbol{p}_i^2/2}}{(2\pi)^N} R(\boldsymbol{r}, t) \tag{94}$$

where $R(\boldsymbol{r}, t)$ is the normalized probability density of positions obtained by integrating over the velocities, such that $\int \Pi_i \mathrm{d}\boldsymbol{r}_i R(\boldsymbol{r}, t) = 1$. For a system with strong time-scale separation, $\epsilon \ll 1$, we assume that the functional dependence on $\{\boldsymbol{p}_i\}$ and R in equation (94) is preserved, and expand the probability distribution as a power series in ϵ:

$$\Psi(\boldsymbol{\Gamma}, t) = \frac{\mathrm{e}^{-\sum_i \boldsymbol{p}_i^2/2}}{(2\pi)^N} R(\boldsymbol{r}, t\,; \epsilon) + \sum_j \epsilon^j \phi^{(j)}(\boldsymbol{\Gamma}, R). \tag{95}$$

The crucial assumption of the Chapman–Enskog method is to still interpret R in equation (95) as the marginal probability distribution embedding the spatio-temporal dependence upon integration over the velocities. This corresponds to impose for $\phi^{(j)}$:

$$\int \Pi_i \mathrm{d}\boldsymbol{p}_i\; \phi^{(j)}(\boldsymbol{\Gamma}, R) = 0 \quad \forall\; j \tag{96}$$

and for R the continuity equation:

$$\frac{\partial R}{\partial t} = -\sum_{i,j} \epsilon^j \partial_i \cdot \int \Pi_i \mathrm{d}\boldsymbol{p}_i\; \boldsymbol{p}_i \phi^{(j)}(\boldsymbol{\Gamma}, R) \equiv \sum_j \epsilon^j \mathcal{F}^{(j)}, \tag{97}$$

where the first equality is obtained by integrating equation (93) over the velocities. The *ansatzs* equation (95) together with equation (97) are inserted into equation (93), resulting in a hierarchy of equations for the various terms in the expansion. We solve the set of equations up to $\sim o(\epsilon^3)$ and obtain the following (now made dimensional) probability distribution [31, 33]:

$$\Psi_0(\boldsymbol{\Gamma}) \simeq \mathcal{N}\,\exp\left[-\beta U - \sum_i \frac{\tau}{2}\left(\frac{\boldsymbol{p}_i^2}{D_0} + \beta^2 D_0 (\partial_i U)^2 - 3\beta D_0 \partial_i^2 U\right)\right.$$
$$\left. - \frac{\tau^2}{2}\left(\beta\left(\sum_j \boldsymbol{p}_j \cdot \partial_j\right)^2 U + \beta D_0 \sum_{i,j} (\boldsymbol{p}_j \cdot \partial_j)\partial_i^2 U\right) + \frac{\tau^3}{6}\beta\left(\sum_j \boldsymbol{p}_j \cdot \partial_j\right)^3 U\right]. \tag{98}$$

J. Stat. Mech. (2021) 043201

Some details on the derivation of equation (98) are given in the appendix for the one dimensional case. The generalization to higher dimensions is straightforward but lengthy (see also [33] for further details).

We are now in the position of computing the non-equilibrium response of AOUP up to second order in ϵ. First, let us decompose Ψ_0, according to equation (50), into its symmetric and antisymmetric parts:

$$\Phi_+ = \beta U + \sum_i \frac{\tau}{2}\left(\frac{\boldsymbol{p}_i^2}{D_0} + \beta^2 D_0(\partial_i U)^2 - 3\beta D_0 \partial_i^2 U\right) + \frac{\tau^2}{2}\beta\left(\sum_j \boldsymbol{p}_j\cdot\partial_j\right)^2 U \tag{99}$$

$$\Phi_- = \sum_{i,j}\frac{\tau^2}{2}\beta D_0(\boldsymbol{p}_j\cdot\partial_j)\partial_i^2 U - \frac{\tau^3}{6}\beta\left(\sum_i \boldsymbol{p}_i\cdot\partial_i\right)^3 U. \tag{100}$$

Note that, indeed Φ_+ does not contain powers of p larger than 2, as forbidden by the non-equilibrium constraint of equation (54). According to the latter condition, the signature of the departure from equilibrium is all embedded in Φ_- being different from zero. Particularly, the latter vanishes for a quadratic potential, as expected from the discussion in the previous section.

We now apply a constant force h along the x-axis on a tagged particle n. In this case, the extended FDR equations (62) and (63) reads

$$\langle A\rangle_t - \langle A\rangle_0 = \delta\mathcal{A}_{p_n^x}\left[-(D_{p_n^x})^{-1}\int_0^t \mathrm{d}s\langle A(s)\mathcal{A}^{\mathrm{irr}}_{p_n^x}(0)\rangle_0 + \int_0^t \mathrm{d}s\langle A(s)\frac{\partial\Phi_-}{\partial p_n^x}(0)\rangle_0\right] \tag{101}$$

where $\delta\mathcal{A}_{p_n^x} = \mu_0 h/\tau$ and $D_{p_n^x} = D_0/\tau^2$, leading to

$$\langle A\rangle_t - \langle A\rangle_0 = h\left[-\tau\beta\int_0^t \mathrm{d}s\langle A(s)\mathcal{A}^{\mathrm{irr}}_{p_n^x}(0)\rangle_0 + \frac{\mu_0}{\tau}\int_0^t \mathrm{d}s\langle A(s)\frac{\partial\Phi_-}{\partial p_n^x}(0)\rangle_0\right]. \tag{102}$$

The first term is directly determined upon making the identification $\mathcal{A}^{\mathrm{irr}}_{p_n^x} = \mu_0(\sum_j \boldsymbol{p}_j\cdot\partial_j)F_n^x - p_n^x/\tau$. The second integral requires the knowledge of the odd-symmetric part of Ψ_0, which, to third order in ϵ is given by equation (100). All in all, we derive the following FDR for AOUP

$$\begin{aligned}\langle A\rangle_t - \langle A\rangle_0 = \beta h\Bigg[&\int_0^t \mathrm{d}s\langle A(s)p_n^x(0)\rangle_0 - \tau\mu_0\int_0^t \mathrm{d}s\langle A(s)\left(\sum_j \boldsymbol{p}_j\cdot\partial_j\right)F_n^x(0)\rangle_0\\ &-\frac{1}{2}\mu_0\tau D_0\left(\int_0^t \mathrm{d}s\langle A(s)\frac{\partial}{\partial x_n}\left(\sum_j \partial_j\cdot\boldsymbol{F}_j\right)(0)\rangle_0\right.\\ &\left.-\frac{\tau}{D_0}\int_0^t \mathrm{d}s\langle A(s)\left(\sum_j \boldsymbol{p}_j\cdot\partial_j\right)^2 F_n^x(0)\rangle_0\right)\Bigg].\end{aligned} \tag{103}$$

By choosing $A\equiv\boldsymbol{p}_n$ we eventually obtain an extended Stokes–Einstein relation:

J. Stat. Mech. (2021) 043201

$$\mu = \beta \left[D - \tau\mu_0 \int_0^\infty \mathrm{d}s \langle p_n^x(s) \left(\sum_j \boldsymbol{p}_j \cdot \partial_j \right) F_n^x(0) \rangle_0 \right.$$
$$- \frac{1}{2}\mu_0 \tau D_0 \left(\int_0^\infty \mathrm{d}s \langle p_n^x(s) \frac{\partial}{\partial x_n} \left(\sum_j \partial_j \cdot \boldsymbol{F}_j \right) (0) \rangle_0 \right.$$
$$\left. \left. - \frac{\tau}{D_0} \int_0^\infty \mathrm{d}s \langle p_n^x(s) \left(\sum_j \boldsymbol{p}_j \cdot \partial_j \right)^2 F_n^x(0) \rangle_0 \right) \right] \tag{104}$$

where D is the many-body diffusivity $D = \int_0^\infty \mathrm{d}s \langle p_n^x(s) p_n^x(0) \rangle_0$. The expression above embeds the violations to the usual Stokes–Einstein relation due to the interplay between activity and inter-particle interactions, up to order $o(\tau^2)$. In the absence of interactions the Stokes–Einstein relation is restored. This is also true for the case of a harmonic potential $U(r) = k|\boldsymbol{r}|^2/2$, for which we are left with

$$\mu = \beta^{\mathrm{eff}} \int_0^t \mathrm{d}s \langle p_n^x(0) p_n^x(s) \rangle_0 \tag{105}$$

being $\beta^{\mathrm{eff}} = \beta(1 + \mu_0 \tau k)$ an effective temperature which depends on the stiffness of the external potential [28]. In contrast with ABP, the existence of a Stokes–Einstein relation for a harmonic potential in AOUP is due to the fact that the model fulfills DB.

5. Conclusions

We started the paper by recalling the pivotal role played by DB as the defining feature of equilibrium dynamics, how its breakdown out-of-equilibrium can be quantified by the presence of irreversible steady-state fluxes, and the symmetry properties of such fluxes under time-reversal. The main outcome of this first part is the derivation of a general expression for the time-reversal operator which incorporates the different parity of variables under time-reversal and allows us to express irreversible steady-fluxes in terms of the difference between the generator of the time-reversed dynamics and the original one.

A second achievement is obtained by making the connection between the breakdown of DB and the different contributions to the entropy production. In this context we derived a constraint of the irreversible steady-state fluxes which applies to any non-equilibrium system and provides non-trivial information for systems with dynamic variables which are odd under time-reversal. In particular, it constraints the functional dependence of the NESS distribution on its odd variables. We then concluded this part by making a connection with such constraints and the extended FDR, shedding light upon the nature of the different terms responsible for violations of the equilibrium FDT. In the second part of the paper we then apply these general results and formalism to ABP and AOUP. While ABP generically break DB, AOUP fulfill DB in the dilute limit, or in the case of a harmonic potential. The non-equilibrium nature of these two model

J. Stat. Mech. (2021) 043201

systems is therefore not equivalent. We then analyze their linear response in the presence of generic many-body interactions in an approximated fashion. In the case of ABP, we recall the Markov approximation method due to Fox and show that the effective dynamics resulting from it fulfills DB, and therefore also the standard FDT. For AOUP we exploit the Chapman–Enskog expansion performed in [33] which allows to derive an extended fluctuation–dissipation relation beyond its effective equilibrium regime. We discuss the violations of the Stokes–Einstein relation in these models of active particles and show the possibility of quantifying them in terms of effective temperatures.

Although some of the results presented here were known, as the existence of an effective equilibrium regime of AOUP and the NESS solution obtained from the Chapman–Enskog expansion, the discussion about the violations of DB and the FDT were scattered and scarce. The extended FDR, as well as the connection with the parity of the NESS distribution and the constraint on the steady-state velocity we derived, enrich previous discussions on non-equilibrium response, clarify the non-equilibrium nature of active model systems and provide a set of analytic results that should be of interest to study non-equilibrium systems in general, well beyond the context of active systems.

Acknowledgments

We warmly thank Jorge Kurchan, Matteo Polettini and Patrick Pietzonka for useful discussions and suggestions. S D C acknowledges funding from the European Union's Horizon 2020 Framework Programme/European Training Programme 674979 NanoTRANS. D L acknowledges MCIU/AEI/FEDER for financial support under Grant agreement RTI2018-099032-J-I00. I P acknowledges support from Ministerio de Ciencia, Innovación y Universidades (Grant No. PGC2018-098373-B-100 AEI/FEDER-EU) and from Generalitat de Catalunya under project 2017SGR-884, and Swiss National Science Foundation Project No. 200021-175719.

Appendix A. Chapman–Enskog method for one dimensional AOUP

Here we report some details of the Chapman–Enskog procedure for a one dimensional system of AOUP. The starting point is the evolution equation:

$$\partial_p \left(p + \partial_p\right) \Psi = \epsilon \partial_t \Psi + \epsilon [p \partial_x \Psi + F \partial_p \Psi + \epsilon F' \partial_p (p \partial_x \Psi)], \tag{106}$$

together with the ansatz for Ψ:

$$\Psi(x, p, t) = \frac{\mathrm{e}^{-p^2/2}}{(2\pi)} R(x, t\,; \epsilon) + \sum_j \epsilon^j \phi^{(j)}(x, p, R). \tag{107}$$

By integrating equation (106) over p we obtain a conservation equation for R:

$$\partial_t R = -\partial_x \sum_j J^{(j)} = -\sum_j \epsilon^j \partial_x \int \mathrm{d}p\, p \phi^{(j)} \equiv \sum_j \epsilon^j \mathcal{F}^{(j)} \tag{108}$$

J. Stat. Mech. (2021) 043201

where $J = \sum_j J^{(j)}$ is the probability flux.

We now substitute equations (108) and (107) into equation (106) and we obtain to first order in ϵ an equation for $\phi^{(1)}$:

$$\mathcal{O}(\epsilon) : \mathcal{L}[\phi^{(1)}] \equiv \partial_p \left(p\phi^{(1)} + \partial_p \phi^{(1)} \right) = p \frac{\mathrm{e}^{-p^2/2}}{\sqrt{2\pi}} (\partial_x R + U' R). \tag{109}$$

In equation (109) $\phi^{(1)}$ has a Gaussian dependence in momenta and we find:

$$\phi^{(1)} = -p \frac{\mathrm{e}^{-p^2/2}}{\sqrt{2\pi}} \mathcal{D}R \tag{110}$$

with $\mathcal{D}R \equiv U'R + \partial_x R$. Note that $\mathcal{L}[\mathrm{e}^{-p^2/2}] = 0$ such that the functions $\phi^{(j)}$ are defined up to a constant in p which is uniquely determined by conditions equation (96) in the main text. The second order in ϵ reads:

$$\mathcal{O}(\epsilon^2) : \mathcal{L}[\phi^{(2)}] = \frac{\mathrm{e}^{-p^2/2}}{\sqrt{2\pi}} [\mathcal{F}^{(1)}(R) + (p^2 - 1) U'' R] + p \partial_x \phi^{(1)} - U' \partial_p \phi^{(1)}. \tag{111}$$

From equations (110) and (111) it follows that $\phi^{(2)}$ is even in p, such that $\mathcal{F}^{(2)} = \int \mathrm{d}p\, p \phi^{(2)} = 0$. Using

$$\mathcal{L} \left[p^2 \frac{\mathrm{e}^{-p^2/2}}{\sqrt{2\pi}} \right] = -2(p^2 - 1) \frac{\mathrm{e}^{-p^2/2}}{\sqrt{2\pi}} \tag{112}$$

we find:

$$\phi^{(2)} = \alpha(x)(p^2 - 1) \frac{\mathrm{e}^{-p^2/2}}{\sqrt{2\pi}} \tag{113}$$

with $\alpha = \frac{1}{2}(\partial_x \mathcal{D}R + U' \mathcal{D}R - U'' R)$.

To third order in ϵ the evolution equation gives:

$$\mathcal{O}(\epsilon^3) : \mathcal{L}[\phi^{(3)}] = \frac{\delta \phi^{(1)}}{\delta R} \mathcal{F}^{(1)}(R) + p \partial_x \phi^{(2)} - U' \partial_p \phi^{(2)} - U'' \partial_p (p \phi^{(1)}) \tag{114}$$

which, using (110) and (113), together with the identity:

$$\mathcal{L} \left[p^3 \frac{\mathrm{e}^{-p^2/2}}{\sqrt{2\pi}} \right] = \frac{\mathrm{e}^{-p^2/2}}{\sqrt{2\pi}} 3p(2 - p^2) \tag{115}$$

gives:

$$\phi^{(3)} = p \frac{\mathrm{e}^{-p^2/2}}{\sqrt{2\pi}} (\beta(x) p^2 + \gamma(x)) \tag{116}$$

with $\beta = -\frac{1}{3}(\mathcal{D}\alpha - U'' \mathcal{D}R)$ and $\gamma = \mathcal{D}\partial_x \mathcal{D}R - \partial_x \alpha + U' \alpha$. We truncate the hierarchy of equations to $\sim o(\epsilon^3)$ and, going back to the conservative equation (108), we find an expression for the first three components of the flux as a function of R:

$$J^{(1)} = -\mathcal{D}R \tag{117}$$

J. Stat. Mech. (2021) 043201

$$J^{(2)} = 0 \tag{118}$$

$$J^{(3)} = \partial_x \left(U'' R \right). \tag{119}$$

We have therefore a closed approximated equation for the evolution of R:

$$\frac{\partial R}{\partial t} = -\epsilon \partial_x [-\partial_x R - U' R + \epsilon^2 \partial_x (U'' R)] + o(\epsilon^3). \tag{120}$$

Equation (120) is readily solved by looking at the zero-flux steady-state with solution:

$$R(x\,;\epsilon) = \exp\left[-U - \epsilon^2 \left(\frac{1}{2} U'^2 - U'' \right) \right] + o(\epsilon^3). \tag{121}$$

In turn equation (121) is used to find the explicit expressions of $\phi^{1,2,3}$ which are substituted in equation (95) to get:

$$\Psi_0(x, p\,;\epsilon) = \Psi_0^{(\epsilon=0)} \left[1 + \frac{\epsilon^2}{2}(1 - p^2) U'' + \epsilon^3 \left(\frac{p^3}{6} U''' - \frac{p}{2} U''' \right) \right] + o(\epsilon^3). \tag{122}$$

Finally, we use equation (121) into equation (122) and interpret the term in square brackets as the linearization of an exponential $\exp(x) \sim 1 + x$ so that:

$$\Psi_0(x, p\,;\epsilon) \sim \exp\left[-\frac{p^2}{2} - U - \frac{\epsilon^2}{2} \left(p^2 U'' + U'^2 - 3U'' \right) + \frac{\epsilon^3}{2} \left(\frac{p^3}{3} U''' - p U''' \right) \right] + o(\epsilon^3). \tag{123}$$

References

[1] Kubo R, Toda M and Hashitsume N 1991 *Statistical Physics II* (*Nonequilibrium Statistical Mechanics*) vol 31 (Berlin: Springer)
[2] Brown W 1993 *Dynamic Light Scattering* (*The Method and Some Applications*) vol 313 (Oxford: Clarendon)
[3] Cipelletti L and Ramos L 2005 Slow dynamics in glassy soft matter *J. Phys.: Condens. Matter* **17** R253
[4] Kubo R 1957 Statistical-mechanical theory of irreversible processes. I. General theory and simple applications to magnetic and conduction problems *J. Phys. Soc. Japan* **12** 570–86
[5] Kubo R 1966 The fluctuation–dissipation theorem *Rep. Prog. Phys.* **29** 255
[6] Hänggi P and Thomas H 1982 Stochastic processes: time evolution, symmetries and linear response *Phys. Rep.* **88** 207–319
[7] Risken H 1996 *The Fokker-Planck Equation* (Berlin: Springer)
[8] Graham R and Haken H 1971 Generalized thermodynamic potential for Markoff systems in detailed balance and far from thermal equilibrium *Z. Phys.* **243** 289–302
[9] Agarwal G S 1972 Fluctuation–dissipation theorems for systems in non-thermal equilibrium and applications *Z. Phys.* **252** 25–38
[10] Haken H 1969 Exact stationary solution of a Fokker–Planck equation for multimode laser action including phase locking *Z. Phys.* **219** 246–68
[11] Harada T and Sasa S-i 2005 Equality connecting energy dissipation with a violation of the fluctuation–response relation *Phys. Rev. Lett.* **95** 130602
[12] Speck T and Seifert U 2006 Restoring a fluctuation–dissipation theorem in a nonequilibrium steady state *Europhys. Lett.* **74** 391
[13] Prost J, Joanny J-F and Parrondo J M R 2009 Generalized fluctuation–dissipation theorem for steady-state systems *Phys. Rev. Lett.* **103** 090601

J. Stat. Mech. (2021) 043201

[14] Baiesi M, Maes C and Wynants B 2009 Fluctuations and response of nonequilibrium states *Phys. Rev. Lett.* **103** 010602
[15] Chetrite R and Gawędzki K 2009 Eulerian and Lagrangian pictures of non-equilibrium diffusions *J. Stat. Phys.* **137** 890
[16] Seifert U and Speck T 2010 Fluctuation–dissipation theorem in nonequilibrium steady states *Europhys. Lett.* **89** 10007
[17] Kurchan J 1998 Fluctuation theorem for stochastic dynamics *J. Phys. A: Math. Gen.* **31** 3719–29
[18] Marconi U M B, Puglisi A, Rondoni L and Angelo V 2008 Fluctuation–dissipation: response theory in statistical physics *Phys. Rep.* **461** 111–95
[19] Baiesi M and Maes C 2013 An update on the nonequilibrium linear response *New J. Phys.* **15** 013004
[20] Sarracino A and Vulpiani A 2019 On the fluctuation–dissipation relation in non-equilibrium and non-Hamiltonian systems *Chaos* **29** 083132
[21] Maes C 2020 Response theory: a trajectory-based approach *Front. Phys.* **8** 229
[22] Cates M E 2012 Diffusive transport without detailed balance in motile bacteria: does microbiology need statistical physics? *Rep. Prog. Phys.* **75** 042601
[23] Bechinger C, Di Leonardo R, Löwen H, Reichhardt C, Volpe G and Volpe G 2016 Active particles in complex and crowded environments *Rev. Mod. Phys.* **88** 045006
[24] Romanczuk P, Bär M, Ebeling W, Lindner B and Schimansky-Geier L 2012 Active Brownian particles *Eur. Phys. J. Spec. Top.* **202** 1–162
[25] Fily Y and Marchetti M C 2012 Athermal phase separation of self-propelled particles with no alignment *Phys. Rev. Lett.* **108** 235702
[26] Redner G S, Hagan M F and Baskaran A 2013 Structure and dynamics of a phase-separating active colloidal fluid *Phys. Rev. Lett.* **110** 055701
[27] Digregorio P, Levis D, Suma A, Cugliandolo L F, Gonnella G and Pagonabarraga I 2018 Full phase diagram of active Brownian disks: from melting to motility-induced phase separation *Phys. Rev. Lett.* **121** 098003
[28] Szamel G 2014 Self-propelled particle in an external potential: existence of an effective temperature *Phys. Rev.* E **90** 012111
[29] Maggi C, Paoluzzi M, Pellicciotta N, Lepore A, Angelani L and Di Leonardo R 2014 Generalized energy equipartition in harmonic oscillators driven by active baths *Phys. Rev. Lett.* **113** 238303
[30] Marconi U M B and Maggi C 2015 Towards a statistical mechanical theory of active fluids *Soft Matter* **11** 8768–81
[31] Fodor É, Nardini C, Cates M E, Tailleur J, Visco P and van Wijland F 2016 How far from equilibrium is active matter? *Phys. Rev. Lett.* **117** 038103
[32] Dabelow L, Bo S and Eichhorn R 2019 Irreversibility in active matter systems: fluctuation theorem and mutual information *Phys. Rev.* X **9** 021009
[33] Bonilla L L 2019 Active Ornstein–Uhlenbeck particles *Phys. Rev.* E **100** 022601
[34] Takatori S C, Yan W and Brady J F 2014 Swim pressure: stress generation in active matter *Phys. Rev. Lett.* **113** 028103
[35] Mallory S A, šarić A, Valeriani C and Cacciuto A 2014 Anomalous thermomechanical properties of a self-propelled colloidal fluid *Phys. Rev.* E **89** 052303
[36] Ginot F, Theurkauff I, Levis D, Ybert C, Bocquet L, Berthier L and Cottin-Bizonne C 2015 Nonequilibrium equation of state in suspensions of active colloids *Phys. Rev.* X **5** 011004
[37] Solon A P, Fily Y, Baskaran A, Cates M E, Kafri Y, Kardar M and Tailleur J 2015 Pressure is not a state function for generic active fluids *Nat. Phys.* **11** 673–8
[38] Winkler R G, Wysocki A and Gompper G 2015 Virial pressure in systems of spherical active Brownian particles *Soft Matter* **11** 6680–91
[39] Wittkowski R, Tiribocchi A, Stenhammar J, Allen R J, Marenduzzo D and Cates M E 2014 Scalar φ 4 field theory for active-particle phase separation *Nat. Commun.* **5** 1–9
[40] Stenhammar J, Marenduzzo D, Allen R J and Cates M E 2014 Phase behaviour of active Brownian particles: the role of dimensionality *Soft Matter* **10** 1489–99
[41] Levis D, Codina J and Pagonabarraga I 2017 Active Brownian equation of state: metastability and phase coexistence *Soft Matter* **13** 8113–9
[42] Paliwal S, Rodenburg J, van Roij R and Dijkstra M 2018 Chemical potential in active systems: predicting phase equilibrium from bulk equations of state? *New J. Phys.* **20** 015003
[43] Steffenoni S, Falasco G and Kroy K 2017 Microscopic derivation of the hydrodynamics of active-Brownian-particle suspensions *Phys. Rev.* E **95** 052142

J. Stat. Mech. (2021) 043201

[44] Solon A P, Stenhammar J, Cates M E, Kafri Y and Tailleur J 2018 Generalized thermodynamics of motility-induced phase separation: phase equilibria, Laplace pressure, and change of ensembles *New J. Phys.* **20** 075001
[45] Hermann S, de las Heras D and Schmidt M 2019 Non-negative interfacial tension in phase-separated active Brownian particles *Phys. Rev. Lett.* **123** 268002
[46] Di Leonardo R *et al* 2010 Bacterial ratchet motors *Proc. Natl Acad. Sci.* **107** 9541–5
[47] Sokolov A, Apodaca M M, Grzybowski B A and Aranson I S 2010 Swimming bacteria power microscopic gears *Proc. Natl Acad. Sci.* **107** 969–74
[48] López H M, Gachelin J, Douarche C, Auradou H and Clément E 2015 Turning bacteria suspensions into superfluids *Phys. Rev. Lett.* **115** 028301
[49] Rafaï S, Jibuti L and Peyla P 2010 Effective viscosity of microswimmer suspensions *Phys. Rev. Lett.* **104** 098102
[50] Paxton W F, Sundararajan S, Mallouk T E and Sen A 2006 Chemical locomotion *Angew. Chem., Int. Ed.* **45** 5420–9
[51] Palacci J, Cottin-Bizonne C, Ybert C and Bocquet L 2010 Sedimentation and effective temperature of active colloidal suspensions *Phys. Rev. Lett.* **105** 088304
[52] Palacci J, Sacanna S, Steinberg A P, Pine D J and Chaikin P M 2013 Living crystals of light-activated colloidal surfers *Science* **339** 936–40
[53] Theurkauff I, Cottin-Bizonne C, Palacci J, Ybert C and Bocquet L 2012 Dynamic clustering in active colloidal suspensions with chemical signaling *Phys. Rev. Lett.* **108** 268303
[54] Sharma A and Brader J M 2016 Communication: Green–Kubo approach to the average swim speed in active Brownian systems *J. Chem. Phys.* **145** 161101
[55] Liluashvili A, Ónody J and Voigtmann T 2017 Mode-coupling theory for active Brownian particles *Phys. Rev.* E **96** 062608
[56] Caprini L, Marconi U M B and Vulpiani A 2018 Linear response and correlation of a self-propelled particle in the presence of external fields *J. Stat. Mech.* 033203
[57] Asheichyk K, Solon A P, Rohwer C M and Krüger M 2019 Response of active Brownian particles to shear flow *J. Chem. Phys.* **150** 144111
[58] Warren P B and Allen R J 2012 Malliavin weight sampling for computing sensitivity coefficients in Brownian dynamics simulations *Phys. Rev. Lett.* **109** 250601
[59] Szamel G 2017 Evaluating linear response in active systems with no perturbing field *Europhys. Lett.* **117** 50010
[60] Cugliandolo L F 2011 The effective temperature *J. Phys. A: Math. Theor.* **44** 483001
[61] Loi D, Mossa S and Cugliandolo L F 2008 Effective temperature of active matter *Phys. Rev.* E **77** 051111
[62] Levis D and Berthier L 2015 From single-particle to collective effective temperatures in an active fluid of self-propelled particles *Europhys. Lett.* **111** 60006
[63] Ben-Isaac E, Park YongK, Popescu G, Brown F L H, Gov N S and Shokef Y 2011 Effective temperature of red-blood-cell membrane fluctuations *Phys. Rev. Lett.* **106** 238103
[64] Cugliandolo L F, Gonnella G and Petrelli I 2019 Effective temperature in active Brownian particles *Fluctuation Noise Lett.* **18** 1940008
[65] Eldeen S, Muoio R, Blaisdell-Pijuan P, La N, Gomez M, Vidal A and Ahmed W 2020 Quantifying the non-equilibrium activity of an active colloid *Soft Matter* **16** 7202–9
[66] Mandal S, Liebchen B and Löwen H 2019 Motility-induced temperature difference in coexisting phases *Phys. Rev. Lett.* **123** 228001
[67] Flenner E and Szamel G 2020 Active matter: quantifying the departure from equilibrium (arXiv:2004.11925)
[68] Petrelli I, Cugliandolo L F, Gonnella G and Suma A 2020 Effective temperatures in inhomogeneous passive and active bidimensional Brownian particle systems *Phys. Rev.* E **102** 012609
[69] Dal Cengio S, Levis D and Pagonabarraga I 2019 Linear response theory and Green–Kubo relations for active matter *Phys. Rev. Lett.* **123** 238003
[70] Van Kampen N G 1989 Langevin-like equation with colored noise *J. Stat. Phys.* **54** 1289–308
[71] Qian H 1998 Vector field formalism and analysis for a class of thermal ratchets *Phys. Rev. Lett.* **81** 3063–6
[72] Qian H, Qian M and Tang X 2002 Thermodynamics of the general diffusion process: time-reversibility and entropy production *J. Stat. Phys.* **107** 1129–41
[73] Wang J, Xu L and Wang E 2008 Potential landscape and flux framework of nonequilibrium networks: robustness, dissipation, and coherence of biochemical oscillations *Proc. Natl Acad. Sci.* **105** 12271–6
[74] Evans D J and Morriss G 2008 *Statistical Mechanics of Nonequilibrium Liquids* 2nd edn (Cambridge: Cambridge University Press)
[75] Gardiner C 2009 *Stochastic Methods* (Berlin: Springer)

J. Stat. Mech. (2021) 043201

[76] Tailleur J, Tănase-Nicola S and Kurchan J 2006 Kramers equation and supersymmetry *J. Stat. Phys.* **122** 557–95
[77] Kurchan J 2009 Six out of equilibrium lectures (arXiv:0901.1271)
[78] Polettini M 2013 Of dice and men. Subjective priors, gauge invariance, and nonequilibrium thermodynamics (arXiv:1307.2057)
[79] Falcioni M, Isola S and Vulpiani A 1990 Correlation functions and relaxation properties in chaotic dynamics and statistical mechanics *Phys. Lett.* A **144** 341–6
[80] Cugliandolo L F, Kurchan J and Parisi G 1994 Off equilibrium dynamics and aging in unfrustrated systems *J. Phys.* I **4** 1641–56
[81] Ge H 2014 Time reversibility and nonequilibrium thermodynamics of second-order stochastic processes *Phys. Rev.* E **89** 022127
[82] Baiesi M, Boksenbojm E, Maes C and Wynants B 2010 Nonequilibrium linear response for Markov dynamics, II: inertial dynamics *J. Stat. Phys.* **139** 492–505
[83] Freitas N, Delvenne J-C and Esposito M 2019 Stochastic and quantum thermodynamics of driven RLC networks (arXiv:1906.11233)
[84] Spinney R E and Ford I J 2012 Entropy production in full phase space for continuous stochastic dynamics *Phys. Rev.* E **85** 051113
[85] Qian H 2001 Mesoscopic nonequilibrium thermodynamics of single macromolecules and dynamic entropy-energy compensation *Phys. Rev.* E **65** 016102
[86] Seifert U 2005 Entropy production along a stochastic trajectory and an integral fluctuation theorem *Phys. Rev. Lett.* **95** 040602
[87] Feng H and Wang J 2011 Potential and flux decomposition for dynamical systems and non-equilibrium thermodynamics: curvature, gauge field, and generalized fluctuation–dissipation theorem *J. Chem. Phys.* **135** 234511
[88] Solon A P, Cates M E and Tailleur J 2015 Active Brownian particles and run-and-tumble particles: a comparative study *Eur. Phys. J. Spec. Top.* **224** 1231–62
[89] Peter H and Jung P 1994 Colored noise in dynamical systems *Adv. Chem. Phys.* **89** 239–326
[90] Maggi C, Marconi U M B, Gnan N and Di Leonardo R 2015 Multidimensional stationary probability distribution for interacting active particles *Sci. Rep.* **5** 10742
[91] Wittmann R, Maggi C, Sharma A, Scacchi A, Brader J M and Marconi U M B 2017 Effective equilibrium states in the colored-noise model for active matter I. Pairwise forces in the fox and unified colored noise approximations *J. Stat. Mech.* 113207
[92] Fox R F 1986 Uniform convergence to an effective Fokker–Planck equation for weakly colored noise *Phys. Rev.* A **34** 4525
[93] Fox R F 1986 Functional-calculus approach to stochastic differential equations *Phys. Rev.* A **33** 467
[94] Farage T F, Krinninger P and Brader J M 2015 Effective interactions in active Brownian suspensions *Phys. Rev.* E **91** 042310
[95] Wittmann R and Brader J M 2016 Active Brownian particles at interfaces: an effective equilibrium approach *Europhys. Lett.* **114** 68004
[96] Rein M and Speck T 2016 Applicability of effective pair potentials for active Brownian particles *Eur. Phys. J.* E **39** 84
[97] Dadhichi L P, Maitra A and Ramaswamy S 2018 Origins and diagnostics of the nonequilibrium character of active systems *J. Stat. Mech.* 123201
[98] Seifert U 2019 From stochastic thermodynamics to thermodynamic inference *Annu. Rev. Condens. Matter Phys.* **10** 171–92
[99] Chapman S, Cowling T G, Burnett D and Cercignani C 1990 *The Mathematical Theory of Non-uniform Gases* (*An Account of the Kinetic Theory of Viscosity, Thermal Conduction and Diffusion in Gases*) Cambridge Mathematical Library (Cambridge: Cambridge University Press)
[100] Grad H 1958 *Principles of the Kinetic Theory of Gases* (Berlin: Springer) pp 205–94
[101] Bonilla L L and Soler J S 2001 High-field limit of the Vlasov–Poisson–Fokker–Planck system: a comparison of different perturbation methods *Math. Models Methods Appl. Sci.* **11** 1457–68

J. Stat. Mech. (2021) 043201