

**Additivity, density fluctuations, and nonequilibrium thermodynamics for active Brownian particles**Subhadip Chakraborti,<sup>\*</sup> Shradha Mishra,<sup>†</sup> and Punyabrata Pradhan<sup>‡</sup>*Department of Theoretical Sciences, S. N. Bose National Centre for Basic Sciences, Block JD, Sector III, Salt Lake, Kolkata 700106, India*

(Received 14 August 2015; revised manuscript received 17 March 2016; published 12 May 2016)

Using an additivity property, we study particle-number fluctuations in a system of interacting self-propelled particles, called active Brownian particles (ABPs), which consists of repulsive disks with random self-propulsion velocities. From a fluctuation-response relation, a direct consequence of additivity, we formulate a thermodynamic theory which captures the previously observed features of nonequilibrium phase transition in the ABPs from a homogeneous fluid phase to an inhomogeneous phase of coexisting gas and liquid. We substantiate the predictions of additivity by analytically calculating the subsystem particle-number distributions in the homogeneous fluid phase away from criticality where analytically obtained distributions are compatible with simulations in the ABPs.

DOI: [10.1103/PhysRevE.93.052606](https://doi.org/10.1103/PhysRevE.93.052606)**I. INTRODUCTION**

The problem of characterization of driven many-particle systems, having a steady state, has drawn much attention in the past decades [1,2]. The problem is, however, quite hard to tackle, mainly due to that the steady-state probability weights of microscopic configurations in a driven system, unlike in equilibrium, are not described by the Boltzmann distribution. Indeed, in most cases, these weights are not known. A simple characterization of the steady-state systems in general would be certainly desirable, and, to this end, various attempts have been made [3,4]. Recently a formulation based on equilibriumlike additivity property provides a framework [5–7], which helps one to describe a broad class of nonequilibrium steady states through fluctuations of a conserved quantity, e.g., mass or particle-number [8–10]. Here we address the question whether an additivity property can be used to obtain large deviation probability for the particle-number or the density fluctuations, the central object in a statistical mechanics theory, in systems of self-propelled particles.

Self-propelled particles (SPPs), also called active matters, are prevalent in nature: in living systems, e.g., bacterial colonies [11], fish schools [12], flocks of birds [13], and insect swarms [14], as well as in nonliving systems, e.g., photoactivated or chemically powered colloids, thermophoretic Janus particles [15], etc. They have been realized in experiments [16] and intensively studied through simulations and theories [17–25]; for reviews, see Refs. [26,27]. These particles propel themselves by converting chemical energy to mechanical one, which is continually dissipated to the medium. The steady flow of energy keeps the system out of equilibrium, and a novel nonequilibrium steady state emerges. Such a nonequilibrium steady state manifests itself by exhibiting rich collective phenomena, e.g., self-assemblies and pattern formations, otherwise impossible in equilibrium.

Recently, we have seen a surge of activities, in search of a suitable statistical mechanics framework which could describe macroscopic properties of the self-propelled particles

in terms of an intensive thermodynamic variable, such as a chemical potential [27–35], pressure [36–41], or an effective temperature [42,43]. However, a complete framework still remains elusive. We propose here a general thermodynamic principle, called additivity, which could enable us to unify fascinatingly broad-ranging phenomena in the systems of self-propelled particles under a unique nonequilibrium thermodynamic theory, directly connecting microscopic fluctuations to the macroscopic properties in the system.

In this paper, using an additivity property, we formulate a thermodynamic theory for a particular class of self-propelled particles, called active Brownian particles (ABPs), consisting of repulsive disks in two dimensions and which have random self-propulsion velocities. We demonstrate, in the regime of homogeneous phase, that additivity property leads to subsystem particle-number distribution  $P_{\mathcal{V}}(\mathcal{N})$ , the probability that a subsystem of volume  $\mathcal{V}$  has  $\mathcal{N}$  number of particles. The logarithm of the probability  $P_{\mathcal{V}}(\mathcal{N})$ , or the large deviation function (LDF), analogous to equilibrium free energy, governs the density fluctuations and thus can immediately connect to the standard statistical mechanics framework. We validate the predictions of the theory regarding density fluctuations by analytically calculating the subsystem particle-number distribution  $P_{\mathcal{V}}(\mathcal{N})$  in the homogeneous fluid phase in the active Brownian particles and comparing them with simulations.

The crucial ingredient of this theory is a nonequilibrium fluctuation-response relation (FR) between compressibility and number-fluctuation or variance  $\sigma_{\mathcal{V}}^2 = \langle \mathcal{N}^2 \rangle - \langle \mathcal{N} \rangle^2$  [see Eq. (5)], which is a direct consequence of additivity. Provided the functional dependence of the variance  $\sigma_{\mathcal{V}}^2(\rho)$  on the particle-number density  $\rho$ , we provide a prescription of how, using additivity, one can calculate the distribution function  $P_{\mathcal{V}}(\mathcal{N})$ .

To illustrate the formalism, we first calculate, within a linearized fluctuating hydrodynamics of the ABPs, the variance  $\sigma_{\mathcal{V}}^2(\rho)$  of particle number in a subsystem of volume  $\mathcal{V}$  as a function of density  $\rho$ . Then we use the standard large deviation methods to obtain the large deviation function, or a nonequilibrium free energy density function  $f(\rho, \text{Pe})$ , and a chemical potential  $\mu(\rho, \text{Pe})$ , as a function of number density  $\rho$  and activity parameter Peclet number  $\text{Pe}$ . Determination of chemical potential leads to a nonequilibrium equation of state, akin to the equilibrium Van der Waals one. Beyond

<sup>\*</sup>subhadip.chakraborti@bose.res.in<sup>†</sup>smishra.phy@itbhu.ac.in<sup>‡</sup>punyabrata.pradhan@bose.res.in

a critical activity, compressibility  $d\rho/d\mu$  becomes negative in a particular density interval, leading to nonmonotonic  $\mu$  as a function of  $\rho$  and hence phase coexistence. In special limits, our theory captures various previous results, e.g., those based on the concept of motility induced phase separation (MIPS) [29–31], indicating the formulation here is indeed consistent with the past studies. Moreover, our analysis suggests that, on a mean-field level, a broad class of self-propelled particles belong to Ising universality.

The organization of the paper is as follows. In Sec. II we discuss additivity and show how subsystem particle-number distribution can be calculated solely from the knowledge of variance of subsystem particle number as a function of number density. In Sec. III we define the model of active Brownian particles and discuss the corresponding fluctuating hydrodynamics. In Sec. IV we calculate, within linearized hydrodynamics, variance of subsystem particle number as a function of density (Sec. IV A) and then characterize noise strengths in the hydrodynamic equations (Sec. IV B). In Sec. V using the functional dependence of the variance on density together with additivity, we formulate a thermodynamic theory of active Brownian particles and substantiate the theory by explicitly calculating subsystem particle-number distributions in homogeneous fluid phase of active Brownian particles. Finally, we summarize in Sec. VI.

## II. ADDITIVITY AND SUBSYSTEM PARTICLE-NUMBER DISTRIBUTION

In this section, we discuss an additivity property which systems having a *finite* correlation length are expected to possess, *irrespective of* whether the systems are *in* or *out* of equilibrium [5–7]. Recently additivity has been used in nonequilibrium mass-transport processes for calculating mass distributions and characterizing macroscopic properties in terms of equilibriumlike thermodynamic potentials [8,9]. Below we discuss how additivity can be used to calculate subsystem particle-number distribution.

First, let us discuss what additivity means in the context of particle-number or density fluctuations in a system. Let us consider  $N$  interacting particles in volume  $V$  where the total number of particles  $N$  is conserved. We divide the system in  $\nu = V/\mathcal{V}$  number of identical subsystems, each having volume  $\mathcal{V}$ , and ask what could be the form of the joint probability distributions for the subsystem particle-numbers  $\{\mathcal{N}_i\} \equiv \{\mathcal{N}_1, \mathcal{N}_2, \dots, \mathcal{N}_\nu\}$ . Provided that the subsystem size is much larger than spatial correlation length  $\xi$ ,  $\nu^{1/d} \gg \xi$  in  $d$  dimensions, additivity implies that the subsystems are statistically almost independent, and therefore, to a very good approximation, the joint subsystem particle-number distribution can be written in a product form [5–7],

$$\mathcal{P}[\{\mathcal{N}_i\}] \simeq \frac{\prod_{k=1}^{\nu} W_{\mathcal{V}}(\mathcal{N}_k)}{Z(N, V)} \delta\left(\sum_k \mathcal{N}_k - N\right), \quad (1)$$

in the thermodynamic limit of  $N, V \rightarrow \infty$  with density  $\rho = N/V$  fixed. In Eq. (1)  $W_{\mathcal{V}}(\mathcal{N}_k)$  is an unknown weight factor, which depends on the subsystem particle number and will be determined later. The normalization constant, or the partition

sum,  $Z(N, V)$  in Eq. (1) can be written as

$$Z(N, V) = \sum_{\{\mathcal{N}_i\}} \left[ \prod_k W_{\mathcal{V}}(\mathcal{N}_k) \right] \delta\left(\sum_k \mathcal{N}_k - N\right). \quad (2)$$

In other words, the property that the joint subsystem particle-number distribution  $\mathcal{P}[\{\mathcal{N}_i\}]$  for a system can be approximately written as a product (i.e., subsystems are statistically almost independent) of individual subsystem weight factors  $W_{\mathcal{V}}(\mathcal{N}_k)$  is called additivity.

In equilibrium, there is a well-defined thermodynamic prescription, which helps us to calculate the weight factor  $W_{\mathcal{V}}(\mathcal{N}_k)$ , i.e.,  $W_{\mathcal{V}}(\mathcal{N}_k)$  can in principle be obtained from the Boltzmann distribution. However, there is no such prescription in nonequilibrium. In fact, in nonequilibrium, the difficulty arises precisely here because, in most cases, the microscopic weights of the configurations in the steady state are *a priori* unknown. At this scenario, additivity, which originates from the simple physical consideration of statistical independence on the coarse-grained level of large subsystems, could help us to bypass the difficulty. As demonstrated recently in Refs. [8,9], to characterize fluctuation properties on a coarse-grained level, one may not actually be required to obtain the steady-state weights of all microscopic configurations. In fact, obtaining coarse-grained probability weights on a large scale (much larger than the microscopic correlation length scale) would suffice to characterize the macroscopic properties of the system, provided that additivity as in Eq. (1) holds.

It is important to note that the weight factor  $W_{\mathcal{V}}(\mathcal{N}_k)$  depends only on the subsystem particle number  $\mathcal{N}_k$  and subsystem volume  $\mathcal{V}$ . Now, provided that Eq. (1) holds, probability distribution function  $P_{\mathcal{V}}(\mathcal{N}) \equiv \text{Prob}[\mathcal{N}_k = \mathcal{N}]$  for large  $\mathcal{V}$  can be written as [5–8],

$$\begin{aligned} P_{\mathcal{V}}(\mathcal{N}) &\simeq W_{\mathcal{V}}(\mathcal{N}) \frac{Z(N - \mathcal{N}, V - \mathcal{V})}{Z(N, V)} \\ &= \frac{W_{\mathcal{V}}(\mathcal{N}) e^{\mu(\rho)\mathcal{N}}}{\mathcal{Z}}, \end{aligned} \quad (3)$$

where  $\mu(\rho)$  is a nonequilibrium chemical potential,

$$\mu(\rho) = \frac{df}{d\rho}, \quad (4)$$

$f(\rho)$  is a nonequilibrium free energy density function with  $Z(N, V) \simeq \exp[-Vf(\rho)]$  and  $\mathcal{Z} = \sum_{\mathcal{N}} W_{\mathcal{V}}(\mathcal{N}) \exp(\mu\mathcal{N})$  is the normalization constant. Importantly, free energy density function  $f(\rho)$ , or equivalently the large deviation function (LDF) which controls the density fluctuations, and chemical potential  $\mu(\rho)$  can now be obtained from a fluctuation-response relation (FR) between compressibility and fluctuation [5–8],

$$\frac{d\rho}{d\mu} = \sigma^2(\rho), \quad (5)$$

where

$$\sigma^2(\rho) = \lim_{\nu \rightarrow \infty} \frac{(\langle \mathcal{N}^2 \rangle - \langle \mathcal{N} \rangle^2)}{\nu}, \quad (6)$$

the scaled variance of subsystem particle number  $\mathcal{N}$ . The above fluctuation-response relation is analogous to the equilibrium fluctuation-dissipation theorem and follows directly from

Eq. (1); for details, see Appendix A. The explicit expression of chemical potential  $\mu(\rho)$  and free energy density function  $f(\rho)$  are obtained by integrating Eq. (5) w.r.t. density  $\rho$ ,

$$\mu(\rho) = \int \frac{1}{\sigma^2(\rho)} d\rho + c_1, \quad (7)$$

and, upon further integration,

$$f(\rho) = \int \mu(\rho) d\rho + c_2, \quad (8)$$

where  $c_1$  and  $c_2$  are arbitrary integration constants.

We now show, following Refs. [10,44], that the weight factor  $W_{\mathcal{V}}(\mathcal{N})$  and, consequently, the particle-number distribution can indeed be calculated using the above chemical potential and free energy function. First, we write Laplace transform (discrete) of the partition sum  $Z(N, V)$  as

$$\begin{aligned} \tilde{Z}(s, V) &= \sum_{N=0}^{\infty} e^{-sN} Z(N, V) \\ &= \sum_{N=0}^{\infty} e^{-sN} \sum_{\{\mathcal{N}_k\}} \left[ \prod_{k=1}^{v=\mathcal{V}/V} W_{\mathcal{V}}(\mathcal{N}_k) \right] \delta\left(\sum_k \mathcal{N}_k - N\right) \\ &= \prod_{k=1}^v \left[ \sum_{\mathcal{N}_k=0}^{\infty} e^{-s\mathcal{N}_k} W_{\mathcal{V}}(\mathcal{N}_k) \right] = [\tilde{W}_{\mathcal{V}}(s)]^v, \end{aligned}$$

where Laplace transform of the weight factor  $W_{\mathcal{V}}(\mathcal{N})$  is written as  $\tilde{W}_{\mathcal{V}}(s) = \sum_{\mathcal{N}=0}^{\infty} e^{-s\mathcal{N}} W_{\mathcal{V}}(\mathcal{N})$ . Now approximating  $\tilde{Z}(s, V) = \sum_{N=0}^{\infty} e^{-sN} Z(N, V) \simeq \int_0^{\infty} dN Z(N, V) e^{-sN}$  where we replace the sum by an integral and then, using  $Z(N, V) \simeq \exp[-Vf(\rho)]$  (by definition), we get

$$e^{-\nu h_{\mathcal{V}}(s)} \equiv V \int e^{-V[f(\rho)+s\rho]} d\rho \simeq [\tilde{W}_{\mathcal{V}}(s)]^{\nu},$$

where the function  $h_{\mathcal{V}}(s)$  is obtained from Legendre transform of free energy density function,

$$h_{\mathcal{V}}(s) = \mathcal{V}[\inf_{\rho} \{f(\rho) + s\rho\}]. \quad (9)$$

The weight factor  $W_{\mathcal{V}}(\mathcal{N})$  can, in principle, be calculated by evaluating the following integral on the complex  $s$  plane along a suitably chosen contour  $C$ :  $W_{\mathcal{V}}(\mathcal{N}) = 1/(2\pi i) \int_C \exp[-h_{\mathcal{V}}(s) + \mathcal{N}s] ds$ . Although, for finite  $\mathcal{V}$ , the explicit calculation of the weight factor may be difficult, the calculation, for large subsystem sizes  $\mathcal{V} \gg \xi$ , simplifies as the function  $-(1/\mathcal{V}) \ln W_{\mathcal{V}}(\mathcal{N})$  is related to  $h_{\mathcal{V}}(s)/\mathcal{V}$  by Legendre transformation [44]. Therefore, in the leading order of  $\mathcal{N} \gg 1$  and  $\mathcal{V} \gg \xi^d$ , the function  $-(1/\mathcal{V}) \ln W_{\mathcal{V}}(\mathcal{N})$  is nothing but the free energy density function  $f(\mathcal{N}/\mathcal{V})$ , implying

$$W_{\mathcal{V}}(\mathcal{N}) \simeq \exp[-\mathcal{V}f(\mathcal{N}/\mathcal{V})]. \quad (10)$$

Consequently, the subsystem particle-number distribution can be written as

$$P_{\mathcal{V}}(\mathcal{N}) \simeq \frac{e^{[-\mathcal{V}f(\mathcal{N}/\mathcal{V}) + \mu(\rho)\mathcal{N}]} }{\mathcal{Z}}, \quad (11)$$

where  $\mathcal{Z}(\mu, V)$  is the normalization constant.

### III. MODEL OF ACTIVE BROWNIAN PARTICLES AND FLUCTUATING HYDRODYNAMICS

To illustrate our theory in a particular model system of self-propelled particles, we study fluctuations in a system of interacting active Brownian particles (ABPs) in two dimensions (2D). We consider  $N$  particles in a 2D periodic box of size  $V = L \times L$ . At time  $t$ , the system is specified by position  $\mathbf{R}_i(t)$  and self-propulsion direction  $\phi_i(t)$  of  $i$ th particle with  $i = 1, \dots, N$ . The system evolves in time through the following overdamped Langevin equations [23], for the positions  $\{\mathbf{R}_i(t)\}$

$$\dot{\mathbf{R}}_i = -\beta D_0 \mathbf{F}_i + v_0 \mathbf{u}_i + \sqrt{2D_0} \tilde{\eta}_i^T$$

and for the orientations  $\{\phi_i(t)\}$  of the velocity vectors

$$\dot{\phi}_i = \sqrt{2D_r} \eta_i^R,$$

where  $\beta = 1/k_B T$  inverse temperature (we set  $\beta = 1$ ), force on  $i$ th particle  $\mathbf{F}_i = \sum_{j \neq i} \nabla U(|\mathbf{R}_i - \mathbf{R}_j|)$ , the WCA interaction potential  $U(r) = 4\epsilon[(a/r)^{12} - (a/r)^6] + \epsilon$  if  $r < 2^{1/6}a$  and zero otherwise,  $\epsilon = \beta^{-1}$ ,  $a$  diameter of the particles,  $v_0$  self-propulsion speed,  $\mathbf{u}_i \equiv \{u_{ix}, u_{iy}\} = \{\cos \phi_i, \sin \phi_i\}$  unit vector along instantaneous self-propulsion direction,  $D_0$  and  $D_r$  translational and rotational diffusion constant, respectively, and the  $\eta$  Gaussian white noises with  $\langle \eta_i \rangle = 0$  and  $\langle \eta_i(t) \eta_j(t') \rangle = \delta_{ij} \delta(t - t')$ .

To analytically study particle-number fluctuations in the active Brownian particles, we resort to a fluctuating hydrodynamic description, representing the system on a coarse-grained level. The following hydrodynamic equations, *without* the noise terms, for a density field  $\rho(\mathbf{r}, t)$  and a polarization density field  $\mathbf{p}(\mathbf{r}, t)$  have been previously obtained and studied for the active Brownian particles [24,27,34],

$$\partial_t \rho = -\nabla \cdot [v(\rho) \mathbf{p} - D(\rho) \nabla \rho + \mathbf{f}_d], \quad (12)$$

$$\partial_t \mathbf{p} = -D_r \mathbf{p} - \frac{1}{2} \nabla(v\rho) + K \nabla^2 \mathbf{p} + \mathbf{f}_p, \quad (13)$$

where  $\rho(\mathbf{r}, t) = \sum_i \delta[\mathbf{r} - \mathbf{R}_i(t)]$  and  $\mathbf{p}(\mathbf{r}, t) = \rho(\mathbf{r}, t) \mathbf{P}(\mathbf{r}, t) = \sum_i \delta[\mathbf{r} - \mathbf{R}_i(t)] \mathbf{u}_i(t)$  are coarse-grained number and polarization densities, respectively, at position  $\mathbf{r}$  and time  $t$  with  $\mathbf{R}_i(t)$  and  $\mathbf{u}_i(t)$  being position and velocity direction of the  $i$ th particle respectively,  $D(\rho)$  bulk diffusion constant,  $v(\rho)$  bulk velocity,  $\mathbf{f}_d$  and  $\mathbf{f}_p$  Gaussian noises specified below. Note that, to study fluctuations, we have added the noise terms  $\mathbf{f}_d$  and  $\mathbf{f}_p$  [20]: Gaussian multiplicative noises with zero mean and correlations  $\langle f_{dv}(\mathbf{r}, t) f_{dv'}(\mathbf{r}', t') \rangle = 2\Delta_d(\rho) \delta_{vv'} \delta(\mathbf{r} - \mathbf{r}') \delta(t - t')$  and  $\langle f_{pv}(\mathbf{r}, t) f_{pv'}(\mathbf{r}', t') \rangle = 2\Delta_p(\rho) \delta_{vv'} \delta(\mathbf{r} - \mathbf{r}') \delta(t - t')$ , with  $v, v' = 1, 2$  denoting Cartesian components. The strengths of the noise terms are not previously known and are characterized later (see Sec. IV B).

When  $v(\rho) \neq 0$ , the steady-state probability functional  $\mathcal{P}[\{\rho(\mathbf{r}), \mathbf{p}(\mathbf{r})\}]$  [47] *neither* has the Boltzmann distribution for the effective probability  $\mathcal{P}_d[\{\rho(\mathbf{r})\}] \propto \exp[-\int f[\rho(\mathbf{r})] d^2 \mathbf{r}]$  for density, *nor* it is in general known; only in special cases, a strictly local free energy functional  $f[\rho(\mathbf{r})]$  can be obtained [29,30,35]. However, additivity in Eq. (1) requires *neither* the existence of any Boltzmann-like distribution *nor* any prior knowledge of the full steady-state structure; it requires only the existence of a finite correlation length  $\xi(\rho)$  [see the relevant length scales  $\xi_{0,1,2}$  defined in Eqs. (16)

and (17)]. When  $v = 0$ ,  $\mathcal{P}_d[\{\rho(\mathbf{r})\}]$  can be exactly calculated; for details, see Appendix E.

#### IV. VARIANCE OF SUBSYSTEM PARTICLE NUMBER

##### A. Calculation of variance using linearized hydrodynamics

As discussed in Sec. II, using the fluctuation-response relation Eq. (5), subsystem particle-number distribution  $P_{\mathcal{V}}(\mathcal{N})$  for large  $\mathcal{V}$  can be determined solely from the variance of particle number, which requires knowledge of only two-point correlation function  $c(\mathbf{r}) = \langle \rho(0)\rho(\mathbf{r}) \rangle - \langle \rho(\mathbf{r}) \rangle^2$ . To this end, we transform the variable  $\theta(\mathbf{r}, t) = \nabla \cdot \mathbf{p}$  in Eq. (13) and, using the standard linear analysis, expand the nonlinear terms in Eqs. (12) and (13) up to linear order of  $\delta\rho$  and  $\delta\mathbf{p}$ , where  $\delta\rho = \rho - \rho_0$ ,  $\delta\mathbf{p} = \mathbf{p} - \mathbf{p}_0$ ,  $\delta\theta = \nabla \cdot (\delta\mathbf{p})$  with  $\rho_0$  and  $\mathbf{p}_0 = 0$  average density and polarization fields, respectively. Using Fourier transform of  $\delta\rho(\mathbf{r}, t)$  and  $\delta\theta(\mathbf{r}, t)$  in the linearized hydrodynamic equations,

$$\delta\tilde{\rho}(\mathbf{q}, \omega) = \int_{\mathbf{r}} \int_t e^{-i\mathbf{q}\cdot\mathbf{r}} e^{-i\omega t} \delta\rho(\mathbf{r}, t) d\mathbf{r} dt, \quad (14)$$

$$\delta\tilde{\theta}(\mathbf{q}, \omega) = \int_{\mathbf{r}} \int_t e^{-i\mathbf{q}\cdot\mathbf{r}} e^{-i\omega t} \delta\theta(\mathbf{r}, t) d\mathbf{r} dt, \quad (15)$$

and proceeding along the lines of Ref. [20], we obtain static structure factor

$$S(\mathbf{q}) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \langle |\delta\tilde{\rho}(\mathbf{q}, \omega)|^2 \rangle d\omega \equiv S_1(q) + S_2(q),$$

where

$$S_1(q) = \frac{V \Delta_d (\Sigma_1)^2}{\mathcal{D} \Sigma_0 \Sigma_2} + \frac{V \Delta_d q^2}{D_r \Sigma_2}, \quad (16)$$

$$S_2(q) = \frac{V \Delta_p v^2}{D_r^2 \mathcal{D} \Sigma_0 \Sigma_2}, \quad (17)$$

an effective diffusivity  $\mathcal{D}(\rho) = D + v\alpha/D_r$ ,  $\alpha(\rho) = (v + \rho dv/d\rho)/2$ ,  $\Sigma_{0,1,2}(q) = (1 + q^2 \xi_{0,1,2}^2)$  and correlation lengths  $\xi_0(\rho) = \sqrt{DK/D_r \mathcal{D}}$ ,  $\xi_1 = \sqrt{K/D_r}$  and  $\xi_2 = \sqrt{(D + K)/D_r}$ ; for details, see Appendix B. Now the variance  $\sigma_{\mathcal{V}}^2(\rho) = \langle \mathcal{N}^2 \rangle - \langle \mathcal{N} \rangle^2$  of particle number  $\mathcal{N} = \int_{\mathcal{V}} \rho(\mathbf{r}) d^2\mathbf{r}$  in a sub-volume  $\mathcal{V}$  can be written as integrated correlations,  $\sigma_{\mathcal{V}}^2(\rho) = \int_{\mathcal{V}} c(\mathbf{r}) d^2\mathbf{r} = S(\mathbf{q} = 0)$ . By defining a scaled variance  $\sigma_{\mathcal{V}}^2(\rho) = \sigma_{\mathcal{V}}^2/\mathcal{V}$ , we finally obtain the variance, albeit within an approximate linearized analysis,

$$\sigma^2(\rho) = \left[ \frac{\Delta_d}{\mathcal{D}} + \frac{\Delta_p v^2}{D_r^2 \mathcal{D}} \right], \quad (18)$$

which can be related to compressibility ( $d\rho/d\mu$ ) through the fluctuation-response relation [Eq. (5)]. The above linear analysis, though approximate, is expected to be valid in the regime of small fluctuations, i.e., far from criticality. A similar expression for structure factor was previously obtained in Ref. [20], though without the part  $S_1(\mathbf{q})$  and without any characterization of the noise strengths  $\Delta_d$  and  $\Delta_p$ . The functional dependence of the noise strengths  $\Delta_d$  and  $\Delta_p$  on density  $\rho$  and self-propulsion speed  $v_0$  will be determined in Sec. IV B. Note that, in the right-hand side of Eq. (18), effective diffusivity  $\mathcal{D}(\rho)$  appears in denominators of both

the terms, which could vanish for suitable parameter values; consequently, both  $S_1(\mathbf{q} = 0)$  and  $S_2(\mathbf{q} = 0)$  can separately diverge.

There are two interesting limiting cases of Eq. (18), which consistently capture various previous results obtained in the context of motility-induced phase separation (MIPS) in self-propelled particles.

*Case I.* To see that  $S_1(\mathbf{q})$  in Eq. (18) can have nontrivial effects, we consider the case when the polarization noise vanishes,  $\Delta_p = 0$ . In that case, the above linear analysis implies that  $S_1(0)$  diverges at a critical density, for any  $\Delta_d$ , whenever  $\mathcal{D} = 0$  ( $\mathcal{D} < 0$  corresponds phase coexistence) depending on the functional form of  $v(\rho)$ . This explains why, in the quasistatic case of  $\mathbf{p}$  where  $\partial_t \mathbf{p} = 0$ ,  $K = 0$  and  $\Delta_p = 0$  in Eq. (13), the variance  $\sigma^2(\rho) \simeq [1/\rho + (1/v)(dv/d\rho)]^{-1}$  obtained from Eq. (18) by choosing  $\Delta_d = (D + v^2/2D_r)\rho$  as in Ref. [29] and assuming  $D \ll v^2/2D_r$  (large velocity regime), can be diverging (for details, see Appendix D). This is because chemical potential  $\mu(\rho) = \ln(\rho v) + c_1$ , obtained using the fluctuation-response relation [Eq. (5)], has a singularity at the critical point where  $dv/d\rho = -v/\rho$  and consequently compressibility  $d\rho/d\mu = (d^2 f/d\rho^2)^{-1}$  diverges; the spinodal line is provided by the condition  $d^2 f/d\rho^2 < 0$ , which is consistent with the previous observations in various systems of self-propelled particles [29,35,38].

*Case II.* On the other hand, in the absence of density noise,  $\Delta_d = 0$  [20], only the second term in the right-hand side of Eq. (18) contributes to  $\sigma^2(\rho) = (2\Delta_p^0/D_r)[D/\rho v^2 + \{1/\rho + (1/v)(dv/d\rho)\}]^{-1}$  as, from the central limit theorem (CLT), the polarization noise strength  $\Delta_p \simeq \Delta_p^0 \rho$  is proportional to the number of particles present in unit volume (for details, see Appendix C). Integrating the fluctuation-response relation [Eq. (5)], we obtain  $\mu(\rho) = \ln(\rho v) + \psi(\rho) + c_1$  and  $f(\rho) = \int \mu(\rho) d\rho = \rho(\ln \rho - 1) + \int^{\rho} [\ln v(\rho) + \psi(\rho)] d\rho + c_1 \rho + c_2$  where  $\psi'(\rho) = D/\rho v^2$  and  $c_1$  and  $c_2$  arbitrary constants of integration. Indeed, the above expressions of  $\mu(\rho)$  and  $f(\rho)$  are quite similar to those obtained for the MIPS in the self-propelled particles (see Case I).

##### B. Calculation of noise strengths and fluctuations

The main difficulty to relate fluctuating hydrodynamic equations (12) and (13) to the microscopic model of active Brownian particles lies in the fact that the noise strengths  $\Delta_d$  and  $\Delta_p$ , bulk diffusion constant  $D$ , and bulk velocity  $v$  could depend on density  $\rho$ , self-propulsion velocity  $v_0$  (even function of  $v_0$ ), and possibly on the norm  $|\mathbf{p}|$ , but their functional forms are not explicitly known. In fact, a systematic derivation of the noise strengths from a microscopic dynamics is a difficult problem and, so far, has not achieved for the active Brownian particles.

In this section, we characterize strengths of the noises in the hydrodynamic equations (12) and (13), in the leading order of self-propulsion velocity  $v_0$ , i.e., when activity is low. To this end, we resort to a near-equilibrium analysis, which, we see later in simulations, however, holds surprisingly well even far from equilibrium where self-propulsion velocity, or the activity, is quite large. We first note that equilibrium compressibility of a two-dimensional hard-disk fluid, known



through virial coefficients [48], has an approximate analytic form [49],

$$[\sigma^2(\rho)]_{v_0=0} = \frac{\Delta_d}{D} \simeq \rho \left(1 - \frac{\rho}{\rho_m}\right)^2, \quad (19)$$

where  $\rho_m \approx 1.15$  close-packing density. Now, we expand  $\Delta_d$  in the leading order of self-propulsion velocity  $v_0$ ,

$$\Delta_d(\rho, v_0) \simeq (\Delta_d^0 + \Delta_d^1 v_0^2) \rho \left(1 - \frac{\rho}{\rho_m}\right)^2 \quad (20)$$

and write  $D \simeq D_0$  where  $\Delta_d^0$ ,  $\Delta_d^1$  and  $D_0$  are all constants (though not independent). These approximations may be the simplest possible ones, but they are quite good in describing the fluctuations in the active Brownian particles, as supported later in the simulations. As the relation in Eq. (19) must be satisfied in the equilibrium limit of  $v_0 = 0$ , we have  $\Delta_d^0/D_0 = 1$ . The dependence of  $D$ ,  $\Delta_d$ , and  $\Delta_p$  on the norm  $|\mathbf{p}|$  is ignored as orientation order  $\langle \mathbf{p} \rangle = 0$  throughout remains absent and the polarization fluctuation is expected to be small. Moreover, the polarization noise strength  $\Delta_p$ , to a good approximation, is expected to have a linear dependence on density,

$$\Delta_p \simeq \Delta_p^0 \rho, \quad (21)$$

where  $\Delta_p^0$  is a constant. This is because the fluctuation  $\sigma_{p_x}^2 = \langle p_x^2 \rangle - \langle p_x \rangle^2 \sim 2\Delta_p/D_r$  in the  $x$  component of polarization density  $\mathbf{p}$  can be written as  $\sigma_{p_x}^2 = \rho[\langle u_{ix}^2 \rangle - \langle u_{ix} \rangle^2] \propto \rho$  (similarly for the  $y$  component) where  $u_{ix}$  is the  $x$  component of the orientation unit vector of the  $i$ th particle and therefore  $\Delta_p \propto \rho$  (for details, see Appendix C). Therefore, the scaled variance as in Eq. (18) can be written as

$$\begin{aligned} \sigma^2(\rho) &= \left[ \frac{(\Delta_d^0 + \Delta_d^1 v_0^2) \rho (1 - \lambda \rho)^2}{D} + \frac{\Delta_p^0 \rho v^2}{D_r^2 \mathcal{D}} \right] \\ &= \frac{(\Delta_d^0 + \Delta_d^1 v_0^2) D_r^2 \rho (1 - \lambda \rho)^2 + \Delta_p^0 \rho v^2}{D^2 \mathcal{D}}. \end{aligned} \quad (22)$$

Now, using the previous results for the bulk velocity in the active Brownian particles,  $v(\rho) \simeq v_0(1 - \lambda \rho)$  [24,34] with  $1/\lambda = \rho_m$  close-packing density, in Eq. (18), effective diffusivity in the above equation can be written as

$$\mathcal{D} = D_0 + \frac{v\alpha}{D_r} = D_0 \left[ 1 + \frac{v_0^2(1 - \lambda \rho)(1 - 2\lambda \rho)}{2D_r D_0} \right],$$

where  $\alpha = (1/2)[v(\rho) + \rho dv/d\rho] = (v_0/2)[1 - 2\lambda \rho]$ . Therefore, Eq. (22) can be written as

$$\begin{aligned} \sigma^2(\rho) &= \rho_m \frac{(A + BP)x(1-x)^2}{1 + P(1-x)(1-2x)} \\ &= \rho_m \frac{(1 + A_1 v_0^2 + BP)x(1-x)^2}{1 + P(1-x)(1-2x)}, \end{aligned} \quad (23)$$

where  $x = \lambda \rho = \rho/\rho_m$  is scaled density, the dimensionless parameters  $A$ ,  $A_1$ , and  $B$  are defined as

$$A = (1 + A_1 v_0^2), \quad A_1 = \frac{\Delta_d^1}{\Delta_d^0}, \quad B = \frac{2\Delta_p^0}{D_r}, \quad (24)$$

and the dimensionless scaled activity parameter

$$P = \frac{v_0^2}{2D_r D_0}. \quad (25)$$

It is customary to define another dimensionless parameter, called activity parameter or Peclet number,  $\text{Pe} = v_0 \tau / a$ , where the microscopic diffusive time scale  $\tau = a^2/D_0$ . Now, using a near-equilibrium linear-response relation  $D_r = cD_0/a^2$  between the orientation (or the polarization) relaxation rate and the translational diffusion constant [50], we express  $P$  in terms of  $\text{Pe}$ ,

$$P \simeq \frac{\text{Pe}^2}{2c}, \quad (26)$$

where  $c$  is a proportionality constant and can be estimated from simulations [see Fig. 2(a) and the corresponding discussions later].

At low activity regime  $\text{Pe} \ll 1$ , one can actually reduce the number of parameters in Eq. (23), from  $A_1$ ,  $B$ , and  $P$  to essentially a single parameter  $P$ , using a constraint these parameters  $A_1$ ,  $B$ , and  $P$  must satisfy. It is not difficult to see that, at low density  $x = \rho/\rho_m \rightarrow 0$ , the particle-number distribution  $P_V(\mathcal{N}) = \exp(-\langle \mathcal{N} \rangle) \langle \mathcal{N} \rangle^{\mathcal{N}} / \mathcal{N}!$ , for any self-propulsion  $v_0$ , must be Poissonian (verified in simulations; see Fig. 3). Therefore, the variance at low density must satisfy the constraint  $\sigma^2(\rho) = \rho$ , implying

$$A_1 v_0^2 + BP = P, \quad (27)$$

or  $A_1 = (1 - B)/2D_r D_0$ . Note that Eq. (27) is exact in the leading order of self-propulsion velocity  $v_0$ . Using Eq. (27) in Eq. (23), we finally obtain the scaled variance as a function of scaled density  $x = \rho/\rho_m$ ,

$$\sigma^2(\rho) = \rho_m \frac{(1 + P)x(1-x)^2}{1 + P(1-x)(1-2x)}, \quad (28)$$

which essentially represents a one-parameter family of curves [see  $\sigma^2(\rho)$  as a function of  $\rho$  for various  $P$  in Fig. 1] with the scaled activity parameter  $P \simeq \text{Pe}^2/2c$  as in Eq. (26). Interestingly, as we find below in the simulations of the active Brownian particles, the form of the variance in Eq. (28) indeed captures quite well the broad features of particle-number fluctuations even when activity is moderately large,  $\text{Pe} \gg 1$ .

Now we show, using the form of the scaled variance in Eq. (28), how the scaled activity parameter  $P$  can be estimated from the simulations of the active Brownian particles. This is done essentially by fitting Eq. (28) for a suitable choice of the fitting parameter  $P$ . In Fig. 1 we plot scaled variance  $\sigma^2$  as a function of  $x = \rho/\rho_m$ , obtained from simulations for various  $\text{Pe} = 0, 5, 10, 20, 50$ , and 100, and then fit the curves with Eq. (28) by suitably choosing  $P \approx 0$  (magenta dashed line), 0.5 (blue dashed double-dotted line), 2.0 (sky-blue dotted line), 4.3 (green double-dashed dotted line), 8.0 (red solid line), and 10 (black dashed dotted line), respectively. To find the dependence of  $P$  on  $\text{Pe}$ , we numerically calculate  $P$  as a function of  $\text{Pe}$ , by solving for  $P$  where we use a particular value of scaled density  $x$  and variance  $\sigma^2$  in Eq. (28). In Fig. 2(a) we plot  $P$  as a function of  $\text{Pe}$ , for a set of two densities  $\rho = 0.26$  and 0.34. The function fits quite well with the form  $P \simeq \text{Pe}^2/(2c + \kappa \text{Pe}^2)$  [see the black solid line in Fig. 2(a)] where  $c \approx 9$ , implying a somewhat larger coarse-grained

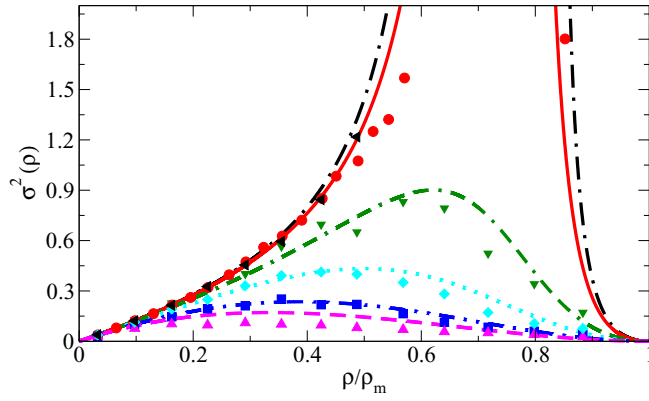


FIG. 1. Simulations in the active Brownian particles. Scaled variance  $\sigma^2$ , for  $Pe = 0$  (magenta triangles), 5 (blue squares), 10 (sky-blue diamonds), 20 (green inverted triangles), 50 (red circles) and 100 (black left-triangles), as a function of scaled density  $\rho/\rho_m$ , with  $\rho_m \approx 1.15$ , is compared with Eq. (28) with  $P \approx 0$  (magenta dashed line), 0.5 (blue dashed double-dotted line), 2.0 (sky-blue dotted line), 4.3 (green double-dashed dotted line), 8.0 (red solid line), and 10 (black dashed dotted line), respectively. Points = simulations, lines = theory.

relaxation rate  $D_r$  for the polarization field than that estimated previously [23], and  $\kappa \approx 0.1$ . In other words, at smaller activity regime  $Pe \lesssim 20$ , the scaled activity parameter  $P \simeq (Pe)^2/2c$  varies quadratically with  $Pe$  as in Eq. (26). However, for very large activity  $Pe \gtrsim 20$ , as discussed above, the scaled activity parameter  $P \simeq 1/\kappa$  eventually saturates.

## V. DENSITY FLUCTUATIONS AND NONEQUILIBRIUM THERMODYNAMICS

### A. Chemical potential and free energy function

In this section, we calculate, using the analytic form of the variance in Eq. (28), nonequilibrium chemical potential  $\mu(\rho)$  and free energy density function  $f(\rho)$ . We use the fluctuation-

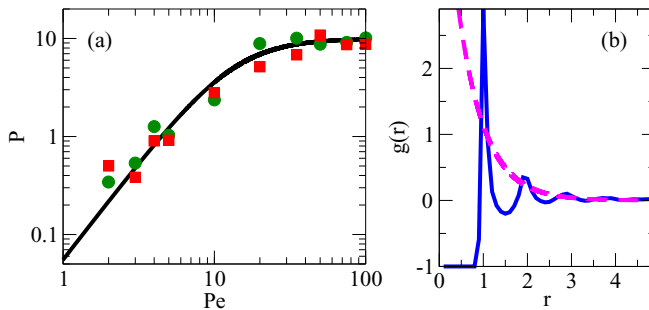


FIG. 2. (a) The scaled activity parameter  $P$  [as defined in Eq. (28)] is plotted as a function of Peclet number  $Pe$  for densities  $\rho \approx 0.26$  (green circles) and  $0.34$  (red squares). (b) Pair correlation  $g(r)$  (blue solid line) is plotted as a function of distance  $r$  for density  $\rho \approx 0.5$  and  $Pe = 50$ ; the magenta dashed line (fitting function) shows an exponential decay of the pair correlation function at large distance, with correlation length  $\xi \sim 1$  (distance is in unit of diameter  $a$  of the particles). Points = simulations, lines = theory with fitting parameter.

response relation [Eq. (5)], change the density variable  $\rho$  to a scaled density  $x = \rho/\rho_m$ , and integrate w.r.t. the scaled density  $x$ ,

$$\frac{d\rho}{d\mu} = \sigma^2(\rho) \Rightarrow \frac{dx}{d\mu} = \frac{(1+P)x(1-x)^2}{1+P(1-x)(1-2x)}, \quad (29)$$

to obtain nonequilibrium chemical potential as a function of the scaled density  $x$ ,

$$\mu(x) = \frac{1}{1+P} \left[ (P-1) \ln(1-x) + (P+1) \ln x + \frac{1}{1-x} \right],$$

which, upon substituting  $x = \rho/\rho_m$ , gives chemical potential as a function of density  $\rho$

$$\begin{aligned} \mu(\rho) = & \frac{1}{1+P} \left[ (P-1) \ln \left( 1 - \frac{\rho}{\rho_m} \right) \right. \\ & \left. + (P+1) \ln \left( \frac{\rho}{\rho_m} \right) + \frac{1}{1 - \frac{\rho}{\rho_m}} \right]. \end{aligned} \quad (30)$$

Now, integrating chemical potential  $\mu(\rho)$  w.r.t. density  $\rho$ , we get free energy density function

$$\begin{aligned} f(\rho) = & \int \mu d\rho = \rho_m \int \mu(x) dx \\ = & \rho_m \frac{[P(x-1) - x] \ln(1-x) + x[(P+1) \ln x - 2P]}{1+P}, \end{aligned} \quad (31)$$

which has the dimension of density and, upon substituting  $x = \rho/\rho_m$ , gives free energy density as a function of density  $\rho$ .

### B. Subsystem particle-number distributions

Nonequilibrium free energy density function in Eq. (31), being a large deviation function, and nonequilibrium chemical potential in Eq. (30), together, govern the particle-number fluctuation in the system. Therefore, based on the analytical result of subsystem particle-number distribution in Eq. (11) which can be explicitly calculated now using Eqs. (28), (30), and (31) (see Sec. II), we finally test in this section the predictions of additivity concerning density fluctuations in actual simulations of the active Brownian particles. The simulations are performed in the fluid phase, which is much away from criticality, where Eq. (28) is expected to hold.

In simulations, we calculate subsystem particle-number distributions  $P_V(\mathcal{N})$  in a subsystem ( $V = 9 \times 9$  in units of  $a$ ) where the rest of the system ( $V = 100 \times 100$ ) acts as a particle *reservoir* of chemical potential  $\mu(\rho)$ . In Fig. 2(b) we plot pair-correlation function  $g(r) = \sum_{i \neq 1} \langle \delta(\mathbf{r} - \mathbf{R}_i(t)) \rangle$  as a function of distance  $r$  at a moderately high density  $\rho \approx 0.5$  and  $Pe = 50$  where correlation length  $\xi \sim \mathcal{O}(a)$ , much smaller than the subsystem size. In Fig. 3 subsystem number distributions  $P_V(\mathcal{N})$  obtained from simulations (points) at  $Pe = 50$  are compared with theory Eq. (11) (lines) at the corresponding scaled activity  $P = 8$ , for several densities  $\rho \approx 0.11$  (black circles, black dashed line), 0.19 (red triangles, red dashed double-dotted line), 0.26 (magenta diamonds, magenta dotted line), 0.34 (green inverted triangles, green double-dashed dotted line), 0.41 (blue squares, blue solid line), and 0.56 (violet asterisks, violet dashed dotted line).

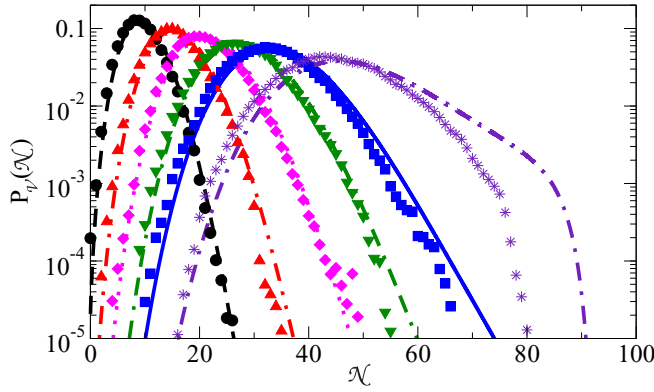


FIG. 3. Subsystem particle-number distributions for activity parameter  $Pe = 50$ , obtained from simulations (points), are compared with theory [Eq. (11)] (lines) with corresponding scaled activity parameter  $P = 8$ , for densities  $\rho \approx 0.11$  (black circles, black dashed line),  $0.19$  (red triangles, red dashed double-dotted line),  $0.26$  (magenta diamonds, magenta dotted line),  $0.34$  (green inverted triangles, green doubled-dashed dotted line),  $0.41$  (blue squares, blue solid line), and  $0.56$  (violet asterisks, violet dashed dotted line).

Agreement between simulations and theory, even at quite large density  $\rho \approx 0.41$ , is reasonably good. Note that, provided the variance  $\sigma^2(\rho)$  as a function of density  $\rho$  [as in Eq. (28)], there is no fitting parameter in the distribution functions  $P_V(\mathcal{N})$  plotted in Figs. 3 and 4. Expectedly, the distributions are Poissonian at low densities. However, the distributions become increasingly non-Poissonian, or non-Gaussian, with increasing density and activity. To emphasize this point, in Fig. 4, we show that, for moderately large density  $\rho \approx 0.41$  and large activity  $Pe = 50$ , the particle-number distribution function  $P_V(\mathcal{N})$  in simulations (blue squares) indeed deviates from the corresponding Poisson (black dashed double-dotted line) as well as Gaussian (red dashed line) distributions. Even then, the numerically obtained distribution (blue diamonds) is indeed quite well described by the analytically obtained distribution Eq. (11) (blue solid line), thus validating additivity, at least in the homogeneous fluid phase, which is sufficiently away from criticality.

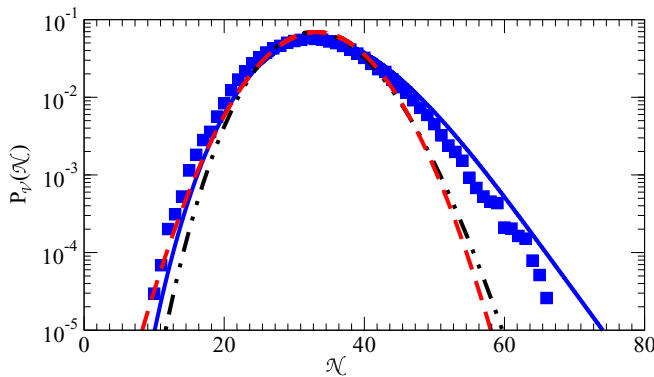


FIG. 4. At moderately large density  $\rho \approx 0.41$  and large activity  $Pe = 50$ , subsystem particle-number distribution  $P_V(\mathcal{N})$  (blue squares), which deviates from Poisson (black dashed double-dotted line) as well as Gaussian (red dashed line) distributions, is quite well captured by theory Eq. (11) (blue solid line).

However, upon approaching closer to the criticality, some discrepancies arise between analytic theory and simulations, presumably due to the linear analysis of Eqs. (12) and (13) and finite-size effects. That the linear analysis breaks down at density  $\rho \approx 0.56$  is evident from Fig. 1 where simulation results (for  $Pe = 50$ ) start deviating from the analytic expressions of Eq. (28) (for corresponding  $P = 8$ ). On the other hand, the finite-size effects originate from the facts that the boundary correlations between subsystem and system (due to increasing correlation length) increase while approaching criticality and the ratio between system and subsystem as well as their individual sizes are finite.

### C. Phase transition

Based on the analysis in the previous sections, one can now formulate a theory of phase transition in the active Brownian particles. We note that the functional form of the scaled variance as in Eq. (28) has many interesting implications. In the regime of large activity where Peclet number  $Pe \gg 1$  (or  $P \gg 1$ ), the scaled variance  $\sigma^2(\rho)$  is independent of  $P$ , which is in quite good agreement with simulations (see Fig. 1 for  $Pe = 50$  and  $100$ ) where  $\sigma^2$  for any  $\rho$  almost saturates at large Peclet number. Moreover, the denominator in Eq. (28) has two roots:

$$x_{2,1} = \frac{3}{4} \pm \frac{\sqrt{P^2 - 8P}}{4P}.$$

In Fig. 5(a) we plot the scaled variance [as in Eq. (28)] and, in Fig. 5(b), chemical potential [as in Eq. (30)] as a function of scaled density  $x = \rho/\rho_m$ , with  $\rho_m \approx 1.15$ , for various values of scaled activity  $P = 0, 1, 2, 4, 8, 12$ , and  $15$ . Below a critical value of the scaled activity  $P < P_c = 8$  (corresponding to activity  $Pe = 50$  in actual simulations), the variance remains positive in the full interval  $0 \leq x \leq 1$ . On the other hand, above a critical value of scaled activity  $P > P_c$ , the variance becomes negative in the interval  $x_1 < x < x_2$  and consequently chemical potential becomes nonmonotonic function of density, which is *not* physical and implies onset of phase coexistence. The coexisting densities can, in principle, be calculated using a Maxwell construction on chemical potential  $\mu(\rho) = \int^\rho 1/\sigma^2 d\rho$  [Eq. (30)] or on free energy density function  $f(\rho) = \int^\rho \mu d\rho$  [Eq. (31)]. Presently the Maxwell construction is however not expected to give an accurate estimate of the coexisting densities as our theory [Eq. (28) and consequent expressions in Eqs. (30) and (31)] have been derived using a linearized hydrodynamics and a near-equilibrium analysis, which do not capture well the fluctuations in the high activity regime.

Somewhat surprisingly, our theory however predicts quite accurately the critical density  $\rho_c$  where compressibility  $d\rho/d\mu$  diverges;  $\rho_c \approx 0.86$ , or critical packing fraction  $\phi_c \approx 0.7$ , obtained from theory is in excellent agreement with simulations [23]. Moreover, we find that compressibility diverges as  $d\rho/d\mu = \sigma^2 \sim 1/(\rho - \rho_c)^{\delta-1}$ , or equivalently chemical potential vanishes as  $\mu \sim (\rho - \rho_c)^\delta$ , with  $\delta = 3$ ; correlation length diverges as  $\xi \sim (\rho - \rho_c)^{-\nu_h}$  with mean field  $\nu_h = 1$ . Clearly, on the mean-field level, the exponents are in accordance with Ising universality;  $P$  and  $\mu$  are analogous to temperature and magnetic field, respectively.

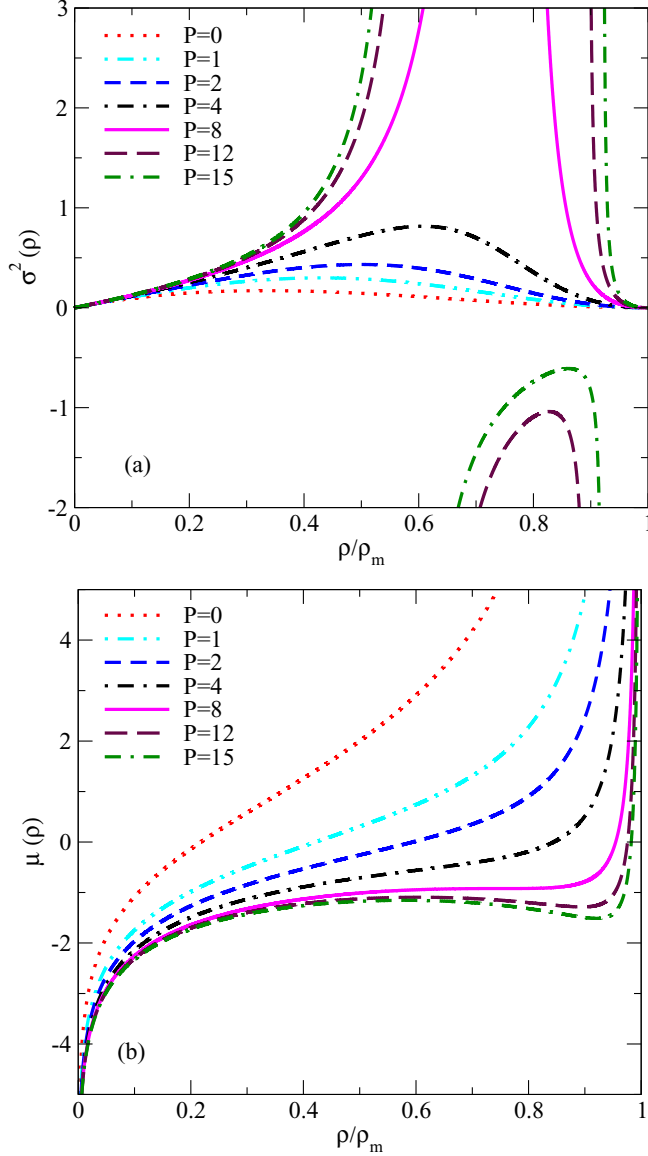


FIG. 5. Scaled variance  $\sigma^2(\rho)$  [Eq. (28)] and corresponding chemical potential  $\mu(\rho) = \int^\rho 1/\sigma^2 d\rho$  [Eq. (30)] are plotted in panels (a) and (b), respectively as a function of the scaled density  $\rho/\rho_m$  for various values of the scaled activity parameter  $P = 0, 1, 2, 4, 8, 12$ , and  $15$ . Chemical potential becomes a nonmonotonic function of density beyond a critical value of the scaled activity parameter,  $P > P_c = 8$ .

## VI. SUMMARY AND CONCLUDING PERSPECTIVE

In this paper, using a nonequilibrium fluctuation-response relation, which is a direct consequence of an additivity property, we formulate a thermodynamic theory for self-propelled particles in the context of a particular model system consisting of active Brownian particles. From the fluctuation-response relation, we demonstrate that subsystem particle-number distributions, which, being related to the density large deviation function and thus analogous to equilibrium free energy, can help us to characterize macroscopic properties in self-propelled particles, in a unified statistical mechanics framework, in terms of a nonequilibrium chemical potential.

Analogous to phase transition in equilibrium, as density and activity (Peclet number) increase, chemical potential becomes nonmonotonic function of density, indicating onset of a gas-liquid phase coexistence.

Importantly, the formalism developed here is solely based on characterization of the variance of subsystem particle number, which is directly related to the two-point (equal-time) density correlations or the structure factor. Provided that one calculates the structure factor accurately, our theory can lead to verifiable predictions concerning the density fluctuations. However, analytically calculating structure factor in a many-particle system is not an easy task. To this end, in the first step, we have calculated, though within a linearized fluctuating hydrodynamics, the structure factor in a microscopic model system of active Brownian particles and, consequently, the variance of subsystem particle-number as a function of density. Our studies of fluctuating hydrodynamics provide some insights in characterizing the noises in the hydrodynamic equations, done in the context of active Brownian particles. For this purpose, we have used a near-equilibrium analysis, which, though approximate, captures reasonably well the broad features of the particle-number fluctuations even in the far-from-equilibrium regime where activity is moderately large.

Furthermore, in the second step, from the functional dependence of the variance of subsystem particle-number on density and then using additivity, we have calculated the subsystem particle-number distribution function and have compared the analytically obtained distribution functions with that obtained from simulations in the active Brownian particles. The agreement between theory and simulations is quite good, except some deviations at the tails. The deviations increase while approaching criticality, indicating the following reasons for these deviations. First, the linear analysis used here breaks down in the regime of high densities and the high activities, where nonlinear effects can induce nontrivial fluctuations; consequently, chemical potential and free energy function obtained from the linear analysis cannot capture the density fluctuations well. Second, there can be significant finite-size effects, originating from the increasing boundary correlations between subsystem and system upon approaching criticality and due to the finite ratio between system and subsystem (simulations here are performed up to total particle number  $N > 10^4$  and roughly for ratio 1:123 between subsystem and system volumes). Thus, larger-scale simulations, though computationally difficult at this stage, would be quite useful for more accurate verification of the predictions of additivity.

For simplicity, here we have restricted ourselves to a particular model system of active Brownian particles. However, the thermodynamic formalism developed here is quite general and could be extended to other active-matter systems, e.g., models with run-and-tumble bacterial dynamics or the Vicsek model [45] and its variants. Moreover, even in the active Brownian particles, it would be quite interesting, though challenging, to go beyond the linear hydrodynamic regime by allowing nonlinear gradient terms (as in Ref. [32]) or self-advective terms (as in Refs. [45,46]), which may be relevant in the large activity regime where fluctuations are large.



From the overall perspective, we believe additivity could be the missing link, providing a unified characterization of a broad range of phenomena in the self-propelled particles observed in the past. Also, it will be interesting to explore if additivity holds in the phase with ‘‘giant number’’ fluctuations which many self-propelled particle systems exhibit or in the presence of inhomogeneities, e.g., a confining potential as in a box with hard walls [39], etc.

### ACKNOWLEDGMENTS

We thank Pradeep K. Mohanty for useful discussions. S.C. and P.P. acknowledge Thematic Unit of Excellence on Computational Materials Science, funded by the Department of Science and Technology (India), for the computational facility used in the present study. S.M. acknowledges financial support (under INSPIRE faculty award 2012) from the Department of Science and Technology (India).

### APPENDIX A: ADDITIVITY AND FLUCTUATION-RESPONSE RELATION

Here we show how additivity, as in Eq. (1), directly leads to the fluctuation-response relation as in Eq. (5). Provided the additivity property is satisfied, the subsystem particle-number distribution can be written, in the thermodynamic limit, as

$$P_{\mathcal{V}}(\mathcal{N}) = \frac{1}{\mathcal{Z}} W_{\mathcal{V}}(\mathcal{N}) e^{\mu \mathcal{N}}, \quad (\text{A1})$$

where  $\mathcal{W}$  is the weight factor for the respective subsystem,  $\mu$  is a chemical potential, and the normalization constant  $\mathcal{Z}$  is given by

$$\mathcal{Z}(\mu) = \sum_{\mathcal{N}=0}^{\infty} W_{\mathcal{V}}(\mathcal{N}) e^{\mu \mathcal{N}}. \quad (\text{A2})$$

Now, the average particle number can be calculated by taking a derivative of logarithm of the normalization constant w.r.t.  $\mu$ ,

$$\langle \mathcal{N} \rangle = \frac{d(\ln \mathcal{Z})}{d\mu}. \quad (\text{A3})$$

By taking another derivative of Eq. (A3) w.r.t.  $\mu$ , one can immediately relate compressibility to the fluctuation,

$$\frac{d\langle \mathcal{N} \rangle}{d\mu} = \frac{d^2(\ln \mathcal{Z})}{d\mu^2} = \langle \mathcal{N}^2 \rangle - \langle \mathcal{N} \rangle^2, \quad (\text{A4})$$

where, in the last step, we have used Eq. (A2). Dividing both side of the above equation by the subsystem volume  $\mathcal{V}$ , we get, in the limit of large  $\mathcal{V}$ , the fluctuation-response relation as in Eq. (13),

$$\frac{d\rho}{d\mu} = \sigma^2, \quad (\text{A5})$$

where the scaled variance is defined as

$$\sigma^2 = \lim_{\mathcal{V} \rightarrow \infty} \frac{(\langle \mathcal{N}^2 \rangle - \langle \mathcal{N} \rangle^2)}{\mathcal{V}}.$$

### APPENDIX B: CALCULATION OF STRUCTURE FACTOR IN THE ACTIVE BROWNIAN PARTICLES WITHIN LINEARIZED FLUCTUATING HYDRODYNAMICS

We consider the following fluctuating hydrodynamic equations, as considered in the main text, for self-propelled particles (SPPs)

$$\partial_t \rho = -\nabla \cdot [v(\rho) \mathbf{p} - D(\rho) \nabla \rho + \mathbf{f}_d], \quad (\text{B1})$$

$$\partial_t \mathbf{p} = -D_r \mathbf{p} - \frac{1}{2} \nabla (\rho v) + K \nabla^2 \mathbf{p} + \mathbf{f}_p, \quad (\text{B2})$$

and perform linear analysis along the lines of Ref. [20]. We transform the variable  $\theta(\mathbf{r}) = \nabla \cdot \mathbf{p}$ , expand the nonlinear terms up to linear order of  $\delta\rho$  and  $\delta\mathbf{p}$ , where  $\delta\rho = \rho - \rho_0$ ,  $\delta\mathbf{p} = \mathbf{p} - \mathbf{p}_0$ ,  $\delta\theta = \nabla \cdot (\delta\mathbf{p})$  with  $\rho_0$  and  $\mathbf{p}_0 = 0$  average density and polarization profile, to obtain

$$\partial_t \delta\rho(\mathbf{r}, t) = -v(\rho_0) \delta\theta(\mathbf{r}, t) + D(\rho_0) \nabla^2 \delta\rho(\mathbf{r}, t) - \nabla \cdot \mathbf{f}_d, \quad (\text{B3})$$

$$\begin{aligned} \partial_t \delta\theta(\mathbf{r}, t) = & -D_r \delta\theta(\mathbf{r}, t) - \alpha(\rho_0) \nabla^2 \delta\rho(\mathbf{r}, t) \\ & + K \nabla^2 \delta\theta(\mathbf{r}, t) + \nabla \cdot \mathbf{f}_p, \end{aligned} \quad (\text{B4})$$

where  $\nabla \cdot (v\mathbf{p}) \simeq v\delta\theta$ ,  $\nabla(\rho v) \simeq 2\alpha(\rho_0)\nabla(\delta\rho)$  with

$$2\alpha(\rho_0) = \frac{d}{d\rho_0} [\rho_0 v(\rho_0)] = v(\rho_0) + \rho_0 \frac{dv(\rho_0)}{d\rho_0}.$$

Using Fourier amplitudes

$$\delta\tilde{\rho}(\mathbf{q}, \omega) = \int_{\mathbf{r}} \int_t e^{-i\mathbf{q}\cdot\mathbf{r}} e^{-i\omega t} \delta\rho(\mathbf{r}, t) d\mathbf{r} dt, \quad (\text{B5})$$

$$\delta\tilde{\theta}(\mathbf{q}, \omega) = \int_{\mathbf{r}} \int_t e^{-i\mathbf{q}\cdot\mathbf{r}} e^{-i\omega t} \delta\theta(\mathbf{r}, t) d\mathbf{r} dt, \quad (\text{B6})$$

and reverting back to global density  $\rho_0 = \rho$  (for notational simplicity), Eqs. (B3) and (B4) can be written as

$$[i\omega + q^2 D(\rho)] \delta\tilde{\rho} + v(\rho) \delta\tilde{\theta} = -i\mathbf{q} \cdot \tilde{\mathbf{f}}_d, \quad (\text{B7})$$

$$\alpha(\rho) q^2 \delta\tilde{\rho} - (D_r + K q^2 + i\omega) \delta\tilde{\theta} = -i\mathbf{q} \cdot \tilde{\mathbf{f}}_p. \quad (\text{B8})$$

Solving for the Fourier modes, we get

$$\begin{aligned} & \begin{bmatrix} \delta\tilde{\rho} \\ \delta\tilde{\theta} \end{bmatrix} \\ & = -i\mathbf{q} \cdot \begin{bmatrix} q^2 D(\rho) + i\omega & v(\rho) \\ \alpha q^2 & -(D_r + K q^2 + i\omega) \end{bmatrix}^{-1} \begin{bmatrix} \tilde{\mathbf{f}}_d \\ \tilde{\mathbf{f}}_p \end{bmatrix} \end{aligned}$$

and therefore

$$\delta\tilde{\rho}(\mathbf{q}, \omega) = \frac{i}{\det(M)} [(D_r + K q^2 + i\omega) \mathbf{q} \cdot \tilde{\mathbf{f}}_d + v \mathbf{q} \cdot \tilde{\mathbf{f}}_p] \quad (\text{B9})$$

with

$$M = \begin{bmatrix} i\omega + q^2 D(\rho) & v(\rho) \\ \alpha q^2 & -(K q^2 + D_r + i\omega) \end{bmatrix}. \quad (\text{B10})$$

Using the noise correlations,  $\langle |\mathbf{q} \cdot \tilde{\mathbf{f}}_d|^2 \rangle = 2V \Delta_d q^2$ ,  $\langle |\mathbf{q} \cdot \tilde{\mathbf{f}}_p|^2 \rangle = 2V \Delta_p q^2$ , and  $\langle (\mathbf{q} \cdot \tilde{\mathbf{f}}_d^*)(\mathbf{q} \cdot \tilde{\mathbf{f}}_p) \rangle = \langle (\mathbf{q} \cdot \tilde{\mathbf{f}}_d)(\mathbf{q} \cdot \tilde{\mathbf{f}}_p^*) \rangle = 0$ ,

we obtain dynamic structure factor

$$\begin{aligned} S(\mathbf{q}, \omega) &= \langle |\delta \tilde{\rho}(\mathbf{q}, \omega)|^2 \rangle \\ &= \frac{\{\omega^2 + (D_r + Kq^2)^2\} \langle |\mathbf{q} \cdot \tilde{\mathbf{f}}_d|^2 \rangle + v^2 \langle |\mathbf{q} \cdot \tilde{\mathbf{f}}_p|^2 \rangle}{|\det(M)|^2} \\ &= \frac{2q^2 V}{|\det(M)|^2} [\Delta_d \{\omega^2 + (D_r + Kq^2)^2\} + \Delta_p v^2], \end{aligned} \quad (\text{B11})$$

where

$$|\det(M)|^2 = (\omega^2 - a)^2 + \omega^2 b^2 \quad (\text{B12})$$

with

$$\begin{aligned} a &= q^2 [D_r \mathcal{D}(\rho) + D(\rho) K q^2], \\ b &= D_r + q^2 [K + D(\rho)], \end{aligned}$$

and

$$\mathcal{D}(\rho) = D(\rho) + v(\rho) \alpha(\rho) / D_r.$$

The static structure factor can be computed as  $S(\mathbf{q}) = (1/2\pi) \int_{-\infty}^{\infty} S(\mathbf{q}, \omega) d\omega$ . Now using the following equalities,

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{d\omega}{(\omega^2 - a)^2 + \omega^2 b^2} &= \frac{\pi}{ab}, \\ \int_{-\infty}^{\infty} \frac{\omega^2 d\omega}{(\omega^2 - a)^2 + \omega^2 b^2} &= \frac{\pi}{b}, \end{aligned} \quad (\text{B13})$$

we obtain  $S(\mathbf{q}) = S_1(\mathbf{q}) + S_2(\mathbf{q})$  where

$$\begin{aligned} S_1(\mathbf{q}) &= V \frac{2q^2 \Delta_d}{2\pi} \left[ \frac{\pi}{b} + (D_r + Kq^2)^2 \frac{\pi}{ab} \right] \\ &= \frac{V \Delta_d q^2}{D_r + q^2(K + D)} \\ &\quad + \frac{V \Delta_d q^2 (D_r + Kq^2)^2}{q^2 [D_r \mathcal{D} + DKq^2] [D_r + q^2(K + D)]} \end{aligned} \quad (\text{B14})$$

and

$$\begin{aligned} S_2(\mathbf{q}) &= V \frac{2q^2 \Delta_p v^2}{2\pi} \frac{\pi}{ab} \\ &= V \frac{\Delta_p v^2}{[D_r \mathcal{D} + DKq^2] [D_r + q^2(K + D)]}. \end{aligned} \quad (\text{B15})$$

The structure factor  $S(\mathbf{q} = 0) = S_1(0) + S_2(0)$  is related to variance  $\sigma_V^2(\rho) = \langle \mathcal{N}^2 \rangle - \langle \mathcal{N} \rangle^2$  of number of particles  $\mathcal{N} = \int_V \rho(\mathbf{r}) d\mathbf{r}$  in a *subvolume*  $\mathcal{V}$ , which can be written as

$$\sigma_V^2(\rho) = S(\mathbf{q} = 0) = \mathcal{V} \left[ \frac{\Delta_d}{\mathcal{D}(\rho)} + \frac{\Delta_p v^2(\rho)}{D_r^2 \mathcal{D}(\rho)} \right], \quad (\text{B16})$$

the desired expression in the main text.

#### APPENDIX C: POLARIZATION FLUCTUATIONS IN THE ACTIVE BROWNIAN PARTICLES

As defined in the equations of motion for the active Brownian particles in the main text,  $\mathbf{u}_i \equiv \{u_{ix}, u_{iy}\} = \{\cos \phi_i, \sin \phi_i\}$  the orientation unit vector for the  $i$ th Brownian particle. From the definition of the polarization density  $\mathbf{p}(\mathbf{r}, t) = \sum_i \delta[\mathbf{r} -$

$\mathbf{R}_i(t)] \mathbf{u}_i(t)$ , we can express total polarization  $\mathbf{P}_{\Delta V}$ , in a small volume  $\Delta V$ , as

$$\mathbf{P}_{\Delta V} = \sum_{i \in \Delta V} \mathbf{u}_i(t), \quad (\text{C1})$$

where the sum is over  $\Delta N$  number of particles in the volume  $\Delta V$  so that

$$\mathbf{p} = \lim_{\Delta V \rightarrow 0} \frac{\mathbf{P}_{\Delta V}}{\Delta V}. \quad (\text{C2})$$

Now, using the CLT, one can estimate the fluctuation or the variance of  $\mathbf{P}_{\Delta V} \equiv \{P_{\Delta V}^x, P_{\Delta V}^y\}$ , which is the sum of  $\Delta N$  random variables (i.e., the sum of random orientation unit vectors of  $\Delta N$  particles in volume  $\Delta V$ ) where the variance of the  $x$  and  $y$  components of each orientation unit vector  $\mathbf{u}_i$ , for any  $i$ , are calculated to be constant. Consequently, the variance of the  $x$  and  $y$  components of polarization density  $\mathbf{p}$  can be calculated as

$$\begin{aligned} \sigma_{p_{ix}}^2 &= \langle p_{ix}^2 \rangle - \langle p_{ix} \rangle^2 \\ &= \lim_{\Delta V \rightarrow 0} \frac{\langle (P_{\Delta V}^x)^2 \rangle - \langle (P_{\Delta V}^x) \rangle^2}{\Delta V} \\ &= \lim_{\Delta V \rightarrow 0} \frac{\langle \Delta N \rangle}{\Delta V} [\langle u_{ix}^2 \rangle - \langle u_{ix} \rangle^2] \propto \rho \end{aligned} \quad (\text{C3})$$

and similarly

$$\sigma_{p_{iy}}^2 \propto \rho. \quad (\text{C4})$$

#### APPENDIX D: STRUCTURE FACTOR IN THE LIMIT OF QUASISTATIC POLARIZATION FIELD ( $\Delta_p = 0$ )

In the quasistatic limit of polarization field, by setting  $\partial_t \mathbf{p} = 0$ ,  $K = 0$ , and polarization noise strength  $\Delta_p = 0$  in Eq. (B2) and then substituting  $\mathbf{p}$  in Eq. (B2) [27,30], we get an effective evolution equation for density field,

$$\partial_t \rho = -\nabla \cdot \left[ -v \frac{\nabla(v\rho)}{2D_r} - D \nabla \rho + \mathbf{f}_d \right] \quad (\text{D1})$$

$$= -\nabla \cdot [\tilde{v} \rho - \tilde{D} \nabla \rho + \mathbf{f}_d], \quad (\text{D2})$$

where effective velocity  $\tilde{v}(\rho) = -v \nabla v / 2D_r$  and effective diffusivity  $\tilde{D} = D + v^2 / 2D_r$ . Now we perform linear analysis of fluctuation  $\delta \rho = \rho - \rho_0$  around the average density  $\rho_0$ ,

$$\partial_t \delta \rho = -\nabla \cdot [-\mathcal{D}(\rho_0) \nabla \delta \rho + \mathbf{f}_d], \quad (\text{D3})$$

where effective diffusivity

$$\mathcal{D} = \tilde{D} + \frac{\rho v}{2D_r} \frac{dv}{d\rho} = D + \frac{\alpha v}{D_r}, \quad (\text{D4})$$

where  $\alpha = (v + \rho \frac{dv}{d\rho}) / 2$ . Taking the Fourier transform of both sides and solving for density mode,

$$\delta \tilde{\rho}(q, \omega) = \frac{-iq}{i\omega + q^2 \mathcal{D}} \tilde{f}_d, \quad (\text{D5})$$

we calculate the dynamic structure factor,

$$S(\mathbf{q}, \omega) = \langle |\delta \tilde{\rho}(q, \omega)|^2 \rangle = \frac{2Vq^2 \Delta_d}{\omega^2 + q^4 \mathcal{D}^2}, \quad (\text{D6})$$

and the static structure factor,

$$S(\mathbf{q}) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\mathbf{q}, \omega) d\omega = \frac{V \Delta_d}{\mathcal{D}}, \quad (\text{D7})$$

which is independent of  $\mathbf{q}$ , i.e., the two-point correlation function  $c(\mathbf{r}) \propto \delta(\mathbf{r})$  and correlation length  $\xi = 0$ . This is the reason why the linear analysis is exact in this case and provides the variance exactly

$$\sigma^2(\rho) = \frac{\Delta_d}{D}, \quad (\text{D8})$$

which is consistent with an integrability condition in Ref. [29] and with the free energy density function  $f(\rho)$  satisfying  $d^2 f/d\rho^2 = 1/\sigma^2(\rho)$ . This could be seen if we choose  $\Delta_d = \tilde{D}\rho$  as in Ref. [29], and we find

$$\sigma^2(\rho) = \frac{\tilde{D}\rho}{\tilde{D} + \frac{\rho v}{2D_r} \frac{dv}{d\rho}} \simeq \frac{v^2 \rho / 2D_r}{\frac{v^2}{2D_r} + \frac{\rho v}{2D_r} \frac{dv}{d\rho}} = \left[ \frac{1}{\rho} + \frac{1}{v} \frac{dv}{d\rho} \right]^{-1}, \quad (\text{D9})$$

by assuming  $D \ll v^2/2D_r$ . Chemical potential  $\mu(\rho)$  can be obtained by integrating fluctuation-response relation [Eq. (5)] w.r.t. density  $\rho$ ,

$$\mu(\rho) = \int \frac{1}{\sigma^2(\rho)} d\rho = \ln(\rho v) + c_1, \quad (\text{D10})$$

$c_1$  being an arbitrary constant of integration.

#### APPENDIX E: FUNCTIONAL FOKKER-PLANCK EQUATION AND $v = 0$ LIMIT

For a functional Langevin equation (stochastic differential equation) having a general form,

$$\partial_t \rho(\mathbf{r}, t) = B[\rho(\mathbf{r}, t)] + g(\mathbf{r}, t), \quad (\text{E1})$$

where  $B[\rho(\mathbf{r}, t)]$  is a functional of  $\rho(\mathbf{r})$  and  $g(\mathbf{r}, t)$  is a Gaussian noise with correlation

$$\langle g(\mathbf{r}, t) g(\mathbf{r}', t') \rangle = G(\mathbf{r}, \mathbf{r}') \delta(t - t'), \quad (\text{E2})$$

the functional Fokker-Planck equation is given by [47]

$$\begin{aligned} \partial_t \mathcal{P}[\rho(\mathbf{r}, t)] &= - \int d^3 \mathbf{r} \frac{\delta}{\delta \rho(\mathbf{r})} \{ B[\rho(\mathbf{r})] \mathcal{P}[\rho(\mathbf{r}, t)] \} \\ &+ \frac{1}{2} \iint d^3 \mathbf{r} d^3 \mathbf{r}' \left[ \frac{\delta^2}{\delta \rho(\mathbf{r}) \delta \rho(\mathbf{r}')} G(\mathbf{r}, \mathbf{r}') \mathcal{P}[\rho(\mathbf{r}, t)] \right]. \end{aligned} \quad (\text{E3})$$

For simplicity, let us consider only one spatial dimension with Cartesian position coordinate  $x$ . Now, we are interested in a Langevin equation having a particular form

$$\dot{\rho} = -\partial_x [\rho v - D(\rho) \partial_x \rho + f_\rho], \quad (\text{E4})$$

where  $g(x, t) = \partial_x f_\rho(x, t)$  with noise correlation  $\langle f_\rho(x, t) f_\rho(x', t') \rangle = 2\Delta(\rho) \delta(x - x') \delta(t - t')$ . Using Eq. (E3), the functional Fokker-Planck equation becomes

$$\partial_t \mathcal{P} = \int dx \frac{\delta}{\delta \rho(x)} \partial_x \left[ \rho v - D(\rho) \partial_x \rho - \Delta(\rho) \partial_x \frac{\delta}{\delta \rho(x)} \right] \mathcal{P}. \quad (\text{E5})$$

For nonzero  $v(\rho) \neq 0$ , solution of the above Fokker-Planck equation is not in general known. In a special case, when an integrability condition is satisfied  $v/D = \partial_x (\delta F / \delta \rho)$  for a functional  $F[\rho(x)] = \int f[\rho(x)] dx$ , the steady-state solution is given by the Boltzmann form  $\mathcal{P} \sim \exp[-F[\rho(x)]]$  [29].

When velocity  $v(\rho) = 0$ , i.e., in equilibrium, the Fokker-Planck equation for the many body probability  $\mathcal{P}[\rho(x), t]$  can be shown to always have the Boltzmann form as follows. The Fokker-Planck equation in this case can be written as

$$\dot{\mathcal{P}} = \int dx \frac{\delta}{\delta \rho(x)} \partial_x \left[ -D(\rho) \frac{\partial \rho}{\partial x} - \Delta(\rho) \left( \frac{\delta}{\delta \rho} \right)' \right] \mathcal{P}. \quad (\text{E6})$$

We start with an ansatz  $\mathcal{P} \sim \exp[-\int f(\rho) dx]$  and, using

$$\left( \frac{\delta \mathcal{P}}{\delta \rho} \right)' = -\mathcal{P} \frac{d^2 f}{d\rho^2} \frac{\partial \rho}{\partial x}, \quad (\text{E7})$$

in Eq. (E6), we obtain  $f(\rho)$  as

$$-D(\rho) \frac{\partial \rho}{\partial x} \mathcal{P} - \Delta(\rho) \left( \frac{\delta \mathcal{P}}{\delta \rho} \right)' = 0 \quad (\text{E8})$$

$$\Rightarrow -D(\rho) \frac{\partial \rho}{\partial x} \mathcal{P} + \Delta(\rho) \mathcal{P} \frac{d^2 f}{d\rho^2} \frac{\partial \rho}{\partial x} = 0 \quad (\text{E9})$$

$$\Rightarrow \frac{\partial \rho}{\partial x} \mathcal{P} \left[ -D(\rho) + \Delta(\rho) \frac{d^2 f}{d\rho^2} \right] = 0 \quad (\text{E10})$$

$$\Rightarrow \frac{d^2 f}{d\rho^2} = \frac{D(\rho)}{\Delta(\rho)}. \quad (\text{E11})$$

Therefore the steady-state probability functional for density fluctuation can be written as  $\mathcal{P}[\{\rho(\mathbf{r})\}] \propto \exp[-\int f[\rho(\mathbf{r})] d^2 \mathbf{r}]$  where  $d^2 f/d\rho^2 = [\Delta_d(\rho)/D(\rho)]^{-1}$ . This is what is expected from the equilibrium fluctuation-dissipation theorem.

- [1] G. Gallavotti and E. G. D. Cohen, *Phys. Rev. Lett.* **74**, 2694 (1995); D. J. Evans and G. P. Morriss, *Statistical Mechanics of Nonequilibrium Fluids* (Academic Press, New York, 1990).  
 [2] B. Derrida, J. L. Lebowitz, and E. R. Speer, *Phys. Rev. Lett.* **87**, 150601 (2001); **89**, 030601 (2002); B. Derrida, *J. Stat. Mech.* (2011) P01030.  
 [3] Y. Oono and M. Paniconi, *Prog. Theor. Phys. Suppl.* **130**, 29 (1998).  
 [4] L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio, and C. Landim, *Rev. Mod. Phys.* **87**, 593 (2015).

- [5] G. L. Eyink, J. L. Lebowitz, and H. Spohn, *J. Stat. Phys.* **83**, 385 (1996).  
 [6] E. Bertin, O. Dauchot, and M. Droz, *Phys. Rev. Lett.* **96**, 120601 (2006).  
 [7] P. Pradhan, C. P. Amann, and U. Seifert, *Phys. Rev. Lett.* **105**, 150601 (2010).  
 [8] S. Chatterjee, P. Pradhan, and P. K. Mohanty, *Phys. Rev. Lett.* **112**, 030601 (2014).  
 [9] S. Chatterjee, P. Pradhan, and P. K. Mohanty, *Phys. Rev. E* **91**, 062136 (2015).

- [10] A. Das, S. Chatterjee, P. Pradhan, and P. K. Mohanty, *Phys. Rev. E* **92**, 052107 (2015).
- [11] E. Ben-Jacob, I. Cohen, and H. Levine, *Adv. Phys.* **49**, 395 (2000); C. Dombrowski, L. Cisneros, S. Chatkaew, R. E. Goldstein, and J. O. Kessler, *Phys. Rev. Lett.* **93**, 098103 (2004).
- [12] S. Hubbard, P. Babak, S. Sigurdsson, and K. Magnusson, *Ecol. Model.* **174**, 359 (2004).
- [13] T. Feder, *Phys. Today* **60**(10), 28 (2007); C. Feare, *The Starling* (Oxford University Press, Oxford, 1984).
- [14] E. Rauch, M. Millonas, and D. Chialvo, *Phys. Lett. A* **207**, 185 (1995).
- [15] J. Palacci, S. Sacanna, A. P. Steinberg, D. J. Pine, P. M. Chaikin, *Science* **339**, 936 (2013); H. R. Jiang, N. Yoshinaga, and M. Sano, *Phys. Rev. Lett.* **105**, 268302 (2010).
- [16] H. H. Wensink, J. Dunkel, S. Heidenreich, K. Drescher, R. E. Goldstein, H. Lowen, and J. M. Yeomans, *Proc. Natl. Acad. Sci. U. S. A.* **109**, 14308 (2012); J. Dunkel, S. Heidenreich, K. Drescher, H. H. Wensink, M. Bar, and R. E. Goldstein, *Phys. Rev. Lett.* **110**, 228102 (2013); J. Palacci, C. Cottin-Bizonne, C. Ybert, and L. Bocquet, *ibid.* **105**, 088304 (2010); F. Peruani, J. Starruss, V. Jakovljevic, L. Sogaard-Andersen, A. Deutsch, and M. Bar, *ibid.* **108**, 098102 (2012).
- [17] S. Mishra, A. Baskaran, and M. C. Marchetti, *Phys. Rev. E* **81**, 061916 (2010).
- [18] F. Peruani, T. Klaus, A. Deutsch, and A. Voss-Boehme, *Phys. Rev. Lett.* **106**, 128101 (2011).
- [19] A. Peshkov, I. S. Aranson, E. Bertin, H. Chate, and F. Ginelli, *Phys. Rev. Lett.* **109**, 268701 (2012).
- [20] Y. Fily and M. C. Marchetti, *Phys. Rev. Lett.* **108**, 235702 (2012); Y. Fily, S. Henkes, and C. Marchetti, *Soft Matter* **10**, 2132 (2014).
- [21] F. D. C. Farrell, M. C. Marchetti, D. Marenduzzo, and J. Tailleur, *Phys. Rev. Lett.* **108**, 248101 (2012); H. H. Wensink, V. Kantsler, R. E. Goldstein, and J. Dunkel, *Phys. Rev. E* **89**, 010302 (2014).
- [22] J. Bialke, T. Speck, and H. Lowen, *Phys. Rev. Lett.* **108**, 168301 (2012).
- [23] G. S. Redner, M. F. Hagan, and A. Baskaran, *Phys. Rev. Lett.* **110**, 055701 (2013).
- [24] J. Bialke, H. Lowen, and T. Speck, *Europhys. Lett.* **103**, 30008 (2013).
- [25] S. Ngo, A. Peshkov, I. S. Aranson, E. Bertin, F. Ginelli, and H. Chate, *Phys. Rev. Lett.* **113**, 038302 (2014).
- [26] M. C. Marchetti, J. F. Joanny, S. Ramaswamy, T. B. Liverpool, J. Prost, M. Rao, and R. A. Simha, *Rev. Mod. Phys.* **85**, 1143 (2013); M. Cates, *Rep. Prog. Phys.* **75**, 042601 (2012).
- [27] A. P. Solon, M. E. Cates, and J. Tailleur, *Eur. Phys. J. Special Topics* **224**, 1231 (2015).
- [28] A. Y. Grosberg and J.-F. Joanny, *Phys. Rev. E* **92**, 032118 (2015).
- [29] J. Tailleur and M. E. Cates, *Phys. Rev. Lett.* **100**, 218103 (2008).
- [30] M. Cates and J. Tailleur, *Europhys. Lett.* **101**, 20010 (2013).
- [31] J. Stenhammar, A. Tiribocchi, R. J. Allen, D. Marenduzzo, and M. E. Cates, *Phys. Rev. Lett.* **111**, 145702 (2013).
- [32] R. Wittkowski, A. Tiribocchi, J. Stenhammar, R. J. Allen, D. Marenduzzo, and M. E. Cates, *Nat. Comm.* **5**, 4351 (2014).
- [33] M. E. Cates, D. Marenduzzo, I. Pagonabarraga, and J. Tailleur, *Proc. Natl. Acad. Sci. U. S. A.* **107**, 11715 (2010).
- [34] T. Speck, J. Bialke, A. M. Menzel, and H. Lowen, *Phys. Rev. Lett.* **112**, 218304 (2014).
- [35] J. Barre, R. Chetrite, M. Muratori, and F. Peruani, *J. Stat. Phys.* **158**, 589 (2015).
- [36] S. C. Takatori, W. Yan, and J. F. Brady, *Phys. Rev. Lett.* **113**, 028103 (2014); S. C. Takatori and J. F. Brady, *Phys. Rev. E* **91**, 032117 (2015).
- [37] S. A. Mallory, A. Saric, C. Valeriani, and A. Cacciuto, *Phys. Rev. E* **89**, 052303 (2014).
- [38] A. P. Solon, J. Stenhammar, R. Wittkowski, M. Kardar, Y. Kafri, M. E. Cates, and J. Tailleur, *Phys. Rev. Lett.* **114**, 198301 (2015).
- [39] A. P. Solon, Y. Fily, A. Baskaran, M. E. Cates, Y. Kafri, M. Kardarand, and J. Tailleur, *Nat. Phys.* **11**, 673 (2015).
- [40] X. Yang, M. L. Manning, and M. Marchetti, *Soft Matter* **10**, 6477 (2014).
- [41] R. G. Winkler, A. Wysocki, and G. Gompper, *Soft Matter* **11**, 6680 (2015).
- [42] D. Levis and L. Berthier, *Europhys. Lett.* **111**, 60006 (2015).
- [43] U. M. B. Marconi and C. Maggi, *Soft Matter* **11**, 8768 (2015).
- [44] H. Touchette, *Phys. Rep.* **478**, 1 (2009).
- [45] T. Vicsek, A. Czirok, E. Ben-Jacob, I. Cohen, and O. Shochet, *Phys. Rev. Lett.* **75**, 1226 (1995); J. Toner and Y. Tu, *ibid.* **75**, 4326 (1995).
- [46] S. Heidenreich, J. Dunkel, S. H. L. Klapp, and M. Bar, [arXiv:1509.08661](https://arxiv.org/abs/1509.08661).
- [47] C. W. Gardiner, *Handbook of Stochastic Methods*, 3rd ed. (Springer-Verlag, Berlin, 2004).
- [48] N. Clisby and B. M. McCoy, *J. Stat. Phys.* **114**, 1343 (2004); **114**, 1361 (2004).
- [49] R. Hoste and W. van Dael, *J. Chem. Soc., Faraday Trans. 2* **80**, 477 (1984).
- [50] M. G. Mazza, N. Giovambattista, H. E. Stanley, and F. W. Starr, *Phys. Rev. E* **76**, 031203 (2007); C. K. Mishra and R. Ganapathy, *Phys. Rev. Lett.* **114**, 198302 (2015); J. Kim and B. J. Sung, *ibid.* **115**, 158302 (2015).