UNIVERSITÄT AUGSBURG
Universitätsbibliothek

OPEN ACCESS
OPUS AUGSBURG

Effective equilibrium states in the colored-noise model for active matter I: pairwise forces in the Fox and unified colored noise approximations

René Wittmann, C. Maggi, Abhinav Sharma, A. Scacchi, J. M. Brader, U. Marini Bettolo Marconi

Angaben zur Veröffentlichung / Publication details:

Wittmann, René, C. Maggi, Abhinav Sharma, A. Scacchi, J. M. Brader, and U. Marini Bettolo Marconi. 2017. "Effective equilibrium states in the colored-noise model for active matter I: pairwise forces in the Fox and unified colored noise approximations." *Journal of Statistical Mechanics: Theory and Experiment* 2017 (11): 113207. https://doi.org/10.1088/1742-5468/aa8c1f.

Nutzungsbedingungen / Terms of use:

licgercopyright

Dieses Dokument wird unter folgenden Bedingungen zur Verfügung gestellt: / This document is made available under these conditions:
Deutsches Urheberrecht
Weitere Informationen finden Sie unter: / For more information see:
https://www.uni-augsburg.de/de/organisation/bibliothek/publizieren-zitieren-archivieren/publiz/

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], C Maggi[2], A Sharma[1,3], A Scacchi[1], J M Brader[1] and U Marini Bettolo Marconi[4]

[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] Leibniz-Institut für Polymerforschung Dresden, D-01069 Dresden, Germany
[4] Scuola di Scienze e Tecnologie, Università di Camerino, Via Madonna delle Carceri, 62032, Camerino, INFN Perugia, Italy
E-mail: rene.wittmann@unifr.ch

Abstract. 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 the approximations necessary to define effective interaction potentials in the low-density limit and study the conditions for which these represent the behavior observed in two-body simulations for the OUPs model and active Brownian particles. The demonstrated limitations of the theory for potentials with a negative slope or curvature can be qualitatively corrected by a new empirical modification. In general, we find that in the presence of translational white noise the Fox approach is more accurate. Finally, we examine an alternative way to define a force-balance condition in the limit of small activity.

Keywords: active matter, coarse-graining, stationary states, Brownian motion

Contents

1. Introduction **2**

2. Theory **4**

2.1. Colored-noise model 4

2.2. Effective equilibrium approach 5

2.3. Two versions of the steady-state condition 6

3. Effective-potential approximation (EPA) **7**

3.1. Calculation of effective potentials 8

3.2. Limitations and possible corrections 8

3.3. Comparison to computer simulations of two active particles 10

3.3.1. The role of approximations, dimensionality and thermal noise. . . . 10

3.3.2. Soft-repulsive Brownian system in three dimensions. 12

3.3.3. Difficulties for non-convex potentials. 13

4. Low-activity approximation **15**

4.1. Alternative derivation of a local force-balance condition 16

4.2. Low-activity approximation of effective potentials 17

5. Conclusions **18**

Acknowledgments **19**

Appendix A. Simulation details **19**

Appendix B. Effective many-body force in the Fox approximation **20**

Appendix C. Integration of the first equality in equation (12) **22**

Appendix D. Integration of the second equality in equation (12) **23**

References **25**

1. Introduction

Active Brownian particles (ABPs) provide a simple, minimal model system to study the collective behavior of active matter. The 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 [6–10], which do not contain all information about the interparticle correlations. Some progress has recently been made in the linear response regime, which allows to decouple the equations of motions of the one-body density and polarization vector [11].

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 [12]). 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 various approximation methods [13–19] which enable the OUP model to be represented using an effective Markovian, and therefore tractable, dynamics. 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, and (ii) the Fox approximation [15, 16], for which an approximate Fokker–Planck equation is developed, have recently been adopted in the context of developing simple theoretical tools to describe active particles [12, 20–26].

When applied to active matter, both the UCNA and Fox approximations are referred to as 'effective equilibrium' approaches. The Markovian character of the dynamics implies that they obey a Fokker–Planck equation from which an effective probability distribution can in principle be obtained. 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. [27]) who observed that the phase separation induced by activity in systems of repulsive ABPs closely resembles 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 [20–24] and Fox [12, 25, 26] approximations, to put the notion of an effective equilibrium description on a firmer footing.

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 the activity-induced modifications of the microstructure and motility-induced phase separation [12]. This analysis clarifies both the nature of the approximations involved and suggests ways in which the description can be improved.

The paper is laid out as follows: in section 2 we first specify the model under consideration and describe the UCNA and Fox approaches to obtaining an effective equilibrium picture highlighting similarities and differences between them. In section 3 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 using a standard soft-repulsive and a non-convex (Gaussian core) potential. In section 4 we consider an alternative approach to obtain pairwise forces, i.e. the low-activity limit, and make contact to the EPA. Finally, we conclude in section 5.

2. Theory

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 [12] or not [20]. We will also clarify some notational issues.

2.1. Colored-noise model

We consider the coupled stochastic (Langevin) differential equations

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

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

$$\mathcal{U}(\mathbf{r}^N) = \left(\nu(\mathbf{r}_i) + \frac{1}{2}\sum_{k=i}^{N} u(\mathbf{r}_i,\mathbf{r}_k)\right), \tag{2}$$

consisting of the one-body external fields $\nu(\mathbf{r}_i)$ 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 '$\cdot$'. Hereafter, we shortly refer to the variable $\boldsymbol{\xi}_i(t)$ as (thermal) noise. The OUPs $\boldsymbol{v}_i(t)$ defined by

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

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

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

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 [28] or ABPs [12, 29]. 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 [23] 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 diameter of a particle and the dimensionless persistence time $\tau=\tau_\mathrm{a}D_\mathrm{t}/d^2$ has been introduced in [12, 25].

Since the dimensionless active diffusivity $\mathcal{D}_\mathrm{a}:=D_\mathrm{a}/D_\mathrm{t}$ implicitly depends on the persistence time, it constitutes the most general measure for the activity (together with the persistence length $v_0\tau_\mathrm{a}$). In order to connect to a system of ABPs [12, 25], it is convenient to consider instead of $\mathcal{D}_\mathrm{a}$ a dimensionless velocity $Pe=v_0d/D_\mathrm{t}$, i.e. the Peclét number. In the literature, some other definitions of a Peclét number are used, which we will not consider here. One peculiar property of a system of active OUPs is that, even at vanishing self-propulsion velocity $v_0=0$, or $\mathcal{D}_\mathrm{a}=0$, there is a contribution of $\boldsymbol{v}_i(t)$ to equation (1) arising from a finite reorientation time τ_a [26]. One thus does not recover the equation of motion of a passive (Brownian) particle, as in the ABPs model. In the long-time limit, however, the contribution to the dynamics becomes irrelevant and the same steady state is described as for a passive Brownian particle, see appendix A for more details. A proper passive system can be recovered from equation (1) only in the limit $\tau_\mathrm{a}\to 0$, in the sense that the velocity correlation in equation (4) reduces to a white noise. A Brownian system is then represented by $\mathcal{D}_\mathrm{a}=1$ when the thermal-noise variable $\boldsymbol{\xi}_i(t)$ is removed, or, trivially, by setting $\mathcal{D}_\mathrm{a}=0$ which amounts to neglecting the contribution of the OUPs.

2.2. 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

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

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

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

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

The UCNA [20, 21] amounts to explicitly inserting the OUPs (3) into the overdamped limit of the time derivative of (1), resulting in the modified Langevin equation $\dot{\mathbf{r}}_i(t)=\Gamma^{-1}_{ij}(\mathbf{r}^N)\left(\gamma^{-1}\mathbf{F}_j(\mathbf{r}^N)+\boldsymbol{\xi}_j(t)+\boldsymbol{\eta}_j(t)\right)$. 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 ABPs with a constant velocity v_0 in the direction of their instantaneous orientation $\mathbf{p}_i$ which is subject to Brownian rotational diffusion [12]. This method directly yields the approximate Smoluchowski equation (superscript (f)) with [26, 30]

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

where the regime of validity is the same as for UCNA. The major difference between equations (7) and (8) only impacts 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 [12] the tensor from equation (5) was incorrectly 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.

2.3. Two versions of the steady-state condition

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 equation (9) with $\mathcal{D}_{ik}^{-1}$ and summing over repeated indices, the steady-state condition takes the more instructive (approximate) form [21]

$$0 = \sum_i \mathcal{D}_{ik}^{-1}\cdot\beta\mathbf{F}_i P_N - \boldsymbol{\nabla}_k P_N - P_N\boldsymbol{\nabla}_k\ln|\det\mathcal{D}_{[N]}| =: \beta\mathbf{F}_k^{\mathrm{eff}}P_N - \boldsymbol{\nabla}_k P_N \tag{12}$$

introducing the effective force $\mathbf{F}_k^{\mathrm{eff}}(\mathbf{r}^N)$. The term $\boldsymbol{\nabla}_k\ln|\det\mathcal{D}_{[N]}|$ is an approximation for $\sum_{ij}\mathcal{D}_{ik}^{-1}\cdot\boldsymbol{\nabla}_j\cdot\mathcal{D}_{ji}$, which becomes exact in the UCNA [21]. For the Fox approach, we argue in appendix B that this is still true in some important special cases, such that equation (12) is accurate enough for our purpose. For high particle numbers N the contribution of the off-diagonal elements to $\mathcal{D}_{[N]}$ becomes increasingly irrelevant

[21], which amounts to setting $\mathcal{D}_{ij} \to \delta_{ij}\mathcal{D}_{ij}$. Assuming this *diagonal form*, the determinant in equation (12) can be replaced according to $\det \mathcal{D}_{[N]} \to \det \mathcal{D}_{kk}$ as we have $\sum_{ij} \mathcal{D}^{-1}_{jk} \cdot \boldsymbol{\nabla}_j \cdot \mathcal{D}_{ij} \equiv \mathcal{D}^{-1}_{kk} \cdot \boldsymbol{\nabla}_k \cdot \mathcal{D}_{kk} \approx \boldsymbol{\nabla}_k \ln|\det \mathcal{D}_{kk}|$ before approximating the expression in the last step (compare appendix B).

Putting aside the dynamical behavior, described in section 2.2, 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 $\tau_\mathrm{a} = 0$. 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 $\tau_\mathrm{a} \to 0$. This appears to be an artifact of the pathological contribution of τ_a to the displacement of the OUPs, whereas the connection to the experimentally more relevant system of ABPs is lost. For the latter it appears more natural to tune v_0 at constant τ_a. In the derivation of the Fox result (11), on the other hand, the explicit time evolution of the OUPs in equation (3) is irrelevant, suggesting a better approximate representation of ABPs [12]. This reflects that we recover the (same) passive system for either $v_0 = 0$ or $\tau_\mathrm{a} = 0$ (in the presence of noise).

Ignoring the noise contribution 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 $\mathcal{D}^{-1}_{[N]} \propto \Gamma_{[N]}$ is pairwise additive, even if $D_\mathrm{t} = 0$. Then the effective many-body potential $\mathcal{H}_{[N]}$ defined as $\mathbf{F}^\mathrm{eff}_k(\mathbf{r}^N) = -\boldsymbol{\nabla}_k \mathcal{H}_{[N]}(\mathbf{r}^N)$ can be written in a closed form [20, 21], admitting the explicit solution $P_N(\mathbf{r}^N) \propto \exp(-\beta \mathcal{H}_{[N]}(\mathbf{r}^N))$ of equation (9). Due to the more nested form of equation (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.

3. Effective-potential approximation (EPA)

Regarding the possible applications using standard methods of equilibrium liquid-state theory a desirable strategy is to approximate $\mathbf{F}^\mathrm{eff}_k$ in equation (12) in terms of pair potentials. This approach allows to describe the phase behavior of ABPs approximated as particles propelled by a set of coupled OUPs, which has been discussed in detail [12, 25] for passive soft-repulsive and Lennard-Jones interactions in three dimensions. However, it can be criticized that (I.i) a system which obeys detailed balance is used to represent the interactions in an active system, (I.ii) the validity criteria of the underlying theory might be violated so that further approximations are required and (I.iii) higher-order particle interactions are neglected, which are believed 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 ABPs 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.

3.1. Calculation of effective potentials

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

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

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

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

for the effective external field $\nu^{\text{eff}}(\mathbf{r})$, since for $N=1$ we have $\mathcal{D}_{[1]}=\mathcal{D}_{11}$. Note that a quite different expression for an effective external potential can be derived starting from the equations of motion for ABPs [25, 31].

Integration of the above equalities yields the desired formulas for the effective potentials depending only on the bare potential $u(r)$ or $\nu(\mathbf{r})$ and the activity parameters D_{a} and τ_{a} [12]. Alternatively, we could have directly defined [21, 23] $\nu^{\text{eff}}(\mathbf{r}) := \mathcal{H}_{[1]}(\mathbf{r})$ 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

$$\beta u^{\text{eff}}(r) = \beta \frac{u(r) + \tilde{\tau}(\partial_r u(r))^2}{1+\mathcal{D}_{\text{a}}} - \ln\left| E_1^{(\mathfrak{d}-1)}(\tau,r)\, E_2(\tau,r)\right|, \tag{16}$$

where $\partial_r = \partial/\partial r$ and

$$E_n(\tau,r) := 1 + 2\tilde{\tau} r^{n-2} \partial_r^n u(r), \;\; n\in\{1,2\} \tag{17}$$

are the Eigenvalues of $\Gamma_{[2]}$. We can further identify $\boldsymbol{\nabla}_1 u(r)$ in the first term of equation (13) as the Eigenvector of $\mathcal{D}_{11}^{-1} - \mathcal{D}_{21}^{-1}$ corresponding to the Eigenvalue $E_2/(1+\mathcal{D}_{\text{a}})$. 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 $(1+\mathcal{D}_{\text{a}})$ [21, 23]. 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.

3.2. Limitations and possible corrections

Studying equation (16) more carefully, we notice that the effective potentials do not always behave in a physical way. This is because, in violation of the validity condition

of both the UCNA and the Fox approximation, the diffusion tensor $\mathcal{D}_{[2]}$ is not positive definite for a large number of relevant potentials. In general, we easily see that the logarithm will diverge whenever one of the Eigenvalues $E_n(\tau, r)$ vanishes. Given a positive and convex bare potential $u(r) > 0$, the eigenvalue E_2 is strictly positive, which also means that the effective attraction solely arises from the term including the logaritm. 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 [26].

First note that there is a broader range of admissible bare potentials when the *diagonal approximation*, equation (14), of the effective pair potential is used, or if we are interested in the one-body external field, equation (15). This can be understood from the explicit formula for $u^{\text{eff}}_{\text{diag}}(r)$, which we obtain from equation (16) by rescaling all terms proportional to τ with a factor $1/2$. In the following, we propose different ways to generally rid the effective potential of possible artifacts of vanishing eigenvalues in the last term of equation (13). A correction of the first term is not necessary and also has no noticeable effect.

Let us first assume that $u(r) > 0$ is convex, i.e. it represents a soft-repulsive interaction. Then a sufficient criterion (due to the presence of the term $\mathbf{1}\delta_{ij}$ in equation (5), some other potentials are allowed that are only slightly negative and slightly non-convex) for the matrix $\Gamma_{[N]}$ to have strictly positive eigenvalues would be that it depends on an elliptic differential operator rather than $\boldsymbol{\nabla}_i \boldsymbol{\nabla}_j$. Therefore, a convenient approximation is to redefine equation (5) by an elliptical operator, the simplest example of which is the Laplacian $\boldsymbol{\Delta} = \boldsymbol{\nabla} \cdot \boldsymbol{\nabla}$. Upon substituting

$$\boldsymbol{\nabla}_i \boldsymbol{\nabla}_j \to \mathbf{1} \boldsymbol{\nabla}_i \cdot \boldsymbol{\nabla}_j \tag{18}$$

the effective potential becomes

$$\beta u^{\text{eff}}_{\boldsymbol{\Delta}}(r) = \beta \frac{u(r) + \tilde{\tau}(\partial_r u(r))^2 - 2(\mathfrak{d}-1)\int_r^\infty \mathrm{d}s\, \tilde{\tau} \frac{(\partial_s u(s))^2}{s}}{1 + \mathcal{D}_{\text{a}}} - \ln\left(1 + 2\tilde{\tau}\partial_r^2 u(r) + 2(\mathfrak{d}-1)\tilde{\tau}\frac{\partial_r u(r)}{r}\right) \tag{19}$$

where the additional term compared to (16) cannot be integrated in general. This *Laplacian approximation* has been successfully employed (together with the Fox and diagonal approximation) in explicit calculations [12, 25]. In $\mathfrak{d} = 1$ dimensions both differential operators reduce to the second derivative and equation (19) is equal to equation (16), which provides a good account of active particles interacting with a soft-repulsive potential [23].

An alternative way is to empirically rectify the explicit formula for $u^{\text{eff}}(r)$ in equation (16). Most intuitively, one can expand the argument of the logarithm up to the first order in τ. In fact, this *small-τ approximation* is quite similar to the Laplacian approximation (19) (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 expression (16) 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.

A more elaborate correction that may also be applied 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 is achieved by substituting in equation (16) $E_n(\tau,r)\to E_n^{(\mathrm{i})}(\tau,r)>0$, where

$$E_n^{(\mathrm{i})}(\tau,r) := \begin{cases} 1/(2-E_n(\tau,r)) & \text{if } E_n(\tau,r)<1 \\ E_n(\tau,r) & \text{otherwise} \end{cases} . \tag{20}$$

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 and the effective potential does not diverge if the bare potential is finite. The empirical motivation behind this correction is that E_n constitutes the two leading terms of the 'resummed' Taylor series of $E_n^{(\mathrm{i})}$ in the case $E_n<1$. As described in appendix B, the most convenient implementation of the inverse-τ approximation for the Fox result is to identify the expression for $E_n(\tau,r)$ in $\det\mathcal{D}_{[2]}^{(\mathrm{f})}$ and use equation (20).

3.3. Comparison to computer simulations of two active particles

In section 3.1 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 according to section 3.2 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 $\beta u_{\mathrm{sim}}^{\mathrm{eff}}(r)=-\ln g(r)$. By doing so, we make the same approximation (I.iii) as in the 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.

3.3.1. The role of approximations, dimensionality and thermal noise. 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 result is qualitatively similar. As expected, the full expression for the effective potential in equation (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 figure 1(a), 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 figure 1(a) 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 figure 1(b). As employing the diagonal form $u_{\mathrm{diag}}^{\mathrm{eff}}(r)$ of the effective potential simply amounts to a rescaling of τ, we observe in figure 1(b) that it results in a smaller effective diameter of the repulsive part but a flatter potential well. Accordingly, r_{div} becomes smaller.

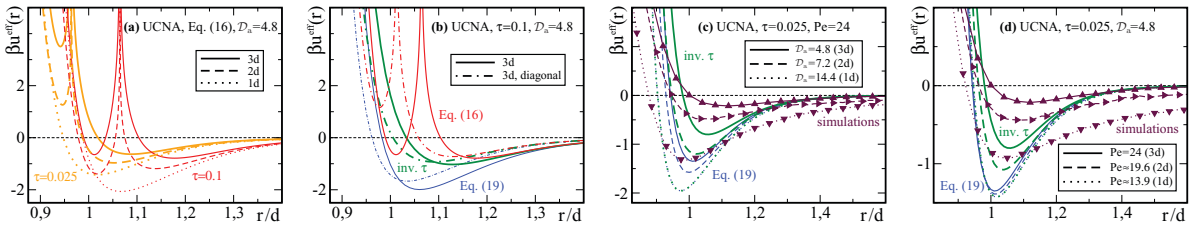

Figure 1. Effective potentials in the presence of thermal noise for a soft-repulsive $\beta u(r) = (r/d)^{-12}$ from the UCNA in $\mathfrak{d} = 1$ (dotted lines), $\mathfrak{d} = 2$ (dashed lines) and $\mathfrak{d} = 3$ (solid lines) dimensions. (a) Full result, equation (16), for the active diffusivity $\mathcal{D}_a = 4.8$ 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 $\mathcal{D}_a = 4.8$ and $\tau = 0.1$. The thick lines correspond to the inverse-τ approximation, equation (20) of equation (16), and the thin lines to the Laplacian approximation, equation (19). (c) Approximate results compared to simulations of active OUPS (lines with triangles) at $\tau = 0.025$ and a constant Peclét number $Pe = \sqrt{\mathfrak{d}\mathcal{D}_a/\tau} = 24$, as $\mathcal{D}_a$ increases with decreasing $\mathfrak{d}$. (d) Approximate results compared to simulations of active OUPS at $\mathcal{D}_a = 4.8$ and $\tau = 0.025$.

Since the definition of the active diffusivity $\mathcal{D}_a$ depends on the dimension $\mathfrak{d}$, not all parameters τ, $\mathcal{D}_a$ and Pe can be kept constant upon varying the dimensionality. For a constant reorientation time τ and propulsion velocity Pe the effective attraction in figure 1(c) 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 [32]. Keeping $\mathcal{D}_a$ constant instead of Pe (which then decreases with decreasing dimension) the same trend is observed in figure 1(d) for the inverse-τ approximation and computer simulations, whereas the result for the Laplacian approximation barely changes with dimensionality. Comparing the qualitative behavior in figures 1(c) 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 [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, equation (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 figures 1(c) 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 [20], where $\boldsymbol{\xi}_i(t)$ in equation (1) was set to zero. Doing so in our simulations also, we observe in figure 2(a) 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 equations (10) and (11), respectively. The full UCNA result is only slightly different in the repulsive regime. Intriguingly, we also recognize in figure 2(a) that the simulation data including the

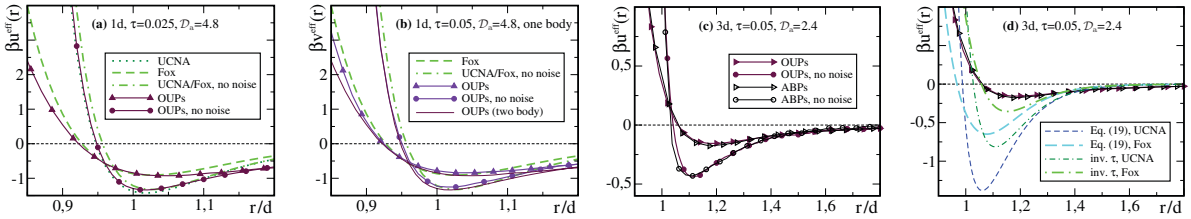

Figure 2. Effective potentials in a soft-repulsive system with and without thermal noise. (a) Comparison to simulations of active OUPS in $\mathfrak{d}=1$ with $\mathcal{D}_a=4.8$ and $\tau=0.025$, as in figure 1(d), including the Fox approximation (dashed lines) and simulations without thermal noise (labeled with dots). In the latter case, UCNA and Fox are equivalent (dot-dashed line). (b) As figure 2(a), but for a single active particle in an external potential $\beta\nu(r)=(r/d)^{-12}$ with $\tau=0.05$ chosen such that the theory predicts the same curves as for two particles with $\tau=0.025$ (the two-body simulation results for $\tau=0.025$ from figure 2(a) are shown as thin solid lines for comparison). (c) Simulations of two ABPs (empty symbols) and OUPs for $\tau=0.05$ and $\mathcal{D}_a=2.4$ in $\mathfrak{d}=3$ dimensions. (d) As figure 2(c) (with noise) compared to the Laplacian approximation, equation (19), and the inverse-τ approximation as labeled.

noise term are excellently represented by the Fox approach. We can understand these observations by recapitulating the idea behind the two approximations. The UCNA amounts to manipulating equation (1) by calculating the second derivative of $\mathbf{r}_i(t)$ in order to eliminate the variable $\boldsymbol{v}_i(t)$ in equation (3). The original discussion of the accuracy of this approximation does not account for the presence of the second stochastic variable $\boldsymbol{\xi}_i(t)$. In contrast, the Fox approach is only dedicated to determine the approximate contribution of the colored noise $\boldsymbol{v}_i(t)$ to the effective probability current in equation (8), which is independent of other terms in equation (1). Therefore, the Fox theory has a broader range of applicability and should be accurate in both the presence and the absence of thermal noise.

Finally, we note that the excellent agreement between theory and computer simulations in one dimension implies that the diagonal approximation is not justified when it comes to describing a two-body system. It is, however, interesting to consider a single particle in an external field of the same form as the interparticle potential considered above. In agreement with the theoretical prediction, the computer simulations in figure 2(b) show nice agreement between the two-body system and a one-body system with the double value $\tau=0.05$ of the persistence time. We further observe that the theoretical result for one body is even closer to the simulation data than for two bodies.

3.3.2. Soft-repulsive Brownian system in three dimensions. We also performed computer simulations of ABPs (described in appendix A) for $\mathfrak{d}=3$. For a finite active diffusivity, figure 2(c) reveals that the numerical effective potentials for the two considered models with and without 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. Quantitatively, this difference is much more pronounced in three dimensions. In the following, we restrict ourselves to systems with thermal noise and compare in figures 2(d) and 3 to the predictions of the theory. As in figures 2(c) and (d), our simulations of ABPs and OUPs, shown in the first column of figure 3, are in nice agreement for all sets of considered

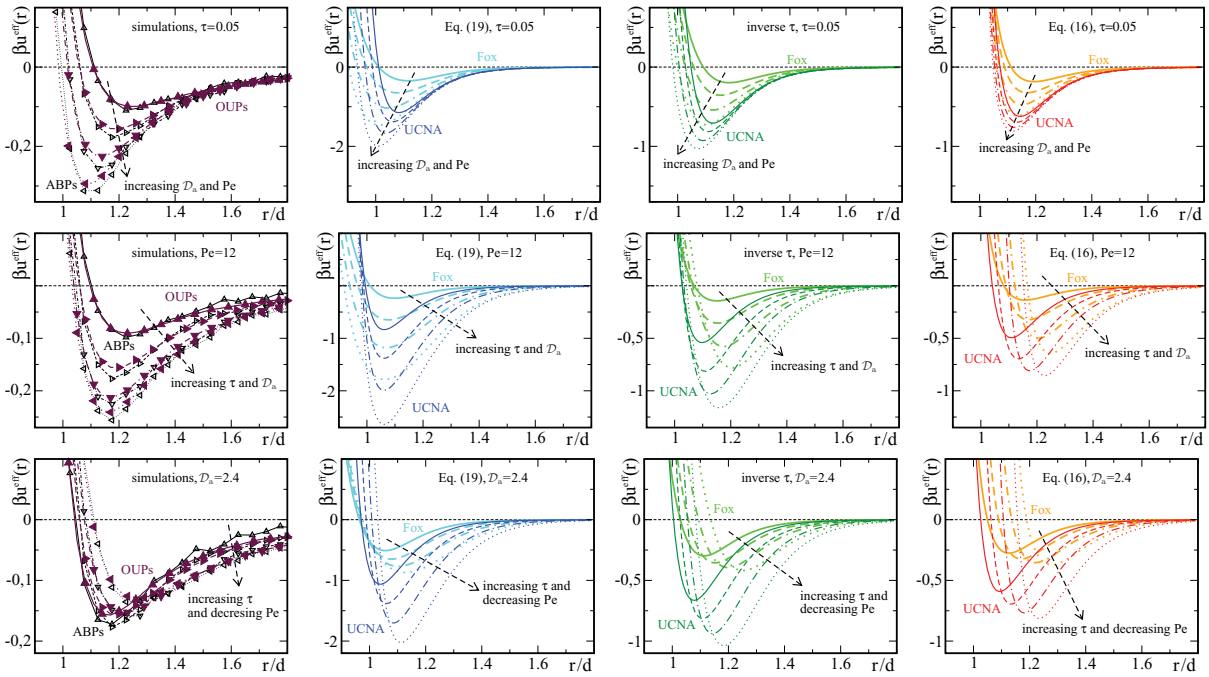

Figure 3. Effective potentials in three dimensions for soft-repulsive spheres, $\beta u(r)=(r/d)^{-12}$, obtained within the Fox (thick, brighter lines) and UCNA (thin, darker lines) and by numerical simulations of ABPs (empty triangles) and active OUPs (filled triangles). Columns from left to right: simulations; Laplacian approximation, equation (19); inverse-τ approximation, equations (16) with (20); equation (16) only for $r>r_{\mathrm{div}}$ (see figure 1(a) for the full result). Rows from top to bottom: increasing the active diffusivity $\mathcal{D}_{\mathrm{a}}$ at constant persistence time $\tau=0.05$; increasing $\mathcal{D}_{\mathrm{a}}$ and τ at constant Peclét number $Pe=\sqrt{3\mathcal{D}_{\mathrm{a}}/\tau}=12$; increasing τ at constant $\mathcal{D}_{\mathrm{a}}=2.4$. The sequence from the solid to the dotted lines corresponds to increasing the respective parameter(s) $\mathcal{D}_{\mathrm{a}}$ from $\mathcal{D}_{\mathrm{a}}=1.2$ or τ from $\tau=0.025$ by a factor of two in each step (dashed lines are always for the same set of parameters).

parameters. This is quite surprising since, on the many-particle level, ABPs and OUPs have different steady states [33, 34]. On the basis of our data for the simplistic two-body system we could rather conclude that OUPs subjected to thermal noise are an excellent model for ABPs at moderate activity [12, 25].

We see in figure 2(d) that the depth of the attractive well of all theoretical versions of the effective potential in $\mathfrak{d}=3$ dimensions is significantly overestimated when compared to the simulations of both ABPs and OUPs. The inverse-τ approximation appears to provide the best guess of the point at which the effective potential changes its sign. To facilitate further qualitative comparison we chose the y axes in figure 3 according to the deviation from the simulations (first column), i.e. by a factor of 10 for the Laplacian approximation (second column) and 5 for the inverse-τ approximation (third column). The chosen approximations exhibit a behavior similar to the full theoretical results of equation (16) in the physical region for $r>r_{\mathrm{div}}$, shown in the last column. The differences arising from using the exact effective force in equation (12) are discussed in appendix B.

For the considered soft-repulsive bare potential, we observe in figure 3 some notable quantitative differences between the UCNA and Fox results, even at relatively high

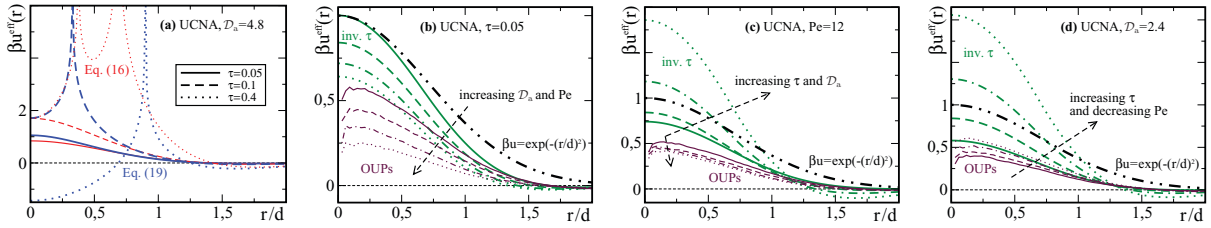

Figure 4. Effective potentials from the UCNA in three dimensions for Gaussian-core particles, $\beta u(r) = \exp(-(r/d)^2)$. (a) Comparison of the divergences in the full result from equation (16) and the Laplacian approximation, equation (19), for different persistence times τ. (b)–(d) Inverse-τ approximation and computer simulations for active OUPs for the same parameters as in rows 1–3 of figure 3, respectively. The bare potential is shown as the thick dot–dot-dashed line.

$\mathcal{D}_a$. 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 figure 3 contains the effective potentials evolving for a constant persistence time τ when the active diffusivity $\mathcal{D}_a$ (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 [12, 25]. The Fox results exhibit a stronger variation with $\mathcal{D}_a$, which is also 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 almost unaffected. The most interesting behavior is observed in the third column at constant $\mathcal{D}_a$. 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. Regarding figure 1(b) one might get the impression that the additional assumption of the diagonal form of the effective potential also results in a slightly better (quantitative) agreement with the simulations. However, we stress that it is not clear 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.

3.3.3. Difficulties for non-convex potentials. The most compelling argument in favor 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. This issue has already been discussed for a Lennard-Jones potential [30] and we will consider here the potential $\beta u(r) = \exp(-(r/d)^2)$ of an active Gaussian-core fluid. Although this model has not received much attention in theories for ABPs, 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 [35, 36], 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 figure 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 $\mathcal{D}_a$ considered here. When the persistence time τ is sufficiently small, the effective potential, equation (16), does not diverge and the different approximations behave in a quite similar way. Interestingly, we observe in figure 4(a) that the divergence of the Laplacian approximation sets in at an even smaller value of $\tau \geqslant 1/12$ than for the full potential. The latter diverges at two points, each related to one of the two eigenvalues, if $\tau \geqslant 1/4$.

In figures 4(b)–(d) we discuss the only suitable form of the effective potential, i.e. 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. Counterintuitively, we observe that in this case the effective interaction becomes more repulsive than in the passive case and also for the inverse-τ approximation. Our computer simulations (also carried out for values of τ much larger than shown in figure 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.

To argue about the validity of the EPA, we consider two classes of bare interactions. Firstly, soft-repulsive and convex potentials lead to quite accurate results in one dimension, but require an empirical correction in higher dimensions. Secondly, 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. At the moment, the only way to obtain a workable theory in this case is by employing the inverse-τ correction introduced in equation (20). Further numerical and theoretical analysis will be needed to clarify this issue fully. Finally, we note that potentials with attractive parts do not *a priori* constitute a problem for the theory, but usually have regions in which they are non convex. For a discussion of such problems see also [26, 30].

4. 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 equation (9) over $N - n$ coordinates [21], which allows defining a mechanical pressure and interfacial tension [22]. Moreover, for the active system evolving according to equations (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 [37]. This suggests that, at leading order in this parameter, the approximations resulting in equation (9) are perfectly justified.

Knowing, however, that equation (9) contains the same information as equation (12), which depends logarithmically on the parameter τ, we should clarify whether (II.i) the low-activity expansion converges, (II.ii) it is sufficient to only consider the leading order and (II.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 equation (12) and discuss the consequences of approximating $\mathbf{F}_k^{\mathrm{eff}}$ 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.

4.1. Alternative derivation of a local force-balance condition

By saying we integrate a multidimensional vector equation (label i) over $N-1$ coordinates we understand multiplying each side by $\delta(\boldsymbol{r}-\boldsymbol{r}_i)$, followed by summation over all particles i and integration over all N spatial coordinates. We further define the average $\langle\mathbf{X}\rangle := \langle\!\langle\sum_{i=1}^N \delta(\boldsymbol{r}-\boldsymbol{r}_i)\mathbf{X}_i\mathbf{X}\rangle\!\rangle/\rho(\mathbf{r})$ of a vector $\boldsymbol{X}_i(\boldsymbol{r}^N)$, where $\langle\!\langle\cdot\rangle\!\rangle$ denotes the full canonical ensemble average and $\rho(\boldsymbol{r}) = \langle\!\langle\hat{\rho}\rangle\!\rangle$ is the average of the density operator $\hat{\rho}=\sum_{i=1}^N \delta(\mathbf{r}-\mathbf{r}_i)$. Approximating now the inverse mobility matrix in equation (10) as $\Gamma_{ij}^{-1}(\mathbf{r}^N)\approx(\mathbf{1}-\tilde{\tau}\boldsymbol{\nabla}_i\boldsymbol{\nabla}_j\mathcal{U})\delta_{ij}$ and integrating equation (9) over $N-1$ coordinates, we find the first member

$$-\rho(\mathbf{r})\langle\boldsymbol{\nabla}\beta\mathcal{U}\rangle = (1+\mathcal{D}_{\mathrm{a}})\left(\boldsymbol{\nabla}\cdot(\mathbf{1}\rho(\mathbf{r})-\tilde{\tau}\rho(\mathbf{r})\langle\boldsymbol{\nabla}\boldsymbol{\nabla}\mathcal{U}\rangle)\right), \tag{21}$$

of a YBG-like hierarchy [21, 22] for the active system, where, explicitly,

$$\langle\mathfrak{D}\mathcal{U}\rangle = \mathfrak{D}\nu(\mathbf{r}) + \int \mathrm{d}\mathbf{r}'\frac{\rho^{(2)}(\mathbf{r},\mathbf{r}')}{\rho(\mathbf{r})}\mathfrak{D}u(\mathbf{r},\mathbf{r}') \tag{22}$$

for any (nontrivial) differential operator $\mathfrak{D}_i$ acting on $\boldsymbol{r}_i$. In the derivation of (21) it turns out that the off-diagonal components of the mobility tensor do not contribute at first order in τ [22]. 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 equation (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 in appendix C. Making use of the equilibrium version of the YBG hierarchy and expanding the expression $\ln(\det\Gamma_{ii}(\mathbf{r}^N))$ up to first order in τ the result is

$$0 = -\boldsymbol{\mathcal{D}}_{\mathrm{I}}^{-1}(\mathbf{r})\rho(\mathbf{r})\langle\boldsymbol{\nabla}\beta\mathcal{U}\rangle - \boldsymbol{\nabla}\rho(\mathbf{r}) + \tilde{\tau}\rho(\mathbf{r})\langle\boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\boldsymbol{\nabla}\mathcal{U}\rangle + \frac{\rho(\mathbf{r})}{1+\mathcal{D}_{\mathrm{a}}}\tilde{\tau}\int \mathrm{d}\mathbf{r}'\left(\boldsymbol{\nabla}\boldsymbol{\nabla}u(\mathbf{r},\mathbf{r}')\right)\cdot\boldsymbol{\nabla}\frac{\rho^{(2)}(\mathbf{r},\mathbf{r}')}{\rho(\mathbf{r})} \tag{23}$$

introducing the averaged inverse diffusion tensor (compare equation (10))

$$\boldsymbol{\mathcal{D}}_{\mathrm{I}}^{-1}(\mathbf{r}) := \frac{\langle\Gamma_{ii}\rangle}{1+\mathcal{D}_{\mathrm{a}}} = \frac{\mathbf{1}+\tilde{\tau}\langle\boldsymbol{\nabla}\boldsymbol{\nabla}\mathcal{U}\rangle}{1+\mathcal{D}_{\mathrm{a}}}. \tag{24}$$

Multiplying equation (23) with $\boldsymbol{\mathcal{D}}_{\mathrm{I}}\approx(1+\mathcal{D}_{\mathrm{a}})(\mathbf{1}-\tilde{\tau}\langle\boldsymbol{\nabla}\boldsymbol{\nabla}\mathcal{U}\rangle)$ it is easy to verify in appendix C that at first order in τ it becomes equivalent to (21) up to a term proportional to the expression in the second line, which we consider as a higher-order contribution.

In order to derive equation (23) in the Fox approach, an additional approximation is required, as the inverse of $\mathcal{D}_{ij}^{(\mathrm{f})}$ from equation (11) is not proportional to Γ_{ij}. We would thus need to redefine $\boldsymbol{\mathcal{D}}_{\mathrm{I}}$ in equation (24) according to equation (11) where $\langle\Gamma_{ii}\rangle^{-1}$ takes the role of Γ_{ij}^{-1}. Regarding the presented alternative derivation of equation (21) in general, its validity appears to be in question. This is because to derive the intermediate result in equation (23) it is necessary to expand a logarithmic term, the Taylor series of which only has a finite radius of convergence. We further assumed explicitly the diagonal form of the mobility tensor to avoid further terms that are not present in the original result. Employing the EPA in the next step will shed more light on these issues.

4.2. 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 $\mathbf{F}_k^{\mathrm{eff}} \approx -\boldsymbol{\nabla}_k \mathcal{U}^{\mathrm{eff}} = -\boldsymbol{\nabla}_k \left(\nu^{\mathrm{eff}}(\mathbf{r}_k) + \sum_{l=k} u^{\mathrm{eff}}(\mathbf{r}_k, \mathbf{r}_l) \right)$ along the lines of (2) but within the EPA using the results derived in section 3.1. As detailed in appendix D, the obvious result is that all correlation functions between more than two particles vanish in the approximate integrated version

$$0 = -\boldsymbol{\nabla}\rho(\mathbf{r}) - \rho(\mathbf{r}) \left\langle \boldsymbol{\nabla}\beta\mathcal{U}^{\mathrm{eff}} \right\rangle \tag{25}$$

of (12). Ignoring the interparticle interactions the approximation involving only $\nu^{\mathrm{eff}}(\mathbf{r})$ becomes exact. This situation is the same as discussed in [21].

Considering the interacting system, we multiply equation (25) with $\boldsymbol{\mathcal{D}}_{\mathrm{I}}$ as done previously for (23). According to appendix D, we can only approximately reproduce equation (21) by doing so. This reflects both the limitations of the EPA and an inconsistency between equations (9) and (12) when they are subject to the same type of approximation, as we discuss in the following. We observe that (III.i) the coupling between external and internal interactions is ignored by equation (25) (III.ii) spurious three-body correlations appear on the left-hand-side of equation (21) (III.iii) the second term on the right-hand side of equation (21) is recovered but involves a seemingly unjustified expansion and (III.iv) if we do not explicitly assume a diagonal diffusion tensor, the last term in equation (21) changes by a factor two. As we are mainly interested in bulk systems, the first point is only briefly commented on in appendix D.

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

$$\sum_{l,j=k} \left(\boldsymbol{\nabla}_k \boldsymbol{\nabla}_k u(\mathbf{r}_k, \mathbf{r}_l)\right) \left(\boldsymbol{\nabla}_k u(\mathbf{r}_k, \mathbf{r}_j)\right) \longrightarrow \sum_{j=k} \left(\boldsymbol{\nabla}_k \boldsymbol{\nabla}_k u(\mathbf{r}_k, \mathbf{r}_j)\right) \left(\boldsymbol{\nabla}_k u(\mathbf{r}_k, \mathbf{r}_j)\right), \tag{26}$$

should not be too crude. Moreover, we have discussed in section 3.2 that such a 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 D that the definition (24) of the averaged diffusion tensor $\boldsymbol{\mathcal{D}}_{\mathrm{I}}$ is not fully compatible with the EPA, resulting in point (III.ii). This is in contrast to the clean derivation of equation (21), where equation (9) is recovered from equation (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. Recall from the discussion in section 3 that an approximation as a pairwise quantity might be quite poor and an expansion of the logarithm does not converge. However, we demonstrate in appendix D that successively employing the EPA and expanding for small $\tilde{\tau}$ according to

$$\begin{aligned}\boldsymbol{\nabla} \ln\left|\det\left(\mathbf{1}+\tilde{\tau}\boldsymbol{\nabla}\boldsymbol{\nabla}\sum_{l>1}u(\mathbf{r},\mathbf{r}_l)\right)\right| &\longrightarrow \boldsymbol{\nabla}\cdot\sum_{l>1}\ln|\det\left(\mathbf{1}+\tilde{\tau}\boldsymbol{\nabla}\boldsymbol{\nabla}u(\mathbf{r},\mathbf{r}_l)\right)| \\ &\longrightarrow \tilde{\tau}\sum_{l>1}\boldsymbol{\nabla}\boldsymbol{\Delta}u(\mathbf{r},\mathbf{r}_l)+\mathcal{O}(\tau^2) \end{aligned} \tag{27}$$

eventually results in full consistency with the respective term in equation (21), stated as point (III.iii). This suggests that the expansion to first order in τ "implies" making the EPA when equation (12) is our starting point. Despite the aforementioned crudity of this expansion, we argue that equation (21) is valid, as its clean derivation from equation (9) does not require dealing with a logarithmic term. The last step in equation (27) is required to recover equation (21) without inducing undesired higher-order terms in τ, as, similar to point (III.ii), the integrated version is incompatible with the chosen $\boldsymbol{\mathcal{D}}_\mathrm{I}$. However, we stress that this approximation should certainly be avoided when calculating the fluid structure.

Finally, we demonstrate in appendix D that the off-diagonal elements of the diffusion tensor entering in equation (12) contribute to equation (25). Hence, the present approach would be even more inconsistent with equation (21) if we did not assume the diagonal form, as noted in point (III.iv). We also note that the same problem occurs for the according generalization of equation (23). In principle we could define in this case an additional averaged diffusion tensor $\boldsymbol{\mathcal{D}}_{\mathrm{I,od}}$, 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.

5. 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 translational Brownian noise cannot be neglected. In the absence of 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 [30]. 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 deterministic Fokker–Planck equation (effective equilibrium picture) and then we define pair forces from the two-particle limit assuming a vanishing probability current. 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 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 equations (9) and (12), although the latter contains a logarithmic term. Relatedly, it was recognized in [22] 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 the route to follow should be carefully chosen for each problem, together with the underlying 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 equation (12) supported by effective pair potentials is most convenient for accessing structural properties [12, 23, 25], whereas the low-activity expansion of equation (9) provides a direct way to define mechanical properties [22]. Moreover, our analysis suggests that the thermodynamic results obtained from equation (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 $1 + \mathcal{D}_a$ of an ideal gas, which can be absorbed into an effective temperature [21, 22, 25]. A more general approach will be detailed in the second paper of this series.

Acknowledgments

R Wittmann, A Scacchi and J M Brader acknowledge funding provided by the Swiss National Science Foundation. C Maggi acknowledges support from the European Research Council under the European Unions Seventh Framework programme (FP7/2007-2013)/ERC Grant agreement No. 307940.

Appendix A. Simulation details

We performed Brownian dynamics simulations of a system composed of two particles of unit diameter $d = 1$ interacting through a soft-repulsive potential or a Gaussian soft-core potential. The potential is truncated at a distance of $r = 2d$. In the simulations of active OUPs, evolving according to equation (1), each particle is subjected to Gaussian thermal noise and non-Gaussian (colored) noise. The latter yields two distinct

contributions to the displacement of each particle: one drift term, proportional to the reorientation time τ and one Gaussian process, proportional to $\sqrt{\mathcal{D}_\mathrm{a}/\tau}$ [26]. For a vanishing active diffusivity $\mathcal{D}_\mathrm{a}$, the drift term decays exponentially in time and is therefore irrelevant in the long-time limit. 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 the simulation is $10^6\tau_\mathrm{B}$. 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 have verified that, for the case of $\mathcal{D}_\mathrm{a} = 0$ and finite τ, the obtained pair-correlation function is independent of τ, although the short-time displacement is not.

We also performed Brownian dynamics simulation of ABPs, for which the colored-noise variable $\boldsymbol{v}_i(t)$ in equation (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. Effective many-body force in the Fox approximation

In this appendix we discuss the accuracy of equation (12) of the main text in the Fox approximation, i.e. choosing the effective diffusion tensor from equation (11). The accurate definition of the effective force is

$$\beta\mathbf{F}_k^\mathrm{eff} = \sum_i \mathcal{D}_{ik}^{-1} \cdot \beta\mathbf{F}_i - \sum_{ij} \mathcal{D}_{ik}^{-1} \cdot \boldsymbol{\nabla}_j \cdot \mathcal{D}_{ji} \tag{B.1}$$

since the conversion

$$\sum_{ij} \mathcal{D}_{ik}^{-1} \cdot \boldsymbol{\nabla}_j \cdot \mathcal{D}_{ji} \approx \boldsymbol{\nabla}_k \ln|\det \mathcal{D}_{[N]}| \tag{B.2}$$

is only correct in the following cases:

1. for a system with no thermal noise. As stated in the main text, in this case the Fox and UCNA results are equivalent. Making use of the symmetry relation $\partial_\gamma \mathcal{D}_{\alpha\beta}^{-1} = \partial_\beta \mathcal{D}_{\alpha\gamma}^{-1}$ (with Greek indices labeling components and particles) and Jacobi's formula the identity in equation (B.2) can be explicitly verified [21].
2. for a passive system, since $\mathcal{D}_{ij} \equiv \mathbf{1}\delta_{ij}$.
3. at leading order in the activity parameter τ, where $\mathcal{D}_{\alpha\beta}^{-1} = \delta_{\alpha\beta}(1+\mathcal{D}_\mathrm{a}) - \tilde{\tau}\mathcal{D}_\mathrm{a}\partial_\alpha\partial_\beta\mathcal{U} + \mathcal{O}(\tilde{\tau}^2)$ and the same arguments as under point 1. can be used.
4. for $N \leqslant 2$ particles in an effectively one-dimensional symmetry, i.e. if there exists a coordinate frame in which the non-trivial contributions to $\mathcal{D}_{[N]}$ reduce to an at most a 2×2 tensor with identical diagonal elements. This can be easily shown by an explicit calculation

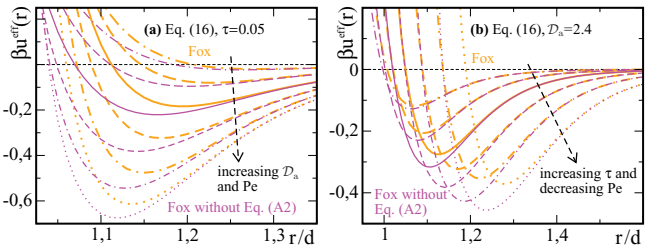

Figure B1. Closeups of the Fox results for the full effective potentials from equation (B.2) as in figure 3 (a) top right with additional data for $\mathcal{D}_a = 0.6$ and $\mathcal{D}_a = 0.3$ and (b) bottom right with additional data for $\tau = 0.0125$ and $\tau = 0.006\,25$. The thinner lines correspond to the results without making the approximation in equation (B.2).

(a) in $\mathfrak{d} = 1$ dimensions

(b) for a planar interaction potential

(c) in the Laplacian approximation (18)

As a simple counter example to the cases listed above, we note that equation (B.2) does not generally hold for $N = 1$ and $\mathfrak{d} = 2$, since $\partial_x \nu(x, y) = -\partial_y \nu(x, y)$, whereas under point 4. we have $\partial_1 u(x_1 - x_2) = -\partial_2 u(x_1 - x_2)$. As for the approximate formulas (13) and (14), we find that the difference between the effective forces, equation (B.1), for $N = 2$ particles with and without diagonal approximation is only a factor 2 in front of each factor τ.

Comparing the requirements for points 1. and 2. we can say that equation (12) is correct for both $\mathcal{D}_a \gg \mathcal{D}_t$ and $\mathcal{D}_a \ll \mathcal{D}_t$, indicating that it should be a good approximation over all ranges of the parameter $\mathcal{D}_a$. Moreover, the assumption of a small persistence time τ is required in the derivation of the effective equilibrium approach [12, 20, 21]. Considering point 3, this means that the approximation in equation (B.2) is consistent with the underlying theory. Indeed, figures B1(a) and (b) show that the approximation is best for either small or large $\mathcal{D}_a$ and small τ, respectively. In general, the difference is not significant compared to other approximations shown in figure 3 of the main text.

If the validity criterion $E_n(\tau, r) > 0$ for the eigenvalues, given by equation (17), of $\Gamma_{[2]}$ is violated, neither side of equation (B.2) results in physically effective potentials. Therefore, the most important benefit of the approximate form on the right-hand side is that it enables the inverse-τ approximation to be employed, as described in section 3.2: the eigenvalues $E_n^{(\mathrm{f})}(\tau, \mathcal{D}_a, r)$ of $\mathcal{D}_{[2]}^{(\mathrm{f})}$ can be written as

$$E_n^{(\mathrm{f})}(\tau, \mathcal{D}_a, r) = 1 + \frac{\mathcal{D}_a}{E_n(\tau, r)} \tag{B.3}$$

so that we can substitute $E_n(\tau, r)$ according to equation (20) of the main text. The substitution of $E_n^{(\mathrm{f})}$ or a more general manipulation of $\mathcal{D}_{[2]}^{(\mathrm{f})}$ is inconvenient since the effective potential would still diverge for $E_n = -\mathcal{D}_a$ although the bare potential does not. Therefore, the third column of figure 3 contains the optimal implementation of the the inverse-τ approximation for the Fox approach. Also recall that, according to point

4.(c), the results in the Laplacian approximation are the same for both expressions in equation (B.2).

Appendix C. Integration of the first equality in equation (12)

The derivation of equation (21) by integrating equation (9) over $N-1$ coordinates is quite similar to that of the YBG hierarchy in a passive system. The first member

$$0 = \boldsymbol{\nabla}\rho(\mathbf{r}) + \rho(\mathbf{r})\boldsymbol{\nabla}\beta\nu(\mathbf{r}) + \int \mathrm{d}\mathbf{r}'\rho^{(2)}(\mathbf{r},\mathbf{r}')\boldsymbol{\nabla}\beta u(\mathbf{r},\mathbf{r}') = \rho(\mathbf{r})\boldsymbol{\nabla}\mu \tag{C.1}$$

is recovered from (21) 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 the chemical potential μ, which is constant in equilibrium. The second member reads

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

and is related via the second equality to the first member. With the help of these exact equilibrium sum rules we will now derive equation (23) by integrating equation (12) over $N-1$ coordinates. Our presentation closely follows the derivation of a dynamical density functional theory including a tensorial diffusivity [38], whereas we only consider the steady-state condition.

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

$$0 = \sum_i \mathcal{D}^{-1}_{ik} \cdot \left(-\boldsymbol{\nabla}_i P_N + \beta \boldsymbol{F}_i P_N\right) + \mathcal{D}^{-1}_{ik} \cdot \boldsymbol{\nabla}_i P_N - \boldsymbol{\nabla}_k P_N + P_N \boldsymbol{\nabla}_k \ln|\det \mathcal{D}^{-1}_{kk}|, \tag{C.3}$$

where we further used the concept 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 (C.3) over $N-1$ coordinates yields (within UCNA)

$$\begin{aligned}
&\frac{1}{1+\mathcal{D}_\mathrm{a}}\Bigg(-\boldsymbol{\nabla}\rho(\mathbf{r}) - \rho(\mathbf{r})\boldsymbol{\nabla}\beta\nu(\mathbf{r}) - \int \mathrm{d}\mathbf{r}'\rho^{(2)}(\mathbf{r},\mathbf{r}')\boldsymbol{\nabla}\beta u(\mathbf{r},\mathbf{r}') \\
&- \tilde{\tau}\boldsymbol{\nabla}\boldsymbol{\nabla}\nu(\mathbf{r}) \cdot \left(\boldsymbol{\nabla}\rho(\mathbf{r}) + \rho(\mathbf{r})\boldsymbol{\nabla}\beta\nu(\mathbf{r}) + \int \mathrm{d}\mathbf{r}'\rho^{(2)}(\mathbf{r},\mathbf{r}')\boldsymbol{\nabla}\beta u(\mathbf{r},\mathbf{r}')\right) \\
&- \tilde{\tau}\int \mathrm{d}\mathbf{r}''\boldsymbol{\nabla}\boldsymbol{\nabla}u(\mathbf{r},\mathbf{r}'') \cdot \left(\boldsymbol{\nabla}\rho^{(2)}(\mathbf{r},\mathbf{r}'') + \rho^{(2)}(\mathbf{r},\mathbf{r}'')\boldsymbol{\nabla}\left(\beta\nu(\mathbf{r}) + \beta u(\mathbf{r},\mathbf{r}'')\right) + \int \mathrm{d}\mathbf{r}'\rho^{(3)}(\mathbf{r},\mathbf{r}',\mathbf{r}'')\boldsymbol{\nabla}\beta u(\mathbf{r},\mathbf{r}')\right) \\
&+ \boldsymbol{\nabla}\rho(\mathbf{r}) + \tilde{\tau}\left(\boldsymbol{\nabla}\boldsymbol{\nabla}\nu(\mathbf{r})\right) \cdot \boldsymbol{\nabla}\rho(\mathbf{r}) + \tilde{\tau}\int \mathrm{d}\mathbf{r}'\left(\boldsymbol{\nabla}\boldsymbol{\nabla}u(\mathbf{r},\mathbf{r}')\right) \cdot \boldsymbol{\nabla}\rho^{(2)}(\mathbf{r},\mathbf{r}')\Bigg) \\
&- \boldsymbol{\nabla}\rho(\mathbf{r}) + N\int \mathrm{d}\mathbf{r}_2 \ldots \int \mathrm{d}\mathbf{r}_N P_N(\mathbf{r},\mathbf{r}_2,\ldots,\mathbf{r}_N)\boldsymbol{\nabla}\ln\left|\det\left(\mathbf{1} + \tilde{\tau}\boldsymbol{\nabla}\boldsymbol{\nabla}\left(\nu(\mathbf{r}) + \sum_{l>1} u(\mathbf{r},\mathbf{r}_l)\right)\right)\right| = 0
\end{aligned} \tag{C.4}$$

As the Fox result (11) for $\mathcal{D}^{-1}_{ik}$ is not a 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 (C.1) and (C.2) and expand the logarithm up to linear order in τ. The result is

$$
\begin{aligned}
&-\frac{1}{1+\mathcal{D}_\mathrm{a}}\left(\mathbf{1}+\tilde{\tau}\boldsymbol{\nabla}\boldsymbol{\nabla}\nu(\mathbf{r})+\tilde{\tau}\int \mathrm{d}\mathbf{r}''\frac{\rho^{(2)}(\mathbf{r},\mathbf{r}'')}{\rho(\mathbf{r})}\boldsymbol{\nabla}\boldsymbol{\nabla}u(\mathbf{r},\mathbf{r}'')\right)\cdot\left(\rho(\mathbf{r})\boldsymbol{\nabla}\beta\nu(\mathbf{r})+\int \mathrm{d}\mathbf{r}'\rho^{(2)}(\mathbf{r},\mathbf{r}')\boldsymbol{\nabla}\beta u(\mathbf{r},\mathbf{r}')\right)\\
&-\boldsymbol{\nabla}\rho(\mathbf{r})+\tilde{\tau}\rho(\mathbf{r})\left(\boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\boldsymbol{\nabla}\nu(\mathbf{r})\right)+\tilde{\tau}\int \mathrm{d}\mathbf{r}'\rho^{(2)}(\mathbf{r},\mathbf{r}')\boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\boldsymbol{\nabla}u(\mathbf{r},\mathbf{r}')\\
&+\frac{1}{1+\mathcal{D}_\mathrm{a}}\tilde{\tau}\int \mathrm{d}\mathbf{r}'\left(\boldsymbol{\nabla}\boldsymbol{\nabla}u(\mathbf{r},\mathbf{r}')\right)\rho(\mathbf{r})\cdot\boldsymbol{\nabla}\frac{\rho^{(2)}(\mathbf{r},\mathbf{r}')}{\rho(\mathbf{r})}=0
\end{aligned}
\tag{C.5}
$$

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 equation (C.4) with $\frac{\rho^{(2)}(\mathbf{r},\mathbf{r}'')}{\rho(\mathbf{r})}\boldsymbol{\nabla}\rho(\mathbf{r})$, which does not cancel out with the expression in the fourth line of equation (C.4).

For convenience we adopt the notational convention of the main text (22) and identify the $\mathfrak{d}\times\mathfrak{d}$ matrix in the first line of equation (C.5) as the inverse of an ensemble-averaged diffusion tensor

$$
\begin{aligned}
\boldsymbol{\mathcal{D}}_\mathrm{I}(\mathbf{r}) &:= (1+\mathcal{D}_\mathrm{a})\left(\mathbf{1}+\tilde{\tau}\boldsymbol{\nabla}\boldsymbol{\nabla}\nu(\mathbf{r})+\tilde{\tau}\int \mathrm{d}\mathbf{r}'\frac{\rho^{(2)}(\mathbf{r},\mathbf{r}')}{\rho(\mathbf{r})}\boldsymbol{\nabla}\boldsymbol{\nabla}u(\mathbf{r},\mathbf{r}')\right)^{-1}\\
&=\left(\frac{\mathbf{1}+\tilde{\tau}\langle\boldsymbol{\nabla}\boldsymbol{\nabla}\mathcal{U}\rangle}{1+\mathcal{D}_\mathrm{a}}\right)^{-1}=(1+\mathcal{D}_\mathrm{a})\left(\mathbf{1}-\tilde{\tau}\langle\boldsymbol{\nabla}\boldsymbol{\nabla}\mathcal{U}\rangle\right)+\mathcal{O}(\tau^2).
\end{aligned}
\tag{C.6}
$$

Now we multiply equation (C.5) with $\boldsymbol{\mathcal{D}}_\mathrm{I}$ and drop all higher-order terms $\propto\tau^2$, which yields

$$
\begin{aligned}
\rho(\mathbf{r})\langle\boldsymbol{\nabla}\beta\mathcal{U}\rangle &= (1+\mathcal{D}_\mathrm{a})\left(-\boldsymbol{\nabla}\rho(\mathbf{r})+\tilde{\tau}\left(\left(\boldsymbol{\nabla}\rho(\mathbf{r})\right)\cdot\langle\boldsymbol{\nabla}\boldsymbol{\nabla}\mathcal{U}\rangle+\rho(\mathbf{r})\langle\boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\boldsymbol{\nabla}\mathcal{U}\rangle\right)\right)\\
&\quad+\tilde{\tau}\int \mathrm{d}\mathbf{r}'\left(\boldsymbol{\nabla}\boldsymbol{\nabla}u(\mathbf{r},\mathbf{r}')\right)\rho(\mathbf{r})\cdot\boldsymbol{\nabla}\frac{\rho^{(2)}(\mathbf{r},\mathbf{r}')}{\rho(\mathbf{r})}\\
&=(1+\mathcal{D}_\mathrm{a})\left(-\boldsymbol{\nabla}\rho(\mathbf{r})+\tilde{\tau}\boldsymbol{\nabla}\cdot\left(\rho(\mathbf{r})\langle\boldsymbol{\nabla}\boldsymbol{\nabla}\mathcal{U}\rangle\right)\right)\\
&\quad-\mathcal{D}_\mathrm{a}\tilde{\tau}\int \mathrm{d}\mathbf{r}'\left(\boldsymbol{\nabla}\boldsymbol{\nabla}u(\mathbf{r},\mathbf{r}')\right)\rho(\mathbf{r})\cdot\boldsymbol{\nabla}\frac{\rho^{(2)}(\mathbf{r},\mathbf{r}')}{\rho(\mathbf{r})}.
\end{aligned}
\tag{C.7}
$$

In the last step we made use of the identity

$$
\boldsymbol{\nabla}\cdot\left(\rho(\mathbf{r})\langle\boldsymbol{\nabla}\boldsymbol{\nabla}\beta\mathcal{U}\rangle\right)=\left(\boldsymbol{\nabla}\rho(\mathbf{r})\right)\cdot\langle\boldsymbol{\nabla}\boldsymbol{\nabla}\beta\mathcal{U}\rangle+\rho(\mathbf{r})\langle\boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\boldsymbol{\nabla}\beta\mathcal{U}\rangle+\rho(\mathbf{r})\int \mathrm{d}\mathbf{r}'\left(\boldsymbol{\nabla}\boldsymbol{\nabla}\beta u(\mathbf{r},\mathbf{r}')\right)\cdot\boldsymbol{\nabla}\frac{\rho^{(2)}(\mathbf{r},\mathbf{r}')}{\rho(\mathbf{r})}
\tag{C.8}
$$

to recover up to the last term the first member of the YGB-like hierarchy stated in equation (21) of the main text. Taking into account the definition $\mathcal{D}_\mathrm{a}\propto\tau$ 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 [21]. The demonstrated equivalence of equations (9) and (12) in the low-activity limit is, however, not obvious and breaks down when higher-order terms in τ are included.

Appendix D. Integration of the second equality in equation (12)

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

$$
\begin{aligned}
0 &= \boldsymbol{\nabla}\rho(\mathbf{r}) + \rho(\mathbf{r})\boldsymbol{\nabla}\beta\nu^{\mathrm{eff}}(\mathbf{r}) + \int \mathrm{d}\mathbf{r}'\rho^{(2)}(\mathbf{r},\mathbf{r}')\boldsymbol{\nabla}\beta u^{\mathrm{eff}}(\mathbf{r},\mathbf{r}') \\
&= \boldsymbol{\nabla}\rho(\mathbf{r}) + \rho(\mathbf{r})\left(\mathcal{D}_{[1]}^{-1}(\mathbf{r})\boldsymbol{\nabla}\beta\nu(r) - \boldsymbol{\nabla}\ln|\det\mathcal{D}_{[1]}(\mathbf{r})|\right) \\
+ \left[\int \mathrm{d}\mathbf{r}'\rho^{(2)}(\mathbf{r},\mathbf{r}')\left(\mathcal{D}_{11}^{-1}(\mathbf{r},\mathbf{r}')\cdot\boldsymbol{\nabla}\beta u(\mathbf{r},\mathbf{r}') - \boldsymbol{\nabla}\ln|\det\mathcal{D}_{11}(\mathbf{r},\mathbf{r}')|\right)\right]_{\nu=0}
\end{aligned}
\tag{D.1}
$$

where in the second step we have inserted the effective external (15) and pair potential (14). To be consistent with appendix C we used the diagonal form of the latter. In the absence of interparticle interactions, $u(r) = 0$, it is easy to verify that both equations (D.1) and (C.4) simplify to the same equality

$$
\boldsymbol{\nabla}\rho(\mathbf{r}) + \rho(\mathbf{r})\left(\frac{(\boldsymbol{\nabla}\beta\nu(\mathbf{r}))\cdot(\mathbf{1}+\tilde{\tau}\boldsymbol{\nabla}\boldsymbol{\nabla}\nu(\mathbf{r}))}{1+\mathcal{D}_{\mathrm{a}}} - \boldsymbol{\nabla}\ln|\det(\mathbf{1}+\tilde{\tau}\boldsymbol{\nabla}\boldsymbol{\nabla}\nu(\mathbf{r}))|\right) = 0, \tag{D.2}
$$

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 (D.1) in the interacting case to equation (C.5), we notice that the EPA ignores the cross terms proportional to $\tilde{\tau}\boldsymbol{\nabla}\boldsymbol{\nabla}\nu(\mathbf{r})\cdot\int \mathrm{d}\mathbf{r}'\rho^{(2)}(\mathbf{r},\mathbf{r}')\boldsymbol{\nabla}\beta u(\mathbf{r},\mathbf{r}')$ and $\tilde{\tau}\boldsymbol{\nabla}\beta\nu(\mathbf{r})\cdot\int \mathrm{d}\mathbf{r}'\rho^{(2)}(\mathbf{r},\mathbf{r}')\boldsymbol{\nabla}\boldsymbol{\nabla}u(\mathbf{r},\mathbf{r}')$, 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 [23]. This amendable difference aside, we now discuss the bulk system.

Setting $\nu(\mathbf{r}) = 0$ in equation (D.1) becomes

$$
\boldsymbol{\nabla}\rho(\mathbf{r}) + \int \mathrm{d}\mathbf{r}'\rho^{(2)}(\mathbf{r},\mathbf{r}')\left(\frac{(\boldsymbol{\nabla}\beta u(\mathbf{r},\mathbf{r}'))\cdot(\mathbf{1}+\tilde{\tau}\boldsymbol{\nabla}\boldsymbol{\nabla}u(\mathbf{r},\mathbf{r}'))}{1+\mathcal{D}_{\mathrm{a}}} - \boldsymbol{\nabla}\ln|\det(\mathbf{1}+\tilde{\tau}\boldsymbol{\nabla}\boldsymbol{\nabla}u(\mathbf{r},\mathbf{r}'))|\right) = 0 \tag{D.3}
$$

This result amounts to setting

$$
\frac{\tilde{\tau}}{1+\mathcal{D}_{\mathrm{a}}}\int \mathrm{d}\mathbf{r}''\,(\boldsymbol{\nabla}\beta u(\mathbf{r},\mathbf{r}''))\cdot\int \mathrm{d}\mathbf{r}'\rho^{(3)}(\mathbf{r},\mathbf{r}',\mathbf{r}'')\boldsymbol{\nabla}\boldsymbol{\nabla}u(\mathbf{r},\mathbf{r}') \to 0, \tag{D.4}
$$

$$
N\int \mathrm{d}\mathbf{r}_2\ldots\mathrm{d}\mathbf{r}_N P_N(\mathbf{r}^N)\boldsymbol{\nabla}\ln\left|\det\left(\mathbf{1}+\tilde{\tau}\boldsymbol{\nabla}\boldsymbol{\nabla}\sum_{l>1}u(\mathbf{r},\mathbf{r}_l)\right)\right| \to \int \mathrm{d}\mathbf{r}'\rho^{(2)}(\mathbf{r},\mathbf{r}')\boldsymbol{\nabla}\ln|\det(\mathbf{1}+\tilde{\tau}\boldsymbol{\nabla}\boldsymbol{\nabla}u(\mathbf{r},\mathbf{r}'))| \tag{D.5}
$$

in (C.4), which is a logical consequence of the higher-order correlations being ignored. Restricting ourselves to the leading order in τ both sides of (D.5) reduce to the equivalent form $\tilde{\tau}\int \mathrm{d}\mathbf{r}'\rho^{(2)}(\mathbf{r},\mathbf{r}')\boldsymbol{\nabla}\cdot\boldsymbol{\nabla}\boldsymbol{\nabla}u(\mathbf{r},\mathbf{r}')$. In contrast, making use of (C.2), the approximation (D.4) to the bare force term is equivalent to setting

$$
\begin{aligned}
&\frac{\tilde{\tau}}{1+\mathcal{D}_{\mathrm{a}}}\int \mathrm{d}\mathbf{r}''\frac{\rho^{(2)}(\mathbf{r},\mathbf{r}'')}{\rho(\mathbf{r})}\boldsymbol{\nabla}\boldsymbol{\nabla}u(\mathbf{r},\mathbf{r}'')\cdot\int \mathrm{d}\mathbf{r}'\rho^{(2)}(\mathbf{r},\mathbf{r}')\boldsymbol{\nabla}\beta u(\mathbf{r},\mathbf{r}') \\
&+\frac{\tilde{\tau}}{1+\mathcal{D}_{\mathrm{a}}}\int \mathrm{d}\mathbf{r}'\,(\boldsymbol{\nabla}\boldsymbol{\nabla}u(\mathbf{r},\mathbf{r}'))\,\rho(\mathbf{r})\cdot\boldsymbol{\nabla}\frac{\rho^{(2)}(\mathbf{r},\mathbf{r}')}{\rho(\mathbf{r})} \\
&\to\frac{\tilde{\tau}}{1+\mathcal{D}_{\mathrm{a}}}\int \mathrm{d}\mathbf{r}'\rho^{(2)}(\mathbf{r},\mathbf{r}')(\boldsymbol{\nabla}\beta u(\mathbf{r},\mathbf{r}'))\cdot(\boldsymbol{\nabla}\boldsymbol{\nabla}u(\mathbf{r},\mathbf{r}'))
\end{aligned}
\tag{D.6}
$$

in (C.5). Thus the factorization in the first line of (C.5) is not possible, such that, upon multiplying with $\boldsymbol{\mathcal{D}}_{\mathrm{I}}$ defined in equation (C.6) there remains an additional term

$$\tilde{\tau}\int \mathrm{d}\mathbf{r}'\rho^{(2)}(\mathbf{r},\mathbf{r}')(\boldsymbol{\nabla}\beta u(\mathbf{r},\mathbf{r}'))\cdot\left(\boldsymbol{\nabla}\boldsymbol{\nabla}u(\mathbf{r},\mathbf{r}')-\int \mathrm{d}\mathbf{r}''\frac{\rho^{(2)}(\mathbf{r},\mathbf{r}'')}{\rho(\mathbf{r})}\boldsymbol{\nabla}\boldsymbol{\nabla}u(\mathbf{r},\mathbf{r}'')\right) \tag{D.7}$$

proportional to τ on the left-hand-side of equation (C.7). This means that the EPA introduces a three-body term to the YBG-like hierarchy (21). The reason for this discrepancy is that the effective diffusion tensor $\boldsymbol{\mathcal{D}}_{\mathrm{I}}$ is defined independently of the approximation made in equation (D.4).

Finally, we note that if we employ in equation (D.1) the effective pair potential (13) that does not correspond to a diagonal diffusion tensor, we will have to modify equation (D.3) by setting $\tau \to 2\tau$. This, in general, reflects the inconsistency between the low-activity expansions of the two versions of the steady-state condition given by equations (9) and (12), which is not a consequence of approximating the effective force in the second form using pair potentials. In particular, we would also have to substitute $\tau \to 2\tau$ in equation (C.7), as both sides in equation (D.5) are equivalent at the linear order in τ. However, in equation (D.7) only the first term should then be multiplied by the factor two, as the second term arises from $\boldsymbol{\mathcal{D}}_{\mathrm{I}}$ and not from the effective potential.

References

[1] Elgeti J, Winkler R G and Gompper G 2015 *Rep. Prog. Phys.* **78** 056601
[2] Kantsler V *et al* 2013 *Proc. Natl Acad. Sci. USA* **110** 1187
[3] Xiao S *et al* 2014 *J. Chem. Phys.* **141** 184902
[4] Yang X, Manning M L and Marchetti M C 2014 *Soft Matter* **10** 6477
[5] Ni R, Cohen Stuart M A and Bolhuis P G 2015 *Phys. Rev. Lett.* **114** 018302
[6] Cates M E and Tailleur J 2015 *Annu. Rev. Condens. Matter Phys.* **6** 219
[7] Cates M E, Marenduzzo D, Pagonabarraga I and Tailleur J 2010 *Proc. Natl Acad. Sci. USA* **107** 11715
[8] Fily Y and Marchetti M C 2012 *Phys. Rev. Lett.* **108** 235702
[9] Speck T, Bialké J, Menzel A M and Löwen H 2014 *Phys. Rev. Lett.* **112** 218304
[10] Gonnella G, Marenduzzo D, Suma A and Tiribocchi A 2015 *C. R. Phys.* **16** 316
[11] Sharma A and Brader J M 2016 *J. Chem. Phys.* **145** 161101
[12] Farage T F F, Krinninger P and Brader J M 2015 *Phys. Rev.* E **91** 042310
[13] Jung P and Hänggi P 1987 *Phys. Rev.* A **35** 4464
[14] Hänggi P and Jung P 1995 *Adv. Chem. Phys.* **89** 239
[15] Fox R F 1986 *Phys. Rev.* A **33** 467
[16] Fox R F 1986 *Phys. Rev.* A **34** 4525
[17] Grigolini P 1986 *Phys. Lett.* A **119** 157
[18] Faetti S, Fronzoni L, Grigolini P and Mannella R 1988 *J. Stat. Phys.* **52** 951
[19] Hasegawa H 2007 *Physica* A **384** 241
[20] Maggi C *et al* 2015 *Sci. Rep.* **5** 10742
[21] Marconi U M B and Maggi C 2015 *Soft Matter* **11** 8768
[22] Marconi U M B, Maggi C and Melchionna S 2016 *Soft Matter* **12** 5727
[23] Marconi U M B, Paoluzzi M and Maggi C 2016 *Mol. Phys.* **114** 2400
[24] Marconi U M B, Gnan N, Paoluzzi M, Maggi C and Di Leonardo R 2015 *Sci. Rep.* **6** 23297
[25] Wittmann R and Brader J M 2016 *Europhys. Lett.* **114** 68004
[26] Sharma A, Wittmann R and Brader J M 2017 *Phys. Rev.* E **95** 012115
[27] Schwarz-Linek J *et al* 2012 *Proc. Natl Acad. Sci. USA* **109** 4052
[28] Tailleur J and Cates M E 2008 *Phys. Rev. Lett.* **100** 218103
[29] Cates M E and Tailleur J 2013 *Europhys. Lett.* **101** 20010
[30] Rein M and Speck T 2016 *Eur. Phys. J.* E **39** 84
[31] Pototsky A and Stark H 2012 *Europhys. Lett.* **98** 50004
[32] Stenhammar J *et al* 2014 *Soft Matter* **10** 1489

[33] Solon A P, Cates M E and Tailleur J 2015 *Eur. Phys. J. Spec. Top.* **224** 1231
[34] Szamel G 2014 *Phys. Rev.* E **90** 012111
[35] Louis A A, Bolhuis P G and Hansen J P 2000 *Phys. Rev.* E **62** 7961
[36] Lang A *et al* 2000 *J. Phys.: Condens. Matter* **12** 5087
[37] Fodor E *et al* 2016 *Phys. Rev. Lett.* **117** 038103
[38] Rex M and Löwen H 2009 *Eur. Phys. J.* E **28** 139