Effective equilibrium states in the colored-noise model for active matter I. Pairwise forces in the Fox and Unified Colored Noise Approximations

René Wittmann,[1, a)] C. Maggi,[2] A. Sharma,[1] A. Scacchi,[1] J. M. Brader,[1] and U. Marini Bettolo Marconi[3]

[1)] *Department of Physics, University of Fribourg, CH-1700 Fribourg, Switzerland*
[2)] *NANOTEC-CNR, Institute of Nanotechnology, Soft and Living Matter Laboratory, Piazzale A. Moro 2, I-00185, Roma, Italy*
[3)] *Scuola di Scienze e Tecnologie, Università di Camerino, Via Madonna delle Carceri, 62032, Camerino, INFN Perugia, Italy*

The equations of motion of active systems can be modeled in terms of Ornstein-Uhlenbeck processes (OUPs) with appropriate correlators. For further theoretical studies, these should be approximated to yield a Markovian picture for the dynamics and a simplified steady-state condition. We perform a comparative study of the Unified Colored Noise Approximation (UCNA) and the approximation scheme by Fox recently employed within this context. We review and discuss in detail the approximations associated with effective interaction potentials that are defined in the low-density limit. We study the conditions for which the behavior observed in two-body simulations for the OUPs model and Active Brownian particles can be represented. Finally, we discuss different ways to define a force-balance condition in the limit of small activity.

I. INTRODUCTION

Active Brownian particles (ABPs) provide a simple, minimal model system with which to study the collective behavior of active matter. Many-body Brownian dynamics simulations of these systems have provided considerable insight into a range of interesting nonequilibrium phenomena, such as the accumulation of particles at boundaries[1–5] and motility-induced phase separation[6]. Much of the phenomenology of ABPs can be captured using coarse-grained, hydrodynamic theories[6–9], however, these approaches are based on general physical arguments which do not enable a direct connection to be made to the microscopic interparticle interactions. Although some progress has recently been made in this direction[10], this remains work in progress.

Due to the inherent difficulty of dealing simultaneously with both the translational and orientational degrees of freedom in active systems, attempts to develop a first-principles theory have largely focused on a simpler, related model, in which the particle dynamics are represented by a set of coupled Ornstein-Uhlenbeck processes (OUPs). Within this model an exponentially correlated noise term, with a given correlation time, serves as proxy for the persistent trajectories of ABPs (connections between the two models were explored in Ref. 11). While the removal of orientational degrees of freedom does indeed simplify the problem, it comes at the cost that one has to deal with the non-Markovian dynamics of the translational coordinates. Fortunately, there exist approximation methods[12–16] which enable the OUP model to be represented using an effective Markovian, and therefore tractable, dynamics. There are two different approaches to doing this, (i) the Unified Colored Noise Approximation (UCNA) of Hänggi *et al.*[13,14], based on

adiabatic elimination on the level of the Langevin equations, (ii) the Fox approximation[15,16], for which an approximate Fokker-Planck equation is developed.

When applied in the context of active matter, both the UCNA and Fox approximations are referred to as 'effective equilibrium' approaches, owing to their Markovian character. Indeed, the possibility of mimicking the behavior of nonequilibrium ABPs using an equilibrium system of passive particles, interacting via *effective* interactions, was suggested by several researchers (see e.g.[17]) who observed that the phase separation induced by activity in systems of repulsive ABPs closely resembled that familiar from passive systems with an attractive interaction. Despite its appeal, several years were required before this observation could be turned into something more concrete. By starting from the simpler OUP model it became possible, via application of the UCNA and Fox approximations, to put the notion of an effective equilibrium description on a firmer footing[11,18–24].

In this paper, we will compare and contrast the two different approaches to effective equilibrium. We will highlight the main approximations involved and assess the validity of the effective-potential approximation (EPA), which has been employed in previous work to investigate activity induced modifications of the microstructure and motility-induced phase separation[11]. This analysis both clarifies the nature of the approximations involved and suggests ways in which the description can be improved.

The paper is laid out as follows: In Sec. II we first specify the model under consideration. The UCNA and Fox approaches to obtaining an effective equilibrium are then described in Sec. III, highlighting similarities and differences between them. In Sec. IV we describe in detail the EPA, where the emphasis is placed on the UCNA due to its simpler structure. The resulting approximate effective potentials are compared to computer simulations in Sec. V using a standard soft-repulsive and a non-convex (Gaussian core) potential. In Sec. VI we consider an al-

[a)]Electronic mail: rene.wittmann@unifr.ch

ternative approach to obtain pairwise forces, i.e., the low-activity limit, and make contact to the EPA in Sec. VII. Finally, we conclude in Sec. VIII.

II. COLORED-NOISE MODEL

In this section, we introduce the common starting point of both the UCNA and Fox approach. Since particles driven by Gaussian colored noise originally were not intended as a model for an active system, the choice of parameters in the literature may depend on the dimensionality and on whether contact to ABPs is made[11] or not[18]. We will also clarify some notational issues.

We consider the coupled stochastic (Langevin) differential equations

$$\dot{\mathbf{r}}_i(t) = \gamma^{-1}\mathbf{F}_i(\mathbf{r}_1, \ldots, \mathbf{r}_N) + \boldsymbol{\xi}_i(t) + \boldsymbol{v}_i(t) \qquad (1)$$

of N particles. The motion of each particle $i \in \{1, \ldots, N\}$ at position $\mathbf{r}_i(t)$ is determined by conservative $\mathbf{F}_i$ and stochastic forces $\gamma\boldsymbol{\xi}_i$ and $\gamma\boldsymbol{v}_i$. The friction coefficient $\gamma = (\beta D_\mathrm{t})^{-1}$ is related to the translational Brownian diffusivity D_t and $\beta = (k_\mathrm{B}T)^{-1}$ is the inverse temperature. We assume that the total interaction force $\mathbf{F}_i(\mathbf{r}^N) = -\boldsymbol{\nabla}_i\mathcal{U}(\mathbf{r}^N)$ can be written as the gradient of a pairwise additive many-body potential

$$\mathcal{U}(\mathbf{r}^N) = \left(V_\mathrm{ext}(\mathbf{r}_i) + \frac{1}{2}\sum_{k \neq i}^N u(\mathbf{r}_i, \mathbf{r}_k) \right), \qquad (2)$$

consisting of the one-body external fields $V_\mathrm{ext}(\mathbf{r}_i)$ and the interparticle potentials $u(\mathbf{r}_i, \mathbf{r}_k) = u(|\mathbf{r}_i - \mathbf{r}_k|)$.

The vector $\boldsymbol{\xi}_i(t)$ represents the translational Brownian diffusion by a Gaussian (white) noise of zero mean and $\langle \boldsymbol{\xi}_i(t)\boldsymbol{\xi}_j(t') \rangle = 2D_\mathrm{t}\mathbf{1}\delta_{ij}\delta(t - t')$ with the unit matrix $\mathbf{1}$. Here and in the following the dyadic product of two vectors with $\mathfrak{d}$ components results in a $\mathfrak{d} \times \mathfrak{d}$ matrix. Any contraction as in a scalar product or a matrix-vector product will be explicitly indicated by a ".". The OUPs $\boldsymbol{v}_i(t)$ defined by

$$\dot{\boldsymbol{v}}_i(t) = -\frac{\boldsymbol{v}_i(t)}{\tau_\mathrm{a}} + \frac{\boldsymbol{\eta}_i(t)}{\tau_\mathrm{a}} \qquad (3)$$

with $\langle \boldsymbol{\eta}_i(t)\boldsymbol{\eta}_j(t') \rangle = 2D_\mathrm{a}\mathbf{1}\delta_{ij}\delta(t-t')$ describe a fluctuating propulsion velocity as a non-Gaussian (colored) noise of zero mean and

$$\langle \boldsymbol{v}_i(t)\boldsymbol{v}_j(t') \rangle = \frac{v_0^2}{\mathfrak{d}}\mathbf{1}\delta_{ij}e^{-\frac{|t-t'|}{\tau_\mathrm{a}}} = \frac{D_\mathrm{a}}{\tau_\mathrm{a}}\mathbf{1}\delta_{ij}e^{-\frac{|t-t'|}{\tau_\mathrm{a}}} . \qquad (4)$$

Here we introduced the active time scale τ_a at which the orientation randomizes and the active diffusion coefficient $D_\mathrm{a} = v_0^2\tau_\mathrm{a}/\mathfrak{d}$, where $v_0^2 = \langle \boldsymbol{v}_i^2(t) \rangle$ is the average squared self-propulsion velocity and $\mathfrak{d}$ the spatial dimension.

The colored-noise model for active particles contains two parameters describing the magnitude and persistence of the self propulsion. We now aim to clarify some notational differences in the literature. The persistence time τ_a can be explicitly related to the equations of motion of run and tumble particles[25] or active Brownian particles[11,26]. The above definitions correspond to the latter case with $\tau_\mathrm{a} = D_\mathrm{r}^{-1}/(\mathfrak{d} - 1)$, where D_r is the rotational Brownian diffusion coefficient. Another common choice[21] amounts to consider $\tau_\mathrm{R} = D_\mathrm{r}^{-1}$ and $D_\mathrm{a} = v_0^2\tau_\mathrm{R}/(\mathfrak{d} - 1)/\mathfrak{d}$. In the following, we use $\tilde{\tau} := \tau_\mathrm{a}/\gamma \equiv \beta\tau d^2$, where d is the typical dimension of a particle and the dimensionless persistence time τ has been introduced in Refs.[11,23].

Since the dimensionless variable $\mathcal{D}_\mathrm{a} := D_\mathrm{a}/D_\mathrm{t}$ implicitly depends on the persistence time, it constitutes the most general measure for the activity. At vanishing propulsion velocity $v_0 = 0$, we have $\mathcal{D}_\mathrm{a} = 0$ and Eq. (1) reduces to the equation of motion of a passive (Brownian) particle with finite reorientation time τ_a. In the limit $\tau_\mathrm{a} \to 0$ we also recover a passive (non-Brownian) system, as the velocity correlation (4) reduces to a white noise. In order to connect to an active Brownian system[11,23], it is convenient to consider instead of $\mathcal{D}_\mathrm{a}$ a dimensionless velocity $Pe = v_0 d/D_\mathrm{t}$, i.e., the Peclét number. In the literature, some other definitions of a Peclét number are used, which we will not consider here.

III. EFFECTIVE EQUILIBRIUM APPROACH

The most important step towards a theoretical study of the OUPs model is to derive from the non-Markovian stochastic process (1) an equation of motion for the N-particle probability distribution $f_N(\mathbf{r}^N, t)$. In this section, we will discuss the differences between the multidimensional generalizations of the UCNA[13,14] and the Fox[15,16] approaches to effective equilibrium and expound the surprising similarities between these two approximations in the (current-free) steady state.

As a central quantity emerging in both cases, we define the $\mathfrak{d}N \times \mathfrak{d}N$ friction tensor $\Gamma_{[N]}$ with the components

$$\Gamma_{ij}(\mathbf{r}^N) = \mathbf{1}\delta_{ij} - \tilde{\tau}\boldsymbol{\nabla}_i\mathbf{F}_j$$
$$= \delta_{ij}\Gamma_{ii}(\mathbf{r}^N) + (1 - \delta_{ij})\tilde{\tau}\boldsymbol{\nabla}_i\boldsymbol{\nabla}_j u(\mathbf{r}_i, \mathbf{r}_j) \qquad (5)$$

resulting in the Hessian of $\mathcal{U}$ and the diagonal $\mathfrak{d} \times \mathfrak{d}$ block

$$\Gamma_{ii}(\mathbf{r}^N) := \mathbf{1} + \tilde{\tau}\boldsymbol{\nabla}_i\boldsymbol{\nabla}_i\left(V_\mathrm{ext}(\mathbf{r}_i) + \sum_{k \neq i}^N u(\mathbf{r}_i, \mathbf{r}_k) \right) \qquad (6)$$

not to be confused with $\Gamma_{[1]}(\mathbf{r}_1)$ for $N = 1$ particle. In the following, we briefly denote by Γ_{ij}^{-1} the ijth block component of the inverse tensor $\Gamma_{[N]}^{-1}$.

The UCNA[18,19] amounts to explicitly inserting the OUPs (3) into the overdamped limit of the time derivative of (1), resulting in the modified Langevin equation $\dot{\mathbf{r}}_i(t) = \Gamma_{ij}^{-1}(\mathbf{r}^N)\left(\gamma^{-1}\mathbf{F}_j(\mathbf{r}^N) + \boldsymbol{\xi}_j(t) + \boldsymbol{\eta}_j(t)\right)$. It is now

straight-forward to obtain for this (approximate) Markovian system driven by white-noise the Smoluchowski equation $\partial f_N(\mathbf{r}^N, t)/\partial t = -\sum_{i=1}^{N} \boldsymbol{\nabla}_i \cdot \mathbf{J}_i(\mathbf{r}^N, t)$ with the probability current (the superscript (u) denotes that the UCNA has been used)

$$\mathbf{J}_i^{(\mathrm{u})} = \sum_k D_\mathrm{t} \Gamma_{ik}^{-1} \cdot \left(\beta \mathbf{F}_k f_N - (1+\mathcal{D}_\mathrm{a}) \sum_j \boldsymbol{\nabla}_j \cdot \left(\Gamma_{jk}^{-1} f_N \right) \right). \tag{7}$$

Note that the UCNA remains valid as long as the friction tensor (5) is positive definite.

The Fox approximation scheme applied to (1), on the other hand, only makes use of the correlator (4) of the OUPs, which, in turn, may also be interpreted as the correlator of $\boldsymbol{v}_i(t) \simeq v_0 \mathbf{p}_i(t)$ corresponding to a coarse-grained equation of motion representing active Brownian particles with a constant velocity v_0 in the direction of their instantaneous orientation $\mathbf{p}_i$ that is subject to Brownian rotational diffusion[11]. This method directly yields the approximate Smoluchowski equation (superscript (f))[24,27]

$$\mathbf{J}_i^{(\mathrm{f})} = D_\mathrm{t} \left(\beta \mathbf{F}_i f_N - \boldsymbol{\nabla}_i f_N - \mathcal{D}_\mathrm{a} \sum_j \boldsymbol{\nabla}_j \cdot (\Gamma_{ji}^{-1} f_N) \right), \tag{8}$$

where the regime of validity is the same as for UCNA. The major difference between Eq. (7) and (8) only affects the effective description of the dynamics as a result of the additional factor Γ_{ik}^{-1} arising on the level of the Langevin equation within the UCNA. Note that in the original generalization of the Fox result[11] the tensor from Eq. (5) was obtained as $\Gamma_{ij} \approx \delta_{ij}(1 - \tilde{\tau} \boldsymbol{\nabla}_i \cdot \mathbf{F}_i)$, which we will later identify as the *(diagonal) Laplacian approximation*. It will turn out that this (or another) approximation is necessary to obtain physical expressions for the effective interaction potentials.

In contrast to the dynamical problem, the (current-free) steady-state conditions

$$\beta \mathbf{F}_i P_N - \sum_j \boldsymbol{\nabla}_j \cdot (\mathcal{D}_{ji} P_N) = 0 \tag{9}$$

for the stationary distribution $P_N(\mathbf{r}^N)$ can be cast in a coherent form, defining the effective diffusion tensor $\mathbf{D}_{[N]}(\mathbf{r}^N) = D_\mathrm{t} \mathcal{D}_{[N]}(\mathbf{r}^N)$, such that only the components

$$\mathcal{D}_{ij}^{(\mathrm{u})}(\mathbf{r}^N) := (1 + \mathcal{D}_\mathrm{a}) \, \Gamma_{ij}^{-1}(\mathbf{r}^N), \tag{10}$$

$$\mathcal{D}_{ij}^{(\mathrm{f})}(\mathbf{r}^N) := \mathbf{1}\delta_{ij} + \mathcal{D}_\mathrm{a} \, \Gamma_{ij}^{-1}(\mathbf{r}^N). \tag{11}$$

differ between the UCNA (u) and Fox (f) results.

Multiplying Eq. (9) with $\mathcal{D}_{ik}^{-1}$ and summing over repeated indices, the steady-state condition takes the more instructive form[19]

$$0 = \sum_i \mathcal{D}_{ik}^{-1} \cdot \beta \mathbf{F}_i P_N - \boldsymbol{\nabla}_k P_N - P_N \boldsymbol{\nabla}_k \ln(\det \mathcal{D}_{[N]})$$

$$=: \beta \mathbf{F}_k^\mathrm{eff} P_N - \boldsymbol{\nabla}_k P_N \tag{12}$$

introducing the effective force $\mathbf{F}_k^\mathrm{eff}(\mathbf{r}^N)$. It has been argued[19] that the contribution of the off-diagonal elements to $\mathcal{D}_{[N]}$ becomes increasingly irrelevant for high particle numbers N, which means that we may set $\mathcal{D}_{ij} \to \delta_{ij} \mathcal{D}_{ij}$. Assuming this diagonal form, the determinant in Eq. (12) needs to be replaced according to $\det \mathcal{D}_{[N]} \to \det \mathcal{D}_{kk}$ as we find $\sum_{ij} \mathcal{D}_{jk}^{-1} \cdot \boldsymbol{\nabla}_j \cdot \mathcal{D}_{ij} \equiv \mathcal{D}_{kk}^{-1} \cdot \boldsymbol{\nabla}_k \cdot \mathcal{D}_{kk} = \boldsymbol{\nabla}_k \ln(\det \mathcal{D}_{kk})$ before rewriting the expression in the last step.

Putting aside the dynamical behavior, the only difference between the UCNA or Fox approximation is manifest in the definitions, (10) and (11), of $\mathcal{D}_{[N]}$. Using UCNA the active diffusivity D_a only appears as part of a prefactor in (10), so that the friction matrix $\Gamma_{[N]}$, representing a correction due to activity, contributes to the steady-state result even in the case $D_\mathrm{a} = 0$, that is when $v_0 = 0$ and $\tau_\mathrm{a} \neq 0$. Hence, the logical parameter suggested by the UCNA to tune the activity is τ_a, with the passive system ($\mathcal{D}_{ij} = \mathbf{1}\delta_{ij}$) restored only in the limit $\tau_\mathrm{a} \to 0$. Note that in this way the connection to the experimental situation with Brownian particles is lost, where it seems more natural to tune v_0 at constant τ_a. The derivation of the Fox result (11), on the other hand, aims to approximately represent active Brownian particles[11]. This reflects that we recover the (same) passive system for both $v_0 = 0$ and $\tau_\mathrm{a} = 0$.

Ignoring the contribution of the Brownian diffusion for $D_\mathrm{a} \gg D_\mathrm{t}$, the UCNA and Fox approximations practically describe the equivalent effective steady states. The major advantage of this approximation, or the UCNA result in general, is that the inverse $\mathcal{D}_{[N]}^{-1} \propto \Gamma_{[N]}$ is pairwise additive, even if $D_\mathrm{t} \neq 0$. Then the effective many-body potential $\mathcal{H}_N$ defined as $\mathbf{F}_k^\mathrm{eff}(\mathbf{r}^N) = -\boldsymbol{\nabla}_k \mathcal{H}_N(\mathbf{r}^N)$ can be written in a closed form[18,19], admitting the explicit solution $P_N(\mathbf{r}^N) \propto \exp(-\beta \mathcal{H}_N(\mathbf{r}^N))$ of Eq. (9). Due to the more nested form of Eq. (11) the Fox approximation does in general not admit an analytic result. As $\mathcal{H}_N$ is not pairwise additive in either approach, some further approximations will become necessary to construct a predictive theory, which we discuss in the following sections.

IV. EFFECTIVE-POTENTIAL APPROXIMATION (EPA)

Regarding the possible applications using standard methods of equilibrium liquid-state theory a desirable strategy is to approximate $\mathbf{F}_k^\mathrm{eff}$ in Eq. (12) in terms of pair potentials. This approach has been used to describe the phase behavior of active Brownian particles (ABPs) approximated as particles propelled by a set of coupled OUPs have been discussed in detail[11,23] for passive soft-repulsive and Lennard-Jones interactions in three dimensions. However, it can be criticized that (IV.i) a system which obeys detailed balance is used to represent the interactions in an active system, (IV.ii) the validity criteria of underlying theory might be violated so that further approximations are required and (IV.iii) higher-order particle interactions are neglected, which are be-

lieved to be important for the phase separation in an active system. In the following, we define the effective pair interaction and motivate different approximations, which we compare to computer simulations of two active Brownian particles and two particles propelled by OUPs. It is our objective to comment on the aforementioned points and illustrate the qualitative differences between the Fox and UCNA. For the sake of simplicity, we will restrict the presentation of technical aspects to the UCNA results.

To identify an effective pair potential $u^{\text{eff}}(r)$, we consider $N = 2$ interacting particles, i.e., the low-density limit of Eq. (12). Ignoring the external forces for now by setting $V_{\text{ext}}(\mathbf{r}) \equiv 0$, it is easy to verify that

$$\boldsymbol{\nabla}_1 \beta u^{\text{eff}}(r) = \left(\mathcal{D}_{11}^{-1} - \mathcal{D}_{21}^{-1}\right) \cdot \boldsymbol{\nabla}_1 \beta u(r) + \boldsymbol{\nabla}_1 \ln(\det \mathcal{D}_{[2]}) \tag{13}$$

and analog equation for $\boldsymbol{\nabla}_2 u^{\text{eff}}(r)$, where we used $\boldsymbol{\nabla}_2 u(r) = -\boldsymbol{\nabla}_1 u(r)$ and $r = |\mathbf{r}_1 - \mathbf{r}_2|$. Keeping in mind that we seek to employ this effective potential to approximately represent the interaction of many particles, it appears undesirable that an equal statistical weight is put to both the diagonal $\mathcal{D}_{11}^{-1}$ and the off-diagonal components $\mathcal{D}_{21}^{-1}$ of the diffusion tensor. As an alternative we propose the effective potential

$$\boldsymbol{\nabla}_k \beta u_{\text{diag}}^{\text{eff}}(r) = \mathcal{D}_{kk}^{-1} \cdot \boldsymbol{\nabla}_k \beta u(r) + \boldsymbol{\nabla}_k \ln(\det \mathcal{D}_{kk}), \tag{14}$$

with $k \in \{1, 2\}$, obtained for a diagonal form of $\mathcal{D}_{[2]}$ with $V_{\text{ext}}(\mathbf{r}) \equiv 0$. For completeness we find in the one-particle limit a quite similar formula

$$\boldsymbol{\nabla} \beta V_{\text{ext}}^{\text{eff}}(\mathbf{r}) = \mathcal{D}_{[1]}^{-1} \cdot \boldsymbol{\nabla} \beta V_{\text{ext}}(\mathbf{r}) + \boldsymbol{\nabla} \ln(\det \mathcal{D}_{[1]}) \tag{15}$$

for the effective external field $V_{\text{ext}}^{\text{eff}}(\mathbf{r})$, as for $N = 1$ we have $\mathcal{D}_{[1]} = \mathcal{D}_{11}$. A quite different expression for an effective external potential has been derived for active Brownian particles[23,28].

Integration of the above equalities yields the desired formulas for the effective potentials depending only on the bare potential $u(r)$ (or $V_{\text{ext}}(\mathbf{r})$) and the activity parameters D_{a} and τ_{a}[11]. Alternatively, we could have directly defined[19,21] $V_{\text{ext}}^{\text{eff}}(\mathbf{r}) := \mathcal{H}_1(\mathbf{r})$ and $u^{\text{eff}}(r) := \mathcal{H}_2(\mathbf{r}_1, \mathbf{r}_2)$ from the many-body potential $\mathcal{H}_N(\mathbf{r})$ identified in the solution of (9), which is, however, inconvenient when the Fox approach is used. Assuming a bare potential $u(r)$ obeying $\lim_{r \to \infty} u(r) = 0$, the integrated form of (13) reads

$$\beta u^{\text{eff}}(r) = \beta \frac{u(r) + \tilde{\tau}(\partial_r u(r))^2}{1 + \mathcal{D}_{\text{a}}}$$
$$- \ln\left(\left|1 + 2\tilde{\tau}\frac{\partial_r u(r)}{r}\right|^{(\mathfrak{d}-1)}\left|1 + 2\tilde{\tau}\partial_r^2 u(r)\right|\right), \tag{16}$$

where $\partial_r = \partial/\partial r$. Assuming a diagonal diffusion tensor, we obtain $u_{\text{diag}}^{\text{eff}}(r)$ from Eq. (16) by rescaling all terms proportional to $\tilde{\tau}$ with a factor $1/2$. Note that in (16) we could equally introduce an effective energy scale $\beta_{\text{eff}} =$

$\beta/(1 + \mathcal{D}_{\text{a}})$ to absorb the factor $(1 + \mathcal{D}_{\text{a}})$[19,21]. We refrain to do so as this interpretation would not be consistent with the the way $\mathcal{D}_{\text{a}}$ enters within the Fox approach.

Studying Eq. (13) more carefully, we notice that the effective potential does not always behave in a physical way. This is because, in violation of the validity condition of both the Fox and the UCNA, the diffusion tensor $\mathcal{D}_{[2]}$ is not positive definite for a large number of relevant potentials. Inspection of (16) shows that we may expect a divergence of the term $\ln(\det \mathcal{D}_{[2]})$ in the effective potential whenever one of the Eigenvalues $E_n[u(r)] := 1 + 2\tilde{\tau}r^{n-2}\partial_r^n u(r)$ for $n \in \{1, 2\}$ vanishes. We further identify $\boldsymbol{\nabla}_1 u(r)$ in the first term of (13) as the Eigenvector of $\mathcal{D}_{11}^{-1} - \mathcal{D}_{21}^{-1}$ corresponding to the Eigenvalue E_2.

Given a positive and convex bare potential $u(r) > 0$, the Eigenvalue E_2 is strictly positive, which means that the effective attraction solely arises from the last term in (13). However, as we have $\partial_r u(r) < 0$ in this case, the Eigenvalue E_1 will vanish at a certain value of r and we require a further approximation to remedy the unphysical behavior of $u^{\text{eff}}(r)$ in $\mathfrak{d} > 1$ dimensions. At a highly non-convex or negative region of the bare potential, the same problem occurs for E_2. Interestingly, if we only require knowledge of an effective potential on a finite interval where the eigenvalues are positive, its overall unphysical behavior is irrelevant[24]. Also, note that there is a broader range of admissible bare potentials when the diagonal form of the effective potential is used. In the following, we propose different ways to rid the effective potential (13) of possible artifacts due to a vanishing determinant. Correcting the other term is neither necessary nor has any noticeable effect. The presented methods can be similarly used to approximate the Fox results, in a way that the equivalent behavior in the high-$\mathcal{D}_{\text{a}}$ limit is retained.

Let us first assume that $u(r) > 0$ is convex. Then a sufficient criterion for the matrix $\mathcal{D}_{[2]}$ to have strictly positive Eigenvalues would require the operator $\boldsymbol{\nabla}_i \boldsymbol{\nabla}_j$ to be elliptic. Due to the presence of the term $\mathbf{1}\delta_{ij}$, some other potentials are allowed that are only slightly negative and slightly non-convex. A very general approximation would thus be to redefine Eq. (5) by an elliptical operator, the simplest example of which is the Laplacian $\boldsymbol{\Delta} = \boldsymbol{\nabla} \cdot \boldsymbol{\nabla}$. Upon representing $\boldsymbol{\nabla}_i \boldsymbol{\nabla}_j \to \mathbf{1}\boldsymbol{\nabla}_i \cdot \boldsymbol{\nabla}_j$ the effective potential becomes

$$\beta u_{\boldsymbol{\Delta}}^{\text{eff}}(r) = \beta \frac{u(r) + \tilde{\tau}(\partial_r u(r))^2 - 2(\mathfrak{d}-1)\int_r^\infty \mathrm{d}s\, \tilde{\tau}\frac{(\partial_s u(s))^2}{s}}{1 + \mathcal{D}_{\text{a}}}$$
$$- \ln\left(1 + 2\tilde{\tau}\partial_r^2 u(r) + 2(\mathfrak{d}-1)\tilde{\tau}\frac{\partial_r u(r)}{r}\right) \tag{17}$$

where the additional term compared to (16) cannot be integrated in general. This *Laplacian approximation* based on the Fox approach has already been employed in the diagonal form in explicit calculations[11,23]. In $\mathfrak{d} = 1$ dimensions both differential operators reduce to the second derivative and Eq. (17) is equal to Eq. (16), which pro-
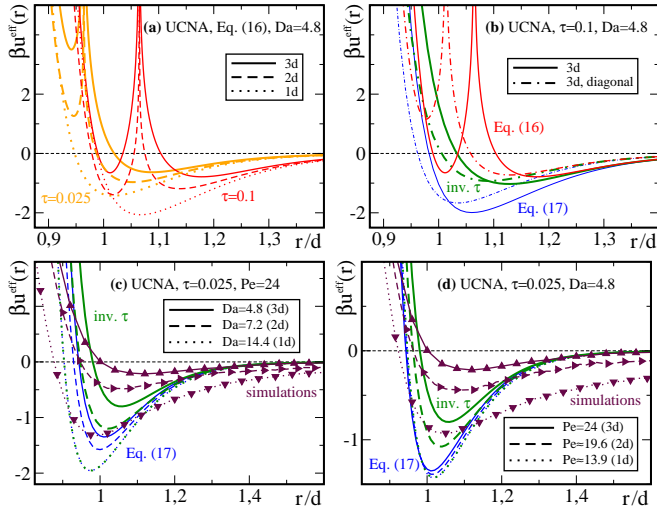
FIG. 1. Effective potentials from the UCNA in $\mathfrak{d}=1$ (dotted lines), $\mathfrak{d}=2$ (dashed lines) and $\mathfrak{d}=3$ (solid lines) dimensions. **(a)** Full result, Eq. (16), for the active diffusivity $\mathcal{D}_\mathrm{a}=4.8$ and persistence times $\tau=0.025$ (thick, brighter lines) or $\tau=0.1$ (thin, darker lines). **(b)** Comparison to the diagonal form (dot-dashed lines) in $\mathfrak{d}=3$ for $\mathcal{D}_\mathrm{a}=4.8$ and $\tau=0.1$. The thick lines correspond to the inverse-τ approximation (see text) and the thin lines to Eq. (17) in the Laplacian approximation. **(c)** Approximate results compared to simulations of active OUPS (lines with triangles) at $\tau=0.025$ and a constant Peclét number $Pe=\sqrt{\mathfrak{d}\mathcal{D}_\mathrm{a}/\tau}=24$, as $\mathcal{D}_\mathrm{a}$ increases with decreasing $\mathfrak{d}$. **(d)** Approximate results compared to simulations of active OUPS at $\mathcal{D}_\mathrm{a}=4.8$ and $\tau=0.025$.

FIG. 2. Effective potentials with and without thermal noise. **(a)** As Fig. 1d for $\mathfrak{d}=1$ including the Fox approximation (dashed lines). Ignoring thermal noise, UCNA and Fox are equivalent (dot-dashed line) and the modified simulation result of active OUPs is labeled with dots. **(b)** Simulations of ABPs (empty symbols) and OUPs for $\tau=0.05$ and $\mathcal{D}_\mathrm{a}=2.4$ in $\mathfrak{d}=3$ dimensions.

vides a good account of active particles interacting with a soft-repulsive potential[21].

An alternative approach is to empirically rectify the explicit formula for $u^\mathrm{eff}(r)$ in Eq. (16). Most intuitively, one may expand the determinant in (16) up to the first order in τ. In fact, this *small-τ approximation* turns out to be quite similar to the Laplacian approximation (17) (and completely equivalent in one dimension), but we do not recover the additional term involving the integral. Performing the small-τ approximation of the full effective potential appears too crude, as an expansion of the logarithm does not converge for $\tilde\tau\sum_{l>1}\boldsymbol{\Delta}u(\mathbf{r},\mathbf{r}_l)>2$. The resulting effective potential will thus become totally uncontrolled for short separations of highly-repulsive particles.

An approximation that may be applied also to highly non-convex potentials is the *inverse-τ approximation*, an empirical strategy maintaining the leading order in τ while not disregarding higher-order terms. This can be achieved by replacing an Eigenvalue E_n with $1/(2-E_n)$ whenever it becomes smaller than one. The major advantage of this approximation is that it yields quite similar results to the full potential whenever the validity condition is only slightly violated.
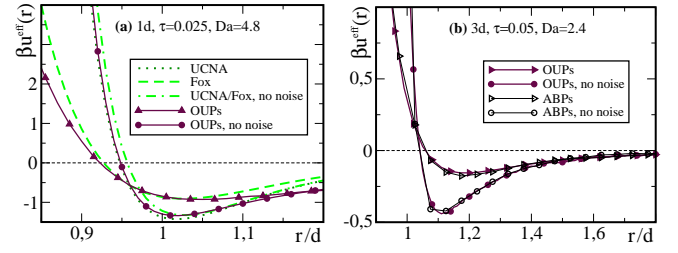
V. COMPUTER SIMULATIONS AND THEORY OF TWO ACTIVE PARTICLES

In Sec. IV we introduced different strategies to define a suitable effective interaction potential in the effective-equilibrium approximation for the colored-noise model. Now we illustrate under which conditions an approximate treatment becomes necessary and compare the theoretical results to computer simulations. The easiest way to determine an effective potential numerically is to set up a two-particle simulation, measure the radial distribution function $g(r)$ and calculate $\beta u^\mathrm{eff}_\mathrm{sim}(r)=-\ln g(r)$. By doing so, we make the same approximation (IV.iii) as in theory to ignore the many-particle character of the interaction. However, the simulations for ABPs and OUPs, detailed in appendix A, take into account the orientation dependence and the non-Markovian character of the dynamics, respectively.

We first discuss some general observations in the UCNA for a soft-repulsive system with the bare potential $\beta u(r)=(r/d)^{-12}$. The behavior of the Fox results is qualitatively similar. As expected, the full expression for the effective potential in Eq. (16) is impractical as it diverges at a certain distance r_div, determined by the condition $r_\mathrm{div}=(24\tau)^{-1/12}d$, which is when the first Eigenvalue E_1 within the logarithm vanishes, whereas E_2 is always positive. As suggested by Fig. 1a, this behavior is most problematic at larger values of τ, where $u^\mathrm{eff}(r_\mathrm{div})$ should be rather negative, as it is the case in $\mathfrak{d}=1$ dimensions. We further see in Fig. 1a that this effect becomes more severe with increasing dimension. Both the inverse-τ and Laplacian approximations successfully cure this unphysical divergence, which we see in Fig. 1b. As employing the diagonal form $u^\mathrm{eff}_\mathrm{diag}(r)$ of the effective potential simply amounts to a rescaling of τ, we observe in Fig. 1b that it results in a smaller effective diameter of the repulsive part but a flatter potential well. Accordingly, r_div becomes smaller.

Since the definition of the active diffusivity $\mathcal{D}_\mathrm{a}$ depends on the dimension $\mathfrak{d}$, not all parameters τ, $\mathcal{D}_\mathrm{a}$ and Pe can be kept constant upon varying the dimension-
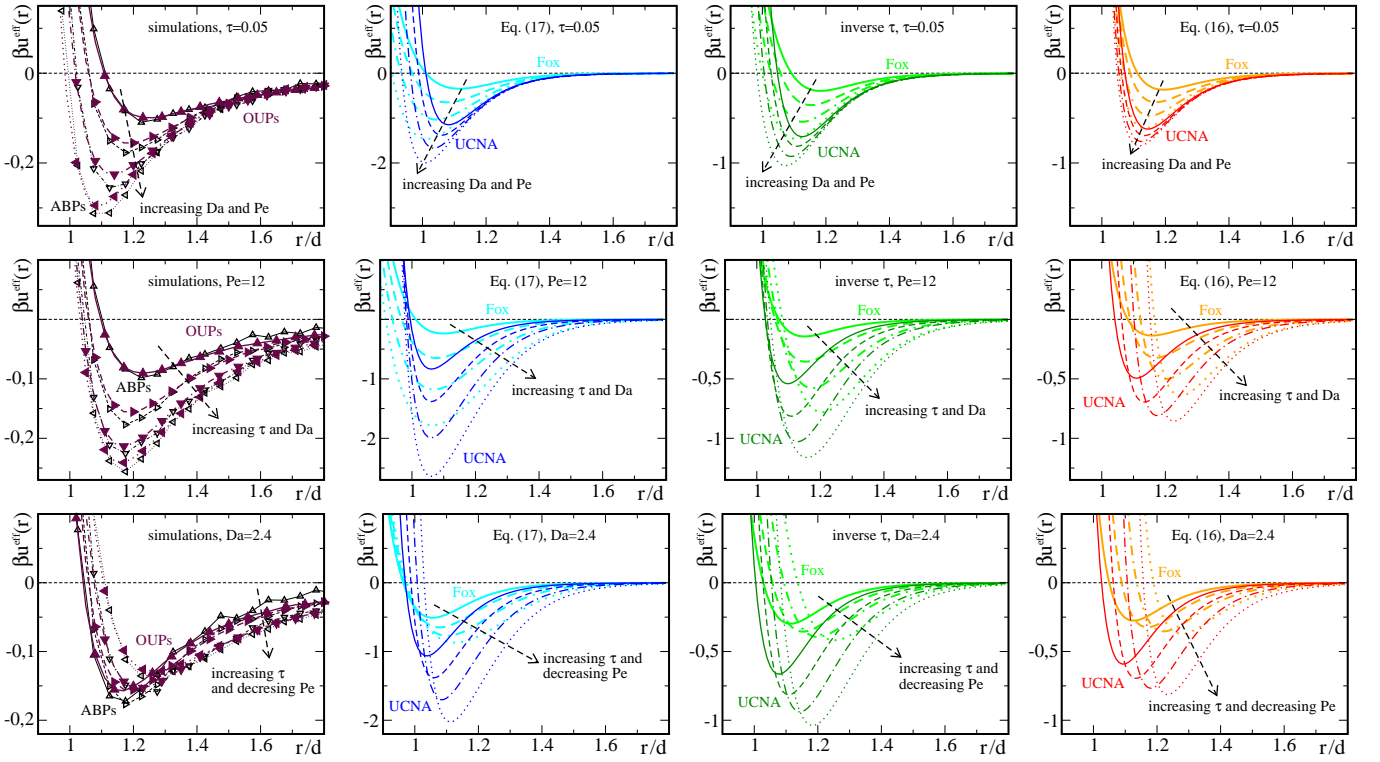
FIG. 3. Effective potentials in three dimensions for soft-repulsive spheres, $\beta u(r) = (r/d)^{-12}$, obtained within the Fox (thick, brighter lines) and UCNA (thin, darker lines) and by numerical simulations of ABPs (empty triangles) and active OUPs (filled triangles). Columns from left to right: simulations; Laplacian approximation, Eq. (17); inverse-τ approximation; Eq. (16) only for $r > r_{\text{div}}$ (see Fig. 1a for the full result). Rows from top to bottom: increasing the active diffusivity $\mathcal{D}_a$ at constant persistence time $\tau = 0.05$; increasing $\mathcal{D}_a$ and τ at constant Peclét number $Pe = \sqrt{3\mathcal{D}_a/\tau} = 12$; increasing τ at constant $\mathcal{D}_a = 2.4$. The sequence from the solid to the dotted lines corresponds to increasing the respective parameter(s) $\mathcal{D}_a$ from $\mathcal{D}_a = 1.2$ or τ from $\tau = 0.025$ by a factor of two in each step (dashed lines are always for the same set of parameters).

ality. For a constant reorientation time τ and propulsion velocity Pe the effective attraction in Fig. 1c is stronger in lower spatial dimensions for both approximations considered, which coincide with the full expression in $\mathfrak{d} = 1$. This behavior appears sensible, as two particles have less possibilities to avoid each other upon collision, and is in qualitative agreement with computer simulations of active OUPs. Moreover, we understand that motility-induced phase separation is harder to observe in higher dimensions[29]. Keeping $\mathcal{D}_a$ constant instead of Pe (which then decreases with decreasing dimension) the same trend is observed in Fig. 1d for the inverse-τ approximation and computer simulations, whereas the result for the Laplacian approximation barely changes with dimensionality. Comparing the qualitative behavior in Figs. 1c and d, we recognize in all spatial dimensions that the numerical effective potential is of longer range than the theoretical predictions in any approximation. This observation confirms the criterion discussed in Ref. 14 that the UCNA is expected to become less accurate for larger separations where a typical length scale of the active motion, closely related to the effective diffusion tensor, Eq. (10), exceeds the spatial scale over which the force field varies.

As the involved approximations become cruder in higher spatial dimensions, the quantitative agreement with the simulation results in Figs. 1c and d becomes worse. For $\mathfrak{d} = 1$, a remarkable agreement between the UCNA and simulation results for the radial distribution of two particles has been reported in Ref. 18, where the Brownian noise $\boldsymbol{\xi}_i(t)$ in Eq. (1) was set to zero. Doing so also in our simulations, we observe in Fig. 2a that the effective potential deepens and its repulsive barrier becomes steeper. This curve is in excellent agreement with the theoretical result for zero noise, obtained by both the UCNA and Fox approach upon dropping the first term in Eq. (10) and Eq. (11), respectively. The full UCNA result is only slightly different in the repulsive regime. Intriguingly, we also recognize in Fig. 2a that the simulation data including the noise term are excellently represented by the Fox approach. To further study the influence of the translational noise, we also performed computer simulations of ABPs (described in appendix A) for $\mathfrak{d} = 3$. Figure 2b reveals that the numerical effective potentials for the two considered models with and without translational noise are nearly identical over the full range of separations. As for $\mathfrak{d} = 1$, the effective potential of active OUPs (and ABPs) in the absence of thermal noise has a deeper well and a larger repulsive diameter. Quantita-

tively, this difference is much more pronounced in three dimensions.

In the following, we restrict ourselves to a three-dimensional system with translational Brownian noise. As in Fig. 2b, our simulations of ABPs and OUPs, shown in the first column of Fig. 3, are in nice agreement for all sets of parameters. This is quite surprising since on the many-particle level ABPs and OUPs have different steady states[30]. On the basis of our data for the simplistic two-body system we could rather conclude that OUPs including translational Brownian noise should be an excellent model for ABPs at moderate activity[11,23]. We see in Fig. 3 that the depth of the attractive well of all theoretical versions of the effective potential is significantly overestimated when compared to the simulations of both ABPs and OUPs. Roughly, this deviation is by a factor of 10 for the Laplacian approximation (second column) and (only) 5 for the inverse-τ approximation (third column). To facilitate a qualitative comparison we chose the y axes accordingly. Moreover, the inverse-τ approximation appears to provide the best guess of the point at which the effective potential changes its sign. The chosen approximations exhibit a similar behavior as the full theoretical results of Eq. (16) in the physical region for $r > r_{\mathrm{div}}$, shown in the last column.

For the considered soft-repulsive bare potential, we observe in Fig. 3 some notable quantitative differences between the UCNA and Fox results, even at relatively high $\mathcal{D}_{\mathrm{a}}$. The effective diameter of the repulsive part is generally smaller than in the UCNA, whereas the overall attraction is weaker in the Fox approach. This becomes most apparent in the Laplacian approximation. The first row of Fig. 3 contains the effective potentials evolving for a constant persistence time τ when the active diffusivity $\mathcal{D}_{\mathrm{a}}$ (or the Peclét number Pe) is increased. All approaches accordingly predict an increased effective attraction and the minimum of the potential is shifted to smaller separations[11,23]. The Fox results exhibit a stronger variation with $\mathcal{D}_{\mathrm{a}}$, which is more consistent with the numerical data. Similarly, the effective potentials in the second row deepen with increasing τ at constant Pe, where the location of the minimum is nearly unaffected. The most interesting behavior is observed in the third column at constant $\mathcal{D}_{\mathrm{a}}$. Again, all approaches agree that the minimum is shifted to larger separations with increasing τ, but the effective attraction predicted by the simulations is nearly constant, as simultaneously the magnitude of the self-propulsion is decreased. This observation is not consistent with the UCNA results.

Based on the presented simple comparison, our conclusion is that the best choice for the theoretical effective potential is the Fox approach in the inverse-τ approximation. Upon further increasing the activity (not shown) the quantitative discrepancy of the Laplacian approximation becomes even more pronounced. Recall from Fig. 1b that the additional assumption of the diagonal form of the effective potential also results in a slightly better (quantitative) agreement with the simulations. However,
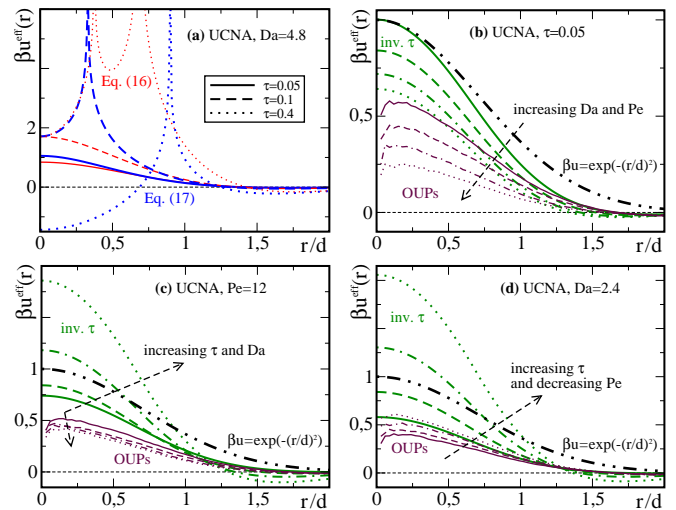

FIG. 4. Effective potentials from the UCNA in three dimensions for Gaussian-core particles, $\beta u(r) = \exp(-(r/d)^2)$. **(a)** Comparison of the divergences in the full result from Eq. (16) and the Laplacian approximation, Eq. (17), for different persistence times τ. **(b)-(d)** Inverse-τ approximation and computer simulations for active OUPs for the same parameters as in rows 1-3 of Fig. 3, respectively. The bare potential is shown as the thick dot-dot-dashed line.

we stress that it is not clear in how far the effective pair potentials can accurately describe the many-body situation, as there are no higher-order interactions present in a two-body simulation.

The most compelling argument in support of the inverse-τ approximation arises from considering non-convex bare potentials, a case in which the Laplacian approximation becomes useless above a certain value of τ. To illustrate this behavior, we discuss the more general potential $\beta u(r) = \exp(-(r/d)^2)$ of an active Gaussian-core fluid. Although this model has not received much attention, it is quite appealing from a theoretical perspective. Most prominently, this bare potential is a known exceptional case in which a simple mean-field theory is particularly accurate[31,32], which might, in a way, also hold for the effective potential of the active system.

The effective potential of an active Gaussian-core fluid is discussed in Fig. 4 within the UCNA. As the absolute value of both the curvature and slope of this model potential is bounded, the Fox results (not shown) are quite similar, even for the moderate values of $\mathcal{D}_{\mathrm{a}}$ considered here. When the persistence time τ is sufficiently small, the effective potential, Eq. (16), does not diverge and the different approximations behave in a quite similar way. Interestingly, we observe in Fig. 4a that the divergence of the Laplacian approximation sets in at an even smaller value of $\tau \geq 1/12$ than for the full potential. The latter diverges at two points, each related to one of the two Eigenvalues, if $\tau \geq 1/4$.

In Fig. 4b, c and d we discuss the only suitable form of the effective potential, i.e., in the inverse-τ approximation. Intriguingly, the predicted behavior depends on

in which way, i.e., by means of which parameter, the activity is modified. Increasing the active diffusivity (or the Peclét number) at a constant value of τ results in a less repulsive core. Upon increasing τ, however, the height of the maximum increases and an attractive well develops at larger separations. Quite counterintuitively, we observe that in this case the effective interaction becomes more repulsive than in the passive case, also for the inverse-τ approximation. Our computer simulations (also carried out for values of τ much larger than shown in Fig. 4) confirm that this is an artifact of the theory, related to the negative curvature of the bare potential. At constant τ the evolution of the theoretical results agrees qualitatively with the simulations. The simulation data are, however, not very sensitive to changes in the persistence time. At constant $\mathcal{D}_a$ the theory predicts the correct trend upon increasing τ, whereas this is not the case at constant Peclét number.

For a discussion of other non-convex bare potentials with attractive parts see, e.g., Refs. 24 and 27. However, the understanding of the behavior of particles interacting with a bare potential which has a negative curvature remains one of the most urgent open problems in our theoretical framework. Further numerical and theoretical analysis will be needed to fully clarify this issue.

VI. LOW-ACTIVITY APPROXIMATION

A second strategy to simplify the steady-state condition is to perform an expansion in the activity parameter τ. At linear order, the effective diffusion tensor $\mathcal{D}_{[N]}(\mathbf{r}^N)$ becomes pairwise additive. In this low-activity approximation, a YBG-like hierarchy can be obtained by successively integrating Eq. (9) over $N - n$ coordinates[19], which allows defining a mechanical pressure and interfacial tension[20]. Moreover, for the active system evolving according to Eqs. (1) and (4) it has been demonstrated that there exists a regime for small values of τ, where the principle of detailed balance is respected[33]. This suggests that, at leading order in this parameter, the approximations resulting in Eq. (9) are perfectly justified.

Knowing, however, that Eq. (9) contains the same information as Eq. (12), which depends logarithmically on the parameter τ, we should clarify whether (VI.i) the low-activity expansion converges, (VI.ii) it is sufficient to only consider the leading order and (VI.iii) one can obtain similar results when employing the EPA. To do so, we demonstrate how the first member ($n = 1$) of the YBG hierarchy can be rederived from Eq. (12) and discuss the consequences of approximating $\mathbf{F}_k^{\text{eff}}$ in terms of pair interactions. Again, we only discuss the UCNA, where, without any further approximation, the inverse diffusion tensor is found to be pairwise additive.

By saying we integrate over $N - 1$ coordinates we understand calculating the average $\langle \mathbf{X} \rangle := \langle\!\langle \hat{\rho} \mathbf{X} \rangle\!\rangle / \rho(\mathbf{r})$, where $\langle\!\langle \mathbf{X} \rangle\!\rangle$ denotes the full canonical ensemble average of an observable $\mathbf{X}(\mathbf{r}^N)$ and $\hat{\rho} = \sum_{i=1}^N \delta(\mathbf{r} - \mathbf{r}_i)$ is the den-

sity operator. Approximating the inverse mobility matrix in Eq. (10) as $\Gamma_{ij}^{-1}(\mathbf{r}^N) \approx \mathbf{1}\delta_{ij} - \tilde{\tau}\boldsymbol{\nabla}_i\boldsymbol{\nabla}_j\mathcal{U}$ and integrating Eq. (9) over $N - 1$ coordinates, we find the first member

$$-\rho(\mathbf{r})\langle\boldsymbol{\nabla}\beta\mathcal{U}\rangle = (1 + \mathcal{D}_a)\left(\boldsymbol{\nabla}\cdot(\mathbf{1}\rho(\mathbf{r}) - \tilde{\tau}\rho(\mathbf{r})\langle\boldsymbol{\nabla}\boldsymbol{\nabla}\mathcal{U}\rangle)\right),$$
(18)

of a YBG-like hierarchy[19,20] for the active system, where

$$\langle\mathfrak{D}\mathcal{U}\rangle = \mathfrak{D}V_{\text{ext}}(\mathbf{r}) + \int d\mathbf{r}' \frac{\rho^{(2)}(\mathbf{r},\mathbf{r}')}{\rho(\mathbf{r})}\mathfrak{D}u(\mathbf{r},\mathbf{r}') \quad (19)$$

for any differential operator $\mathfrak{D}$. In the derivation of (18) it turns out that the off-diagonal components of the mobility tensor do not contribute at first order in τ[20]. Hence, we might as well have assumed the diagonal form $\Gamma_{ij} \approx \delta_{ij}\Gamma_{ii}$ at linear order in τ beforehand.

In order to connect to the EPA, we derive a YBG-like hierarchy from Eq. (12). Assuming the diagonal form $\Gamma_{ij} \approx \delta_{ij}\Gamma_{ii}$, the integration over $N - 1$ coordinates of the first equality is carried out i n appendix B. Making use of the equilibrium version of the YBG hierarchy and expanding the $\ln(\det\Gamma_{ii}(\mathbf{r}^N))$ up to first order in τ the result is

$$0 = -\boldsymbol{\mathcal{D}}_{\text{I}}^{-1}(\mathbf{r})\rho(\mathbf{r})\langle\boldsymbol{\nabla}\beta\mathcal{U}\rangle - \boldsymbol{\nabla}\rho(\mathbf{r}) + \tilde{\tau}\rho(\mathbf{r})\langle\boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\boldsymbol{\nabla}\mathcal{U}\rangle$$

$$+ \frac{\rho(\mathbf{r})}{1 + \mathcal{D}_a}\tilde{\tau}\int d\mathbf{r}'\left(\boldsymbol{\nabla}\boldsymbol{\nabla}u(\mathbf{r},\mathbf{r}')\right)\cdot\boldsymbol{\nabla}\frac{\rho^{(2)}(\mathbf{r},\mathbf{r}')}{\rho(\mathbf{r})} \quad (20)$$

introducing the averaged inverse diffusion tensor

$$\boldsymbol{\mathcal{D}}_{\text{I}}^{-1}(\mathbf{r}) := \left\langle\mathcal{D}_{[1,\cdot]}^{-1}\right\rangle = \frac{\langle\Gamma_{[1,\cdot]}\rangle}{1 + \mathcal{D}_a} = \frac{\mathbf{1} + \tilde{\tau}\langle\boldsymbol{\nabla}\boldsymbol{\nabla}\mathcal{U}\rangle}{1 + \mathcal{D}_a}. \quad (21)$$

Multiplying Eq. (20) with $\boldsymbol{\mathcal{D}}_{\text{I}} \approx (1 + \mathcal{D}_a)(\mathbf{1} - \tilde{\tau}\langle\boldsymbol{\nabla}\boldsymbol{\nabla}\mathcal{U}\rangle)$ it is easy to verify in appendix B that at first order in τ it becomes equivalent to (18) up to a term proportional to the expression in the second line, which we consider as a higher-order contribution.

In order to derive Eq. (20) in the Fox approach, an additional approximation is required, as the inverse of $\mathcal{D}_{ij}^{(f)}$ from Eq. (11) is not proportional to Γ_{ij}. We would thus need to redefine $\boldsymbol{\mathcal{D}}_{\text{I}}$ in Eq. (21) according to Eq. (11) where $\langle\Gamma_{[1,\cdot]}\rangle^{-1}$ takes the role of Γ_{ij}^{-1}. Regarding in general the presented alternative derivation of Eq. (18), its validity appears to be in question. This is because to derive the intermediate result in Eq. (20) it is necessary to expand a logarithmic term, the Taylor series of which only has a finite radius of convergence. We further explicitly assumed the diagonal form of the mobility tensor to avoid further terms that are not present in the original result. Employing in the next step the EPA will shed more light on these issues.

VII. LOW-ACTIVITY APPROXIMATION OF EFFECTIVE POTENTIALS

Having established a connection between (9) and (12) also at linear order in τ, we now turn to the

case in which the second equality in (12) does not hold. This is when we assume $\mathbf{F}_k^{\text{eff}} \approx -\boldsymbol{\nabla}_k \mathcal{U}^{\text{eff}} = -\boldsymbol{\nabla}_k \left(V_{\text{ext}}^{\text{eff}}(\mathbf{r}_k) + \sum_{l \neq k} u^{\text{eff}}(\mathbf{r}_k, \mathbf{r}_l) \right)$ along the lines of (2) but within the EPA using the results derived in Sec. IV. As detailed in appendix C, the obvious result is that all correlation functions between more than two particles vanish in the approximate integrated version

$$0 = -\boldsymbol{\nabla}\rho(\mathbf{r}) - \rho(\mathbf{r}) \left\langle \boldsymbol{\nabla} \beta \mathcal{U}^{\text{eff}} \right\rangle \qquad (22)$$

of (12). Ignoring the interparticle interactions the approximation involving only $V_{\text{ext}}^{\text{eff}}(\mathbf{r})$ becomes exact. This situation is the same as discussed in Ref. 19.

Considering the interacting system, we multiply Eq. (22) with $\boldsymbol{\mathcal{D}}_{\text{I}}$ as done previously for (20). As detailed in appendix C, we can only approximately reproduce Eq. (18) by doing so. This reflects both the limitations of the EPA and an inconsistency between Eq. (9) and Eq. (12) when being subject to the same type of approximation, as we discuss in the following. We observe that (VII.i) the coupling between external and internal interactions is ignored by Eq. (22) (VII.ii) spurious three-body correlations appear on the left-hand-side of Eq. (18) (VII.iii) the second term on the right-hand side of Eq. (18) is recovered but involves a seemingly unjustified expansion and (VII.iv) if we do not explicitly assume a diagonal diffusion tensor, the last term in Eq. (18) changes by a factor two. As we are mainly interested in bulk systems, the first point is only briefly commented on in appendix C.

The term including the bare interaction force in Eq. (12) depends on the position of three bodies. Hence, the pairwise approximation, which amounts to setting

$$\sum_{l,j \neq k} \left(\boldsymbol{\nabla}_k \boldsymbol{\nabla}_k u(\mathbf{r}_k, \mathbf{r}_l) \right) \left(\boldsymbol{\nabla}_k u(\mathbf{r}_k, \mathbf{r}_j) \right)$$
$$\rightarrow \sum_{j \neq k} \left(\boldsymbol{\nabla}_k \boldsymbol{\nabla}_k u(\mathbf{r}_k, \mathbf{r}_j) \right) \left(\boldsymbol{\nabla}_k u(\mathbf{r}_k, \mathbf{r}_j) \right), \qquad (23)$$

should not be too crude. Moreover, we have discussed in Sec. IV that this contribution to the effective force is usually purely repulsive and thus plays only a minor role in characterizing a possible phase transition. However, we show in appendix C that definition (21) of the averaged diffusion tensor $\boldsymbol{\mathcal{D}}_{\text{I}}$ is not fully compatible with the EPA, resulting in point (VII.ii). This is in contrast to the clean derivation of Eq. (18), where Eq. (9) is recovered from Eq. (12) by multiplication with the many-body effective diffusion tensor $\mathcal{D}_{[N]}$ before integrating over $N-1$ positions.

The last term $P_N \boldsymbol{\nabla}_k \ln(\det \mathcal{D}_{kk}^{-1}(\mathbf{r}^N))$ in (12), although considered here for a diagonal diffusion tensor, constitutes a full N-body quantity. Recalling the conclusions of Sec. IV, the approximation as a pairwise quantity might be quite poor and an expansion of the logarithmic term is not justified. However, we demonstrate in appendix C that successively employing the EPA and expanding for

small $\tilde{\tau}$ according to

$$\boldsymbol{\nabla} \ln \det \left(\mathbf{1} + \tilde{\tau} \boldsymbol{\nabla}\boldsymbol{\nabla} \sum_{l>1} u(\mathbf{r}, \mathbf{r}_l) \right)$$
$$\rightarrow \boldsymbol{\nabla} \cdot \sum_{l>1} \ln \det \left(\mathbf{1} + \tilde{\tau} \boldsymbol{\nabla}\boldsymbol{\nabla} u(\mathbf{r}, \mathbf{r}_l) \right)$$
$$\rightarrow \tilde{\tau} \sum_{l>1} \boldsymbol{\nabla}\boldsymbol{\Delta} u(\mathbf{r}, \mathbf{r}_l) + \mathcal{O}(\tau^2) \qquad (24)$$

eventually results in full consistency with the respective term in Eq. (18), stated as point (VII.iii). This suggests that the expansion to first order in τ "implies" making the EPA when Eq. (12) is our starting point. Despite the aforementioned crudity of this expansion, we argue that Eq. (18) is valid, as its clean derivation from Eq. (9) does not require dealing with a logarithmic term. The last step in Eq. (24) is required to recover Eq. (18) without inducing undesired higher-order terms in τ, as, similar to point (VII.ii), the integrated version is incompatible with the chosen $\boldsymbol{\mathcal{D}}_{\text{I}}$. However, we stress that this approximation should certainly be avoided when calculating the fluid structure.

Finally, we demonstrate in appendix C that the off-diagonal elements of the diffusion tensor entering in Eq. (12) contribute to Eq. (22). Hence, the present approach would be even more inconsistent with Eq. (18) if we did not assume the diagonal form, as noted in point (VII.iv). We also note that the same problem occurs for the according generalization of Eq. (20). In principle we could define in this case an additional averaged diffusion tensor $\boldsymbol{\mathcal{D}}_{\text{I,od}}$, corresponding to the off-diagonal elements, which could counteract this inconsistency. Such a calculation would, however, not be useful when an effective pair potential is employed.

VIII. CONCLUSIONS

In this paper we studied different ways to define an effective pair interaction potential between active particles. Our numerical investigation reveals that a two-particle system of ABPs and active OUPs exhibits a quite similar behavior. These results serve as a benchmark to test the approximations involved in recent effective equilibrium approaches, which have been reviewed and compared in detail. For spatial dimensions higher than one we introduced an empirical way to rid the theoretical result of possible divergences, which also appears to yield the best agreement with the simulation data, although the effective attraction is still significantly overestimated. Regarding the quite accurate one-dimensional results and the qualitative features of the effective potentials in three dimensions, the Fox approximation is superior to the UCNA when the thermal Brownian noise cannot be neglected. In the absence of thermal noise both approximation schemes admit the same steady-state solution.

Further analysis is needed to better understand the role of the neglected many-body interactions in both the two-body simulations and the theory, which are thought to be imperative for a quantitative description of active systems[27]. The presented theoretical approach follows two major approximate steps to define the effective pair potential. First, we map the equation of motion (1) onto a system respecting detailed balance (effective equilibrium picture) and then we define pair forces from the two-particle limit. It could well be that the mapping in the first step breaks down parts of the many-body nature of the interactions in the active system, such that the effective attraction in the many-body system becomes accessible already on the level of pair interactions. As a logical next step, it seems worthwhile to study the effective potential extracted from computer simulations of a many-particle system, in order to clarify in how far the strong attraction of the effective potential needs to be seen as the result of a fortuitous cancellation of errors.

The low-activity limit in the effective equilibrium picture also results in pairwise forces. Under this assumption, we revealed some minor inconsistencies between the two equivalent steady-state conditions in Eqs. (9) and (12), although the latter contains a logarithmic term. Relatedly, it was recognized in Ref. 20 that different routes to define the active pressure only coincide at lowest order in the activity parameter τ. We suspect that further differences will occur at higher orders in τ and when employing further approximations, such as the EPA. We conclude that it should be carefully chosen for each problem which equality to consider under which approximations.

The obvious purpose of both the low-activity approximation and the EPA is to allow for an analytically tractable theory. It appears that the condition given by Eq. (12) supported by effective pair potentials is most convenient for accessing structural properties[11,21,23], whereas the low-activity expansion of Eq. (9) provides a direct way to define mechanical properties[20]. Moreover, our analysis suggests that the thermodynamic results obtained from Eq. (12) can be rescaled in order to obtain a workable definition of mechanical active pressure and surface tension. Arguably, the most simplistic scaling factor would be the diffusivity $1 + \mathcal{D}_{\mathrm{a}}$ of an ideal gas, which can be absorbed into an effective temperature[19,20,23]. A more general approach will be detailed in the second paper of this series.

Appendix A: Simulation details

We performed Brownian Dynamics simulations of a system composed of two particles of unit diameter $d=1$ interacting through soft-repulsive potential or Gaussian soft-core potential. The potential is truncated at a distance of $r = 2d$. In the simulations of active OUPs, evolving according to Eq. (1), each particle is subjected to Gaussian thermal noise and a non-Gaussian (colored) noise[24]. The integration time step is fixed to $\mathrm{d}t = 10^{-4}\tau_{\mathrm{B}}$ where $\tau_{\mathrm{B}} = d^2/D_{\mathrm{t}}$ is the time scale of translational diffusion. The total run time of simulation is $10^4\tau_{\mathrm{B}}$. For every $\mathrm{d}t$, we calculate the distance between the two particles. The pair-correlation function is obtained in a standard way from the distance distribution.

We also performed Brownian Dynamics simulation of ABPs, for which the colored-noise variable $\mathbf{v}_i(t)$ in Eq. (1) is replaced with the vector $v_0\,\mathbf{p}_i(t)$ describing a constant velocity v_0 of the self-propulsion in the direction of the instantaneous orientation. The equation $\dot{\mathbf{p}}_i(t) = \boldsymbol{\eta}_i(t) \times \mathbf{p}_i(t)$ for the time evolution for the orientation vector $\mathbf{p}_i(t)$ of each particle i is evaluated as an Ito integral, where $\boldsymbol{\eta}_i(t)$ is a white noise describing rotational diffusion. The integration time step is fixed to $\mathrm{d}t = 10^{-4}$ and the total run time is $10^4\tau_{\mathrm{B}}$.

Appendix B: Integration of the first equality in Eq. (12)

The derivation of Eq. (18) by integrating Eq. (9) over $N-1$ coordinates is quite similar to that of the YBG hierarchy in a passive system. The first member

$$0 = \boldsymbol{\nabla}\rho(\mathbf{r}) + \rho(\mathbf{r})\boldsymbol{\nabla}\beta V_{\mathrm{ext}}(\mathbf{r}) + \int \mathrm{d}\mathbf{r}'\rho^{(2)}(\mathbf{r},\mathbf{r}')\boldsymbol{\nabla}\beta u(\mathbf{r},\mathbf{r}') = \rho(\mathbf{r})\boldsymbol{\nabla}\mu \tag{B1}$$

is recovered from (18) when setting $\tau = \mathcal{D}_{\mathrm{a}} = 0$. The second equality reflects the interpretation of the term on the left-hand side as the gradient of a chemical potential μ, which is constant in equilibrium. The second member reads

$$0 = \boldsymbol{\nabla}\rho^{(2)}(\mathbf{r},\mathbf{r}'') + \rho^{(2)}(\mathbf{r},\mathbf{r}'')\boldsymbol{\nabla}\left(\beta V_{\mathrm{ext}}(\mathbf{r}) + \beta u(\mathbf{r},\mathbf{r}'')\right) + \int \mathrm{d}\mathbf{r}'\rho^{(3)}(\mathbf{r},\mathbf{r}',\mathbf{r}'')\boldsymbol{\nabla}\beta u(\mathbf{r},\mathbf{r}') = \rho^{(2)}(\mathbf{r},\mathbf{r}'')\boldsymbol{\nabla}\mu\,. \tag{B2}$$

and is related via the second equality to the first member. With the help of these exact equilibrium sum rules we will now derive Eq. (20) by integrating Eq. (12) over $N - 1$ coordinates. Our presentation follows closely the derivation of a dynamical density functional theory including a tensorial diffusivity[34], whereas we only consider the steady-state condition.

10

We start by writing the first equality in (12) as

$$0 = \sum_i \mathcal{D}_{ik}^{-1} \cdot (\beta \mathbf{F}_i P_N - \boldsymbol{\nabla}_i P_N) + \mathcal{D}_{ik}^{-1} \cdot \boldsymbol{\nabla}_i P_N - \boldsymbol{\nabla}_k P_N + P_N \boldsymbol{\nabla}_k \ln(\det \mathcal{D}_{kk}^{-1}), \tag{B3}$$

where we further used that the negative logarithm is the logarithm of the inverse argument and replaced $\det \mathcal{D}_{[N]}$ with $\det \mathcal{D}_{kk}$ by assuming the diagonal form. Integration of (B3) over $N-1$ coordinates yields (within UCNA)

$$\frac{1}{1+\mathcal{D}_a} \left(-\boldsymbol{\nabla}\rho(\mathbf{r}) - \rho(\mathbf{r})\boldsymbol{\nabla}\beta V_{\text{ext}}(\mathbf{r}) - \int d\mathbf{r}' \rho^{(2)}(\mathbf{r},\mathbf{r}')\boldsymbol{\nabla}\beta u(\mathbf{r},\mathbf{r}') \right.$$

$$- \tilde{\tau}\boldsymbol{\nabla}\boldsymbol{\nabla} V_{\text{ext}}(\mathbf{r}) \cdot \left(\boldsymbol{\nabla}\rho(\mathbf{r}) + \rho(\mathbf{r})\boldsymbol{\nabla}\beta V_{\text{ext}}(\mathbf{r}) + \int d\mathbf{r}' \rho^{(2)}(\mathbf{r},\mathbf{r}')\boldsymbol{\nabla}\beta u(\mathbf{r},\mathbf{r}') \right)$$

$$- \tilde{\tau}\int d\mathbf{r}'' \boldsymbol{\nabla}\boldsymbol{\nabla} u(\mathbf{r},\mathbf{r}'') \cdot \left(\boldsymbol{\nabla}\rho^{(2)}(\mathbf{r},\mathbf{r}'') + \rho^{(2)}(\mathbf{r},\mathbf{r}'')\boldsymbol{\nabla}(\beta V_{\text{ext}}(\mathbf{r}) + \beta u(\mathbf{r},\mathbf{r}'')) + \int d\mathbf{r}' \rho^{(3)}(\mathbf{r},\mathbf{r}',\mathbf{r}'')\boldsymbol{\nabla}\beta u(\mathbf{r},\mathbf{r}') \right)$$

$$\left. + \boldsymbol{\nabla}\rho(\mathbf{r}) + \tilde{\tau}(\boldsymbol{\nabla}\boldsymbol{\nabla} V_{\text{ext}}(\mathbf{r})) \cdot \boldsymbol{\nabla}\rho(\mathbf{r}) + \tilde{\tau}\int d\mathbf{r}' (\boldsymbol{\nabla}\boldsymbol{\nabla} u(\mathbf{r},\mathbf{r}')) \cdot \boldsymbol{\nabla}\rho^{(2)}(\mathbf{r},\mathbf{r}') \right)$$

$$- \boldsymbol{\nabla}\rho(\mathbf{r}) + N\int d\mathbf{r}_2 \dots \int d\mathbf{r}_N P_N(\mathbf{r},\mathbf{r}_2,\dots,\mathbf{r}_N)\boldsymbol{\nabla}\ln\det\left(\mathbf{1} + \tilde{\tau}\boldsymbol{\nabla}\boldsymbol{\nabla}\left(V_{\text{ext}}(\mathbf{r}) + \sum_{l>1} u(\mathbf{r},\mathbf{r}_l) \right) \right) = 0 \tag{B4}$$

As the Fox result (11) for $\mathcal{D}_{ik}^{-1}$ is no pairwise quantity, a further approximation is required to obtain a similar hierarchy. Now we eliminate the term in brackets within the third line containing the three-body correlation function with the help of (B1) and (B2) and expand the logarithm up to linear order in τ. The result is

$$- \frac{1}{1+\mathcal{D}_a}\left(\mathbf{1} + \tilde{\tau}\boldsymbol{\nabla}\boldsymbol{\nabla} V_{\text{ext}}(\mathbf{r}) + \tilde{\tau}\int d\mathbf{r}'' \frac{\rho^{(2)}(\mathbf{r},\mathbf{r}'')}{\rho(\mathbf{r})}\boldsymbol{\nabla}\boldsymbol{\nabla} u(\mathbf{r},\mathbf{r}'') \right) \cdot \left(\rho(\mathbf{r})\boldsymbol{\nabla}\beta V_{\text{ext}}(\mathbf{r}) + \int d\mathbf{r}' \rho^{(2)}(\mathbf{r},\mathbf{r}')\boldsymbol{\nabla}\beta u(\mathbf{r},\mathbf{r}') \right)$$

$$- \boldsymbol{\nabla}\rho(\mathbf{r}) + \tilde{\tau}\rho(\mathbf{r})(\boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\boldsymbol{\nabla} V_{\text{ext}}(\mathbf{r})) + \tilde{\tau}\int d\mathbf{r}' \rho^{(2)}(\mathbf{r},\mathbf{r}')\boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\boldsymbol{\nabla} u(\mathbf{r},\mathbf{r}')$$

$$+ \frac{1}{1+\mathcal{D}_a}\tilde{\tau}\int d\mathbf{r}' (\boldsymbol{\nabla}\boldsymbol{\nabla} u(\mathbf{r},\mathbf{r}'))\rho(\mathbf{r})\cdot\boldsymbol{\nabla}\frac{\rho^{(2)}(\mathbf{r},\mathbf{r}')}{\rho(\mathbf{r})} = 0 \tag{B5}$$

where we used the identity $\boldsymbol{\nabla}\boldsymbol{\Delta} = \boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\boldsymbol{\nabla}$. The term in the last line stems from replacing $\boldsymbol{\nabla}\rho^{(2)}(\mathbf{r},\mathbf{r}'')$ in the third line of Eq. (B4) with $\frac{\rho^{(2)}(\mathbf{r},\mathbf{r}'')}{\rho(\mathbf{r})}\boldsymbol{\nabla}\rho(\mathbf{r})$, which does not cancel with the expression in the fourth line of Eq. (B4).

For convenience we adopt the notational convention of the main text (19) and identify the $\mathfrak{d}\times\mathfrak{d}$ matrix in the first line of Eq. (B5) as the inverse of an ensemble-averaged diffusion tensor

$$\mathcal{D}_{\text{I}}(\mathbf{r}) := (1+\mathcal{D}_a)\left(\mathbf{1} + \tilde{\tau}\boldsymbol{\nabla}\boldsymbol{\nabla} V_{\text{ext}}(\mathbf{r}) + \tilde{\tau}\int d\mathbf{r}' \frac{\rho^{(2)}(\mathbf{r},\mathbf{r}')}{\rho(\mathbf{r})}\boldsymbol{\nabla}\boldsymbol{\nabla} u(\mathbf{r},\mathbf{r}') \right)^{-1}$$

$$= \left(\frac{\mathbf{1} + \tilde{\tau}\langle\boldsymbol{\nabla}\boldsymbol{\nabla}\mathcal{U}\rangle}{1+\mathcal{D}_a} \right)^{-1} = (1+\mathcal{D}_a)(\mathbf{1} - \tilde{\tau}\langle\boldsymbol{\nabla}\boldsymbol{\nabla}\mathcal{U}\rangle) + \mathcal{O}(\tau^2). \tag{B6}$$

Now we multiply Eq. (B5) with $\mathcal{D}_{\text{I}}$ and drop all higher-order terms $\propto \tau^2$, which yields

$$\rho(\mathbf{r})\langle\boldsymbol{\nabla}\beta\mathcal{U}\rangle = (1+\mathcal{D}_a)(-\boldsymbol{\nabla}\rho(\mathbf{r}) + \tilde{\tau}((\boldsymbol{\nabla}\rho(\mathbf{r}))\cdot\langle\boldsymbol{\nabla}\boldsymbol{\nabla}\mathcal{U}\rangle + \rho(\mathbf{r})\langle\boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\boldsymbol{\nabla}\mathcal{U}\rangle)) + \tilde{\tau}\int d\mathbf{r}' (\boldsymbol{\nabla}\boldsymbol{\nabla} u(\mathbf{r},\mathbf{r}'))\rho(\mathbf{r})\cdot\boldsymbol{\nabla}\frac{\rho^{(2)}(\mathbf{r},\mathbf{r}')}{\rho(\mathbf{r})}$$

$$= (1+\mathcal{D}_a)(-\boldsymbol{\nabla}\rho(\mathbf{r}) + \tilde{\tau}\boldsymbol{\nabla}\cdot(\rho(\mathbf{r})\langle\boldsymbol{\nabla}\boldsymbol{\nabla}\mathcal{U}\rangle)) - \mathcal{D}_a\tilde{\tau}\int d\mathbf{r}' (\boldsymbol{\nabla}\boldsymbol{\nabla} u(\mathbf{r},\mathbf{r}'))\rho(\mathbf{r})\cdot\boldsymbol{\nabla}\frac{\rho^{(2)}(\mathbf{r},\mathbf{r}')}{\rho(\mathbf{r})}. \tag{B7}$$

In the last step we made use of the identity

$$\boldsymbol{\nabla}\cdot(\rho(\mathbf{r})\langle\boldsymbol{\nabla}\boldsymbol{\nabla}\beta\mathcal{U}\rangle) = (\boldsymbol{\nabla}\rho(\mathbf{r}))\cdot\langle\boldsymbol{\nabla}\boldsymbol{\nabla}\beta\mathcal{U}\rangle + \rho(\mathbf{r})\langle\boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\boldsymbol{\nabla}\beta\mathcal{U}\rangle + \rho(\mathbf{r})\int d\mathbf{r}' (\boldsymbol{\nabla}\boldsymbol{\nabla}\beta u(\mathbf{r},\mathbf{r}'))\cdot\boldsymbol{\nabla}\frac{\rho^{(2)}(\mathbf{r},\mathbf{r}')}{\rho(\mathbf{r})} \tag{B8}$$

to recover up to the last term the first member of the YGB-like hierarchy stated in Eq. (18) of the main text. Taking into account the definition $\mathcal{D}_a \propto \tau$ of the active diffusion coefficient, we argue that the additional term is not relevant at linear order in τ. Alternatively, taking the mean-field approximation $\rho^{(2)}(\mathbf{r},\mathbf{r}') \approx \rho(\mathbf{r})\rho(\mathbf{r}')$, this term will also vanish. We thus have rederived a result obtained in a much simpler way in Ref. 19. The demonstrated equivalence of Eqs. (9) and (12) in the low-activity limit is, however, not obvious and breaks down when higher-order terms in τ are included.

Appendix C: Integration of the second equality in Eq. (12)

Assuming pairwise interaction potentials, the integration of the second equality in Eq. (12) over $N-1$ coordinates results in

$$0 = \boldsymbol{\nabla}\rho(\mathbf{r}) + \rho(\mathbf{r})\boldsymbol{\nabla}\beta V_{\text{ext}}^{\text{eff}}(\mathbf{r}) + \int d\mathbf{r}'\rho^{(2)}(\mathbf{r},\mathbf{r}')\boldsymbol{\nabla}\beta u^{\text{eff}}(\mathbf{r},\mathbf{r}')$$

$$= \boldsymbol{\nabla}\rho(\mathbf{r}) + \rho(\mathbf{r})\left(\mathcal{D}_{[1]}^{-1}(\mathbf{r})\boldsymbol{\nabla}\beta V_{\text{ext}}(r) - \boldsymbol{\nabla}\ln(\det\mathcal{D}_{[1]}(\mathbf{r}))\right)$$

$$+ \left[\int d\mathbf{r}'\rho^{(2)}(\mathbf{r},\mathbf{r}')\left(\mathcal{D}_{11}^{-1}(\mathbf{r},\mathbf{r}')\cdot\boldsymbol{\nabla}\beta u(\mathbf{r},\mathbf{r}') - \boldsymbol{\nabla}\ln(\det\mathcal{D}_{11}(\mathbf{r},\mathbf{r}'))\right)\right]_{V_{\text{ext}}=0} \tag{C1}$$

where in the second step we have inserted the effective external (15) and pair potential (14). To be consistent with appendix B we used the diagonal form of the latter. In the absence of interparticle interactions, $u(r) = 0$, it is easy to verify that both Eqs. (C1) and (B4) simplify to the same equality

$$\boldsymbol{\nabla}\rho(\mathbf{r}) + \rho(\mathbf{r})\left(\frac{(\boldsymbol{\nabla}\beta V_{\text{ext}}(\mathbf{r}))\cdot(\mathbf{1}+\tilde{\tau}\boldsymbol{\nabla}\boldsymbol{\nabla}V_{\text{ext}}(\mathbf{r}))}{1+\mathcal{D}_{\text{a}}} - \boldsymbol{\nabla}\ln\det\left(\mathbf{1}+\tilde{\tau}\boldsymbol{\nabla}\boldsymbol{\nabla}V_{\text{ext}}(\mathbf{r})\right)\right) = 0\,, \tag{C2}$$

which is a trivial consequence of the fact that the many-body potential is the sum of single-particle contributions: the friction tensor (5) is diagonal and all equations decouple. Comparing the result (C1) in the interacting case to Eq. (B5), we notice that the EPA ignores the cross terms proportional to $\tilde{\tau}\boldsymbol{\nabla}\boldsymbol{\nabla}V_{\text{ext}}(\mathbf{r})\cdot\int d\mathbf{r}'\rho^{(2)}(\mathbf{r},\mathbf{r}')\boldsymbol{\nabla}\beta u(\mathbf{r},\mathbf{r}')$ and $\tilde{\tau}\boldsymbol{\nabla}\beta V_{\text{ext}}(\mathbf{r})\cdot\int d\mathbf{r}'\rho^{(2)}(\mathbf{r},\mathbf{r}')\boldsymbol{\nabla}\boldsymbol{\nabla}u(\mathbf{r},\mathbf{r}')$, coupling the external and internal interactions on the level of pair correlations. It is, however, possible to capture these terms within a generalized effective external two-body field in the spirit of Ref. 21. This amendable difference aside, we now discuss the bulk system.

Setting $V_{\text{ext}}(\mathbf{r}) = 0$ in Eq. (C1) becomes

$$\boldsymbol{\nabla}\rho(\mathbf{r}) + \int d\mathbf{r}'\rho^{(2)}(\mathbf{r},\mathbf{r}')\left(\frac{(\boldsymbol{\nabla}\beta u(\mathbf{r},\mathbf{r}'))\cdot(\mathbf{1}+\tilde{\tau}\boldsymbol{\nabla}\boldsymbol{\nabla}u(\mathbf{r},\mathbf{r}'))}{1+\mathcal{D}_{\text{a}}} - \boldsymbol{\nabla}\ln\det\left(\mathbf{1}+\tilde{\tau}\boldsymbol{\nabla}\boldsymbol{\nabla}u(\mathbf{r},\mathbf{r}'))\right)\right) = 0 \tag{C3}$$

This result amounts to setting

$$\frac{\tilde{\tau}}{1+\mathcal{D}_{\text{a}}}\int d\mathbf{r}''\left(\boldsymbol{\nabla}\beta u(\mathbf{r},\mathbf{r}'')\right)\cdot\int d\mathbf{r}'\rho^{(3)}(\mathbf{r},\mathbf{r}',\mathbf{r}'')\boldsymbol{\nabla}\boldsymbol{\nabla}u(\mathbf{r},\mathbf{r}') \to 0\,, \tag{C4}$$

$$N\int d\mathbf{r}_2\ldots d\mathbf{r}_N P_N(\mathbf{r}^N)\boldsymbol{\nabla}\ln\det\left(\mathbf{1}+\tilde{\tau}\boldsymbol{\nabla}\boldsymbol{\nabla}\sum_{l>1}u(\mathbf{r},\mathbf{r}_l)\right) \to \int d\mathbf{r}'\rho^{(2)}(\mathbf{r},\mathbf{r}')\boldsymbol{\nabla}\ln\det\left(\mathbf{1}+\tilde{\tau}\boldsymbol{\nabla}\boldsymbol{\nabla}u(\mathbf{r},\mathbf{r}')\right) \tag{C5}$$

in (B4), which is a logical consequence of the higher-order correlations being ignored. Restricting ourselves to the leading order in τ both sides of (C5), we find that reduce to the equivalent form $\tilde{\tau}\int d\mathbf{r}'\rho^{(2)}(\mathbf{r},\mathbf{r}')\boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\boldsymbol{\nabla}u(\mathbf{r},\mathbf{r}')$. In contrast, making use of (B2), the approximation (C4) to the bare force term is equivalent to setting

$$\frac{\tilde{\tau}}{1+\mathcal{D}_{\text{a}}}\int d\mathbf{r}''\frac{\rho^{(2)}(\mathbf{r},\mathbf{r}'')}{\rho(\mathbf{r})}\boldsymbol{\nabla}\boldsymbol{\nabla}u(\mathbf{r},\mathbf{r}'')\cdot\int d\mathbf{r}'\rho^{(2)}(\mathbf{r},\mathbf{r}')\boldsymbol{\nabla}\beta u(\mathbf{r},\mathbf{r}') + \frac{\tilde{\tau}}{1+\mathcal{D}_{\text{a}}}\int d\mathbf{r}'\left(\boldsymbol{\nabla}\boldsymbol{\nabla}u(\mathbf{r},\mathbf{r}')\right)\rho(\mathbf{r})\cdot\boldsymbol{\nabla}\frac{\rho^{(2)}(\mathbf{r},\mathbf{r}')}{\rho(\mathbf{r})}$$

$$\to \frac{\tilde{\tau}}{1+\mathcal{D}_{\text{a}}}\int d\mathbf{r}'\rho^{(2)}(\mathbf{r},\mathbf{r}')(\boldsymbol{\nabla}\beta u(\mathbf{r},\mathbf{r}'))\cdot(\boldsymbol{\nabla}\boldsymbol{\nabla}u(\mathbf{r},\mathbf{r}')) \tag{C6}$$

in (B5). Thus the factorization in the first line of (B5) is not possible, such that, upon multiplying with $\mathcal{D}_{\text{I}}$ defined in Eq. (B6) there remains an additional term

$$\tilde{\tau}\int d\mathbf{r}'\rho^{(2)}(\mathbf{r},\mathbf{r}')(\boldsymbol{\nabla}\beta u(\mathbf{r},\mathbf{r}'))\cdot\left(\boldsymbol{\nabla}\boldsymbol{\nabla}u(\mathbf{r},\mathbf{r}') - \int d\mathbf{r}''\frac{\rho^{(2)}(\mathbf{r},\mathbf{r}'')}{\rho(\mathbf{r})}\boldsymbol{\nabla}\boldsymbol{\nabla}u(\mathbf{r},\mathbf{r}'')\right) \tag{C7}$$

proportional to τ on the left-hand-side of Eq. (B7). This means that the EPA introduces a three-body term to the YBG-like hierarchy (18). The reason is that the effective diffusion tensor $\mathcal{D}_{\text{I}}$ is defined independently of the approximation made in Eq. (C4).

Finally, we note that if we employ in Eq. (C1) the effective pair potential (13) that does not correspond to a diagonal diffusion tensor, we will have to Modify Eq. (C3) by setting $\tau \to 2\tau$. This reflects in general the inconsistency between the low-activity expansions of the two versions of the steady-state condition given by Eqs. (9) and (12), which is not

12

due to approximating the effective force in the second form using pair potentials. In particular, we also need to substitute $\tau \to 2\tau$ in Eq. (B7), as both sides in Eq. (C5) are equivalent at linear order in τ. However, in Eq. (C7) only the first term is to be multiplied by the factor two, as the second term arises from $\mathcal{D}_I$ and not from the effective potential.

REFERENCES

[1] J. Elgeti, R.G. Winkler and G. Gompper, Rep. Prog. Phys. **78**, 056601 (2015).

[2] V. Kantsler *et al.*, Proc. Natl. Acad. Sci. U.S.A. **110**, 1187 (2013).

[3] S. Xiao *et al.*, J. Chem. Phys. **141**, 184902 (2014).

[4] X. Yang, M.L. Manning and M.C. Marchetti, Soft Matter, **10**, 6477 (2014).

[5] R. Ni, M.A. Cohen Stuart and P.G. Bolhuis, Phys. Rev. Lett., **114**, 018302 (2015).

[6] M.E. Cates and J. Tailleur, Annu. Rev. Condens. Matter Phys. **6**, 219 (2015).

[7] M.E. Cates, D. Marenduzzo, I. Pagonabarraga, and J. Tailleur, Proc. Natl. Acad. Sci. U. S. A. **107**, 11715 (2010).

[8] Y. Fily and M.C. Marchetti, Phys. Rev. Lett. **108**, 235702 (2012).

[9] T. Speck, J. Bialk, A.M. Menzel, and H. Löwen, Phys. Rev. Lett. **112**, 218304 (2014).

[10] A. Sharma and J.M. Brader, J.Chem.Phys. **145**, 161101 (2016).

[11] T. F. F. Farage, P. Krinninger, and J. M. Brader, Phys. Rev. E **91**, 042310 (2015).

[12] S. Faetti, L. Fronzoni, P. Grigolini, and R. Mannella, J. Stat. Phys. **52**, 951 (1988).

[13] P. Jung and P. Hnggi, Phys. Rev. A **35**, 4464 (1987).

[14] P. Hänggi and P. Jung, Adv.Chem. Phys. **89**, 239 (1995).

[15] R.F. Fox, Phys. Rev. A **33**, 467 (1986).

[16] R.F. Fox, Phys. Rev. A **34**, 4525(R) (1986).

[17] J. Schwarz-Linek *et al.*, Proc. Natl. Acad. Sci. USA **109**, 4052 (2012).

[18] C. Maggi *et al.* Scientific Reports **5**, 10742 (2015).

[19] U. M. B. Marconi and C. Maggi, Soft Matter, **11**, 8768 (2015).

[20] U. M. B. Marconi, C. Maggi and S. Melchionna, Soft Matter **12**, 5727 (2016).

[21] U. M. B. Marconi, M. Paoluzzi and C. Maggi, Mol. Phys. 10.1080/00268976.2016.1155777 (2016).

[22] U. M. B. Marconi, N. Gnan, M. Paoluzzi, C. Maggi and R. Di Leonardo, Scientific Reports **6** 23297 (2015).

[23] R. Wittmann and J. M. Brader, EPL **114** 68004 (2016).

[24] A. Sharma, R. Wittmann and J. M. Brader, Phys. Rev. E, **95**, 012115 (2017).

[25] J. Tailleur and M. E. Cates, Phys. Rev. Lett. **100**, 218103 (2008).

[26] M. E. Cates and J. Tailleur, EPL **101**, 20010 (2013).

[27] M. Rein and T. Speck, Eur. Phys. J. E **39**, 84 (2016).

[28] A. Pototsky and H. Stark, EPL **98**, 50004 (2012).

[29] J. Stenhammar *et al.*, Soft Matter, **10**, 1489 (2014).

[30] A. P. Solon, M. E. Cates and J. Tailleur, Eur. Phys. J. Special Topics **224**, 1231 (2015).

[31] A. A. Louis, P. G. Bolhuis and J. P. Hansen, Phys. Rev. E **62**, 7961 (2000).

[32] A. Lang *et al.*, J. Phys.: Condend. Matter **12**, 5087 (2000).

[33] E. Fodor *et al.*, Phys. Rev. Lett. **117**, 038103 (2016).

[34] M. Rex and H. Löwen, Eur. Phys. J. E **28**,139 (2009).