Ensemble average and nearest particle statistics in disperse multiphase flows

Duan Z. Zhang†

Theoretical Division, Fluid Dynamics and Solid Mechanics Group, Los Alamos National Laboratory, T–3, Los Alamos, NM 87545, USA

A relation between the ensemble average and the nearest particle statistics is derived. The relation accounts for interactions among all the particles, not only the nearest one, and can be used to study long-range particle interactions without the difficulty of a divergent integral. As an example, this relation is applied to calculate the particle sedimentation velocity to the first order of the particle volume fraction. Using the relation, an important particle–fluid–particle stress is introduced for general multiphase flows.

Key words: multiphase flow, particle/fluid flow

---

1. Introduction

Particulate systems involving long-range interactions are difficult to study, especially with the presence of inhomogeneity. In the field of multiphase flows, numerical simulations have been performed by many researchers (e.g. Patankar & Joseph 2001; Zhang & Prosperetti 2005; Wang *et al.* 2016; Subramanian & Balachandar 2017; Theofanous & Chang 2018; Yao & Capecelatro 2018). Most, if not all, of them are limited to homogeneous systems. The assumption of statistical homogeneity is needed in many numerical calculations not only for convenience in the post-processing of the numerical results, but for the validity of the numerical methods in dealing with the long-range interactions among the particles. For instance, the use of periodic boundary conditions in numerical simulations and the treatment (O'Brien 1979) of the surface integral at infinity in Stokesian dynamics (Brady & Bossis 1988) all require the absence of volume fraction gradient because of the long-range hydrodynamic interactions. Similar situations also exist in other fields of physics. For instance, numerical simulations of charged particles often rely on the Ewald sum (Nestler, Pippig & Pott 2015; Rackers *et al.* 2018; Yao & Capecelatro 2018) in a periodic domain and the assumption of charge neutrality to avoid divergent integrals caused by Coulomb forces among the particles. To study an inhomogeneous particulate system, one inevitably needs to consider interactions between the macroscopic length scale at which the mean fields change and the scale of mean particle separation. The long-range interactions cannot be categorized into either of the scales. This difficulty is reflected in the calculation of mean properties, where divergent integrals are

† Email address for correspondence: dzhang@lanl.gov

often encountered. It will be useful to have a method capable of calculating averages without the difficulties of the divergent integrals.

The main objective of this work is to rigorously establish a relation between the ensemble average and the nearest particle statistics. Because of the rapid far-field decay of the nearest particle probability density function, the integral relating the ensemble average and the average conditional on the nearest particle converges absolutely; therefore it is potentially a powerful method to study long-range particle interactions without the difficulty of a divergent integral, especially in processing data from numerical simulations.

The nearest particle probability is an old concept (Hertz 1909; Chandrasekhar 1943). Its use to study effective properties in multiphase flows has also been proposed (Drew & Mandyam 1997). Without a rigorously derived relation between the ensemble average and the nearest particle statistics, there are many questions regarding the legitimacy of its use. These questions include the effect of other nearby particles and the order of the asymptotes of the average quantities at the limit of a small particle volume fraction θ_p. For instance, the typical distance between particles scales as $\theta_p^{-1/3}$. For a method based on the interparticle distance, does this imply that the primary effect of the volume fraction to an average quantity, such as the drag force, is of order $\theta_p^{1/3}$? For a particle, between the typical distance to its nearest particle and the distance twice of that, often there are ten or more particles. How can one account for the effects of these particles?

These questions will be answered at the end of the next section after deriving the relation between the ensemble phase average and the nearest particle statistics and by the example in § 3. As an application of the relation, in § 3, the Stokes drag and the particle sedimentation velocity are calculated following the steps of Batchelor (1972) for a flow with a dilute particle phase, but without using the renormalization techniques. In this way the new method is compared with the renormalization method.

Although the particle sedimentation velocity is a half-century-old problem, and the theory of multiphase flows has progressed considerably beyond the work of Batchelor (1972). For instance, it is now known that the relative motion between the particle and the fluid phases results in microstructures of the particles affecting the sedimentation velocity (Ham & Homsy 1988; Cichocki & Sadlej 2005; Felderhof 2008). The divergent nature associated with long-range particle interactions and their treatments is still a subject of modern research (Piazza 2014; Rackers *et al.* 2018). The work of Batchelor (1972) is still viewed as a significant reference point because of the physical insights, despite the simplification assumptions used. We revisit this sedimentation problem with the new relation to understand the physics contained in the quantities conditionally averaged on the nearest particle. In the process, we also learn how the new relation avoids the need for the difficult renormalizations, and how the effects of the particles other than the nearest one are included. For this purpose, before applying the relation to more complex realistic systems (Fiore, Wang & Swan 2018), we choose to start with the simplified system studied by Batchelor (1972), assuming a statistically homogeneous and isotropic particle distribution.

In § 4, we show that the correlation product of the nearest particle distance and the fluid force on a particle is a stress tensor containing information of particle interactions at the interparticle length scale. This stress is similar to the potential part of the virial stress in a molecular system. Its divergence is a force density important for statistically inhomogeneous flows. This application shows that the nearest particle statistics cannot only recover classical results, but can also be used to study new physics. Furthermore, the renormalization technique is limited to linear problems, while the new relation is rather general. Without the divergence difficulty, the nearest particle statistics can be used to

study a large variety of long-range particle interactions, such as charged particles in plasma and colloidal systems (Allen & Tildesley 2017) and gravitational effects on galaxies.

2. Ensemble phase average and nearest particle statistics

2.1. *Average for the continuous phase*

Let $\mathscr{P}$ be the probability measure (Ash 1972) defined on subsets of a collection (ensemble, or sample space) of disperse multiphase flows, and $\mathscr{F}$ be a multiphase flow in the sample space, which is often represented by a list of variables, including the fields of particle positions, shapes, velocities, boundary values, etc., that uniquely and completely describes the multiphase flow. To allow for generality while simplifying our notation, here we use the abstract concept (Ash 1972) of probability $\mathscr{P}$. If flow $\mathscr{F}$ can be represented by a finite set of parameters (λ), then $\mathscr{F} = \{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ and $\mathrm{d}\mathscr{P} = P(\lambda_1, \lambda_2, \ldots, \lambda_n) \, \mathrm{d}\lambda_1 \, \mathrm{d}\lambda_2 \ldots \mathrm{d}\lambda_n$, with $P(\lambda_1, \lambda_2, \ldots, \lambda_n)$ being the probability density in the n-dimensional phase space formed by the parameters. Unfortunately, most multiphase flows cannot be described simply by a finite number of parameters because of the degrees of freedom associated with the continuous phase. The only exceptions are the Stokes and potential flows with boundary conditions specified without uncertainties.

The indicator function $\chi_c(\boldsymbol{x}, t, \mathscr{F})$ for the continuous phase is defined such that $\chi_c(\boldsymbol{x}, t, \mathscr{F}) = 1$, if at time t position $\boldsymbol{x}$ is occupied by the continuous phase in flow $\mathscr{F}$, and $\chi_c(\boldsymbol{x}, t, \mathscr{F}) = 0$, otherwise. The volume fraction of the continuous phase is defined as (Zhang & Prosperetti 1994, 1997; Zhang *et al.* 2007)

$$\theta_c(\boldsymbol{x}, t) = \int \chi_c(\boldsymbol{x}, t, \mathscr{F}) \, \mathrm{d}\mathscr{P}. \tag{2.1}$$

For a generic continuous phase quantity $q_c(\boldsymbol{x}, t, \mathscr{F})$ in flow $\mathscr{F}$ at position $\boldsymbol{x}$ and time t, its ensemble phase average is defined as

$$\langle q_c \rangle(\boldsymbol{x}, t) = \frac{1}{\theta_c(\boldsymbol{x}, t)} \int \chi_c(\boldsymbol{x}, t, \mathscr{F}) q_c(\boldsymbol{x}, t, \mathscr{F}) \, \mathrm{d}\mathscr{P}. \tag{2.2}$$

These definitions are extensions of the ensemble average of Batchelor (1972). For Stokes flows containing N identical particles, we can write $\mathrm{d}\mathscr{P} = P(\mathscr{C}_N) \, \mathrm{d}\mathscr{C}_N / N!$ if the notation of Batchelor (1972) is used, where $\mathscr{C}_N$ is the particle configuration. Besides the advantage of simplified notations, direct use of the probability $\mathscr{P}$ implies that our results are independent of the description for the physical system.

To consider the nearest particle statistics, we introduce a function,

$$h_i(\boldsymbol{x}, t, \mathscr{F}) = \frac{1}{N_x(\boldsymbol{x}, t, \mathscr{F})} \prod_j H(|\boldsymbol{\xi}_j(t, \mathscr{F}) - \boldsymbol{x}| - |\boldsymbol{\xi}_i(t, \mathscr{F}) - \boldsymbol{x}|), \tag{2.3}$$

where $\boldsymbol{\xi}_i(t, \mathscr{F})$ is the location of the i-th particle centre in flow $\mathscr{F}$ at time t, $H(\cdot)$ is the Heaviside function ($H(x) = 0$ for $x < 0$; $H(x) = 1$ for $x \geqslant 0$) and

$$N_x(\boldsymbol{x}, t, \mathscr{F}) = \sum_i \prod_j H(|\boldsymbol{\xi}_j(t, \mathscr{F}) - \boldsymbol{x}| - |\boldsymbol{\xi}_i(t, \mathscr{F}) - \boldsymbol{x}|) \geqslant 1. \tag{2.4}$$

Indices i and j in both (2.3) and (2.4) run through all the particles in flow $\mathscr{F}$. The products of the Heaviside functions in (2.3) and (2.4) are unity if and only if particle i is the

nearest (from the particle centre $\boldsymbol{\xi}_i$) to spatial position $\boldsymbol{x}$, and are zero otherwise. Function $h_i(\boldsymbol{x}, t, \mathscr{F})$ is non-zero if and only if particle i is one of the nearest particles to $\boldsymbol{x}$. In (2.4), N_x is the number of the nearest particles to a spatial point $\boldsymbol{x}$ at time t in flow $\mathscr{F}$. Here $N_x(\boldsymbol{x}, t, \mathscr{F}) \geqslant 1$, because for a given position $\boldsymbol{x}$ in a disperse multiphase flow, there is always at least one the nearest particle. For a dispersion of randomly placed particles, $N_x = 1$ in almost all flows and all $\boldsymbol{x}$ and t. With a probability of zero, a position $\boldsymbol{x}$ has two or more nearest particles at the same distance away. For ordered systems, N_x can be greater than 1. One example is a simple cubic array of particles. If $\boldsymbol{x}$ is the centre of the cube, $N_x = 8$.

For a spatial point $\boldsymbol{y}$ using the property $\int \delta[\boldsymbol{\xi}_i(t, \mathscr{F}) - \boldsymbol{y}] \, d^3 y = 1$ of the δ-function, and then exchanging the order of the integral and the summation over i, we have

$$\int \sum_i h_i(\boldsymbol{x}, t, \mathscr{F}) \delta[\boldsymbol{\xi}_i(t, \mathscr{F}) - \boldsymbol{y}] \, d^3 y = \sum_i h_i(\boldsymbol{x}, t, \mathscr{F}) = \frac{N_x(\boldsymbol{x}, t, \mathscr{F})}{N_x(\boldsymbol{x}, t, \mathscr{F})} = 1. \quad (2.5)$$

The mathematical steps in (2.5) may appear trivial, but the left-hand side provides a useful way of performing the integral. For a given position $\boldsymbol{x}$ and flow $\mathscr{F}$, one can choose any position $\boldsymbol{y}$, surround it with an infinitesimal differential volume (with the value of the volume) $d^3 y$, and calculate $\sum_i h_i(\boldsymbol{x}, t, \mathscr{F}) \delta[\boldsymbol{\xi}_i(t, \mathscr{F}) - \boldsymbol{y}] \, d^3 y$. For most of the choices of $\boldsymbol{y}$, the value is zero. The non-zero value appears only when the differential volume around $\boldsymbol{y}$ contains one of the nearest particle centres ($\boldsymbol{\xi}_i$) to $\boldsymbol{x}$ in the flow $\mathscr{F}$, otherwise $h_i(\boldsymbol{x}, t, \mathscr{F}) = 0$. There are only $N_x(\boldsymbol{x}, t, \mathscr{F})$ of them. The key point here is that one can freely choose a position $\boldsymbol{y}$ first and then ask whether this position is the centre of the nearest particle to $\boldsymbol{x}$ in the flow. Relation (2.5) states that after integrating contributions from all $\boldsymbol{y}$ in the entire space, every flow is accounted for 100 % and only 100 %. This property ensures that the nearest statistics developed in the following does not miscount or overcount any flow in the ensemble.

Multiplying the left-hand side of (2.5) with the right-hand side of (2.2) and then exchanging the order of the integrations, one finds a major conclusion of this work, the relation between the ensemble phase average and the nearest particle statistics,

$$\langle q_c \rangle(\boldsymbol{x}, t) = \int \langle q_c \rangle_{nst}(\boldsymbol{x}, \boldsymbol{y}, t) P_{nst}^c(\boldsymbol{y}|\boldsymbol{x}, t) \, d^3 y, \quad (2.6)$$

where

$$\langle q_c \rangle_{nst}(\boldsymbol{x}, \boldsymbol{y}, t) = \frac{1}{\theta_c(\boldsymbol{x}, t) P_{nst}^c(\boldsymbol{y}|\boldsymbol{x}, t)} \int q_c(\boldsymbol{x}, t, \mathscr{F}) \chi_c(\boldsymbol{x}, t, \mathscr{F})$$

$$\sum_i h_i(\boldsymbol{x}, t, \mathscr{F}) \delta[\boldsymbol{\xi}_i(t, \mathscr{F}) - \boldsymbol{y}] \, d\mathscr{P} \quad (2.7)$$

and

$$P_{nst}^c(\boldsymbol{y}|\boldsymbol{x}, t) = \frac{1}{\theta_c(\boldsymbol{x}, t)} \int \chi_c(\boldsymbol{x}, t, \mathscr{F}) \sum_{i=1} h_i(\boldsymbol{x}, t, \mathscr{F}) \delta[\boldsymbol{\xi}_i(t, \mathscr{F}) - \boldsymbol{y}] \, d\mathscr{P}, \quad (2.8)$$

defined such that when $q_c(\boldsymbol{x}, t, \mathscr{F}) = 1$, $\langle q_c \rangle_{nst}(\boldsymbol{x}, \boldsymbol{y}, t) = 1$.

Similar to the discussion above, in (2.7) and (2.8) for a chosen $\boldsymbol{y}$, the δ-function $\delta[\boldsymbol{\xi}_i(t, \mathscr{F}) - \boldsymbol{y}]$ restricts that the contributions are only from the flows in which there is a particle centred at position $\boldsymbol{y}$ at time t. Function $h_i(\boldsymbol{x}, t, \mathscr{F})$ in the integrals then

ensures that the particle at y, which is particle i, is the nearest to spatial point x at time t. Finally, the indicator functions $\chi_c(x, t, \mathscr{F})$ restrict that the contributions to the integrals only come from the flows in which position x is occupied by the continuous phase at time t. Since in the sense of ensemble average, the volume fraction $\theta_c(x, t)$ defined in (2.1) is the probability of finding position x being occupied by the continuous phase, the quantity defined in (2.8) is then the probability density of finding the nearest particle to x at position y, conditional on position x being occupied by the continuous phase at time t. Similarly, the quantity defined in (2.7) is the continuous phase quantity q_c at position x and time t averaged conditionally on the particle at y being the nearest particle to x. It is important to note that $q_c(x, t, \mathscr{F})$ in definition (2.7) is the quantity at position x in flow $\mathscr{F}$ at time t, containing effects from all the particles, including, but not only, the nearest one centred at y. Furthermore, using (2.1) and (2.5) in (2.8), one finds

$$\int P_{nst}^c(y|x, t) \, \mathrm{d}^3 y = 1, \tag{2.9}$$

as required for a probability density. This normalization property of P_{nst}^c implies that for any bounded function $\langle q_c \rangle_{nst}$, the ensemble phase average calculated from integral (2.6) converges absolutely since $P_{nst}^c(y|x, t) \geqslant 0$ as defined in (2.8). We are then free of the difficulties of divergent integrals, even for long-range particle interactions.

Relation (2.6) states that the ensemble average of a continuous phase quantity at position x and time t can be calculated in two steps. The first step is to average over all the flows in which the nearest particle is centred at position y. This average is $\langle q_c \rangle_{nst}$ defined in (2.7), containing effects from all the particles in the flow. The second step is to average over all possible positions of the nearest particle centres (y). The ensemble phase average is then the expected value of $\langle q_c \rangle_{nst}$ calculated using (2.6) based on the probability density $P_{nst}^c(y|x, t)$ of the nearest particle.

2.2. *Average for the particle phase*

The simple derivation above can also be performed for quantities associated with particles. We now list the steps in this subsection. In an ensemble average, the particle number density is defined as (Irving & Kirkwood 1950; Zhang & Prosperetti 1994; Zhang, Ma & Rauenzahn 2006)

$$n_p(x, t) = \int \sum_i \delta[\xi_i(t, \mathscr{F}) - x] \, \mathrm{d}\mathscr{P}. \tag{2.10}$$

The ensemble average for a particle quantity $q_i(t, \mathscr{F})$ is defined as

$$\bar{q}_p(x) = \frac{1}{n_p(x)} \int \sum_i q_i(t, \mathscr{F}) \delta[\xi_i(t, \mathscr{F}) - x] \, \mathrm{d}\mathscr{P}. \tag{2.11}$$

Similar to the procedures for the continuous phase above, to relate the ensemble average to the nearest particle statistics, we introduce a function

$$h_{ij}(t, \mathscr{F}) = \frac{1}{N_i(t, \mathscr{F})} \prod_{k \neq i, k \neq j} H\left[|\xi_k(t, \mathscr{F}) - \xi_i(t, \mathscr{F})| - |\xi_j(t, \mathscr{F}) - \xi_i(t, \mathscr{F})|\right], \tag{2.12}$$

where

$$N_i(t, \mathscr{F}) = \sum_{j \neq i} \prod_{k \neq i, k \neq j} H\left[|\xi_k(t, \mathscr{F}) - \xi_i(t, \mathscr{F})| - |\xi_j(t, \mathscr{F}) - \xi_i(t, \mathscr{F})|\right] \geqslant 1, \tag{2.13}$$

is the number of the nearest particles to particle i. The products in (2.12) and (2.13) are unity if and only if particle j is one of the nearest particles to particle i. Otherwise the products are zero. Similar to (2.5), using the property of δ-function, for any flow $\mathscr{F}$ and particle i we have

$$\int \sum_{j \neq i} \delta[\mathbf{y} - \boldsymbol{\xi}_j(t, \mathscr{F})]h_{ij}(t, \mathscr{F})\, \mathrm{d}^3 y = \sum_{j \neq i} h_{ij}(t, \mathscr{F}) = \frac{N_i}{N_i} = 1. \qquad (2.14)$$

Multiplying the left-hand side of (2.14), with the right-hand side of (2.11) and then exchanging the order of the integrations, similar to (2.6) one finds

$$\bar{q}_p(\mathbf{x}, t) = \int \bar{q}_{p,nst}(\mathbf{x}, \mathbf{y}, t)P^p_{nst}(\mathbf{y}|\mathbf{x}, t)\, \mathrm{d}^3 y, \qquad (2.15)$$

where

$$P^p_{nst}(\mathbf{y}|\mathbf{x}, t) = \frac{1}{n_p(\mathbf{x}, t)} \int \sum_i \delta[\boldsymbol{\xi}_i(t, \mathscr{F}) - \mathbf{x}] \sum_{j \neq i} \delta[\boldsymbol{\xi}_j(t, \mathscr{F}) - \mathbf{y}]h_{ij}(t, \mathscr{F})\, \mathrm{d}\mathscr{P}, \quad (2.16)$$

is the probability density of finding the nearest particle at $\mathbf{y}$ given a particle already in $\mathbf{x}$, and

$$\bar{q}_{p,nst}(\mathbf{x}, \mathbf{y}, t) = \frac{1}{n_p(\mathbf{x}, t)P^p_{nst}(\mathbf{y}|\mathbf{x}, t)} \int \sum_i q_i(t, \mathscr{F})\delta[\boldsymbol{\xi}_i(t, \mathscr{F}) - \mathbf{x}]$$

$$\sum_{j \neq i} h_{ij}(t, \mathscr{F})\delta[\boldsymbol{\xi}_j(t, \mathscr{F}) - \mathbf{y}]\, \mathrm{d}\mathscr{P} \qquad (2.17)$$

is the particle value q on the particle at $\mathbf{x}$ averaged conditionally on the particle at $\mathbf{y}$ being the nearest to the particle at $\mathbf{x}$. Again, this conditionally averaged quantity includes effects from all particles in the flow, not only the pair interaction between the nearest neighbours. Similar to (2.9), integrating (2.16) over $\mathbf{y}$, and noting (2.14) and (2.10), we have

$$\int P^p_{nst}(\mathbf{y}|\mathbf{x}, t)\, \mathrm{d}^3 y = 1, \qquad (2.18)$$

as required for a probability density function. Similar to the ensemble average for the continuous phase, this relation implies the absolute convergence of integral (2.15) for any bounded $\bar{q}_{p,nst}$.

In an ensemble average, the pair distribution function $P_2(\mathbf{x}, \mathbf{y}, t)$ can be expressed as

$$P_2(\mathbf{x}, \mathbf{y}, t) = \int \sum_i \delta[\boldsymbol{\xi}_i(t, \mathscr{F}) - \mathbf{x}] \sum_{j \neq i} \delta[\boldsymbol{\xi}_j(t, \mathscr{F}) - \mathbf{y}]\, \mathrm{d}\mathscr{P}. \qquad (2.19)$$

Let $q_{ij}(t, \mathscr{F})$ be a quantity pertaining to a pair of particles i and j. For pairs located at $\mathbf{x}$ and $\mathbf{y}$, the average of the pair quantity can be defined as

$$\bar{q}_2(\mathbf{x}, \mathbf{y}, t) = \frac{1}{P_2(\mathbf{x}, \mathbf{y}, t)} \int \sum_i \delta[\boldsymbol{\xi}_i(t, \mathscr{F}) - \mathbf{x}] \sum_{j \neq i} \delta[\boldsymbol{\xi}_j(t, \mathscr{F}) - \mathbf{y}]q_{ij}(t, \mathscr{F})\, \mathrm{d}\mathscr{P}.$$

$$(2.20)$$

This is an average conditional on a pair of particles located at $\mathbf{x}$ and $\mathbf{y}$. Letting the pair quantity $q_{ij} = h_{ij}$, denoting the corresponding average $\bar{q}_2$ as $\bar{h}_2$ in (2.20), and comparing

the relation with (2.16) we find

$$\bar{h}_2(\boldsymbol{x}, \boldsymbol{y}, t) = \frac{P_{nst}^p(\boldsymbol{y}|\boldsymbol{x}, t) n_p(\boldsymbol{x}, t)}{P_2(\boldsymbol{x}, \boldsymbol{y}, t)}. \qquad (2.21)$$

For $q_{ij} = h_{ij}$ in (2.20), the integrand is non-zero if and only if $\boldsymbol{x}$ is occupied by a particle (i), $\boldsymbol{y}$ is occupied by another (j), and particle j is the nearest neighbour to particle i, otherwise $h_{ij} = 0$. Such calculated $\bar{h}_2$ is the probability of finding the particle at $\boldsymbol{y}$ being the nearest neighbour to the particle at $\boldsymbol{x}$, conditional on knowing a pair of particles at $\boldsymbol{x}$ and $\boldsymbol{y}$. In other words, under the condition that a particle pair locates at positions $\boldsymbol{x}$ and $\boldsymbol{y}$, $\bar{h}_2(\boldsymbol{x}, \boldsymbol{y}, t)$ is the probability that no particle, other than the one at $\boldsymbol{x}$, is inside the spherical region centred at $\boldsymbol{x}$ with radius $|\boldsymbol{y} - \boldsymbol{x}|$.

The relations and equations shown so far are exact. They only rely on the assumptions that the probability $\mathscr{P}$ exists and that the related functions are well-behaved, so that the order of integrations can be exchanged freely. There is no assumption about the volume fraction or the nature of the particle interactions; therefore, the relations are applicable in general particulate systems.

Probability densities $P_{nst}^c(\boldsymbol{y}|\boldsymbol{x}, t)$ and $P_{nst}^p(\boldsymbol{y}|\boldsymbol{x}, t)$ are different. In the former, the position $\boldsymbol{x}$ is occupied by the continuous phase, and in the latter, the position $\boldsymbol{x}$ is occupied by a particle centre. In the rest of this work, we mainly focus on the spatial part of the particle distributions. For simplicity, we now drop the variable t for time in the rest of the text.

2.3. *Dilute particle phase*

We now use (2.21) to calculate $P_{nst}^p(\boldsymbol{y}|\boldsymbol{x})$. In a dilute dispersion of randomly distributed hard spheres of radii a, correct to the leading order of the particle number density n_p, $P_2(\boldsymbol{x}, \boldsymbol{y}) = 0$, if $r = |\boldsymbol{y} - \boldsymbol{x}| < 2a$, and $P_2(\boldsymbol{x}, \boldsymbol{y}) = n_p^2$, if $r = |\boldsymbol{y} - \boldsymbol{x}| > 2a$ (Batchelor 1972). The probability $\bar{h}_2(\boldsymbol{x}, \boldsymbol{y})$ of no particle centre other than $\boldsymbol{x}$ inside the spherical region $|\boldsymbol{z} - \boldsymbol{x}| < r$ can be calculated as the follows. For hard spheres, the probability is the same as the probability of no particle centre in the spherical shell $2a < |\boldsymbol{z} - \boldsymbol{x}| < r$, since no particle centre other than $\boldsymbol{x}$ can be inside of the spherical region $|\boldsymbol{z} - \boldsymbol{x}| < 2a$ as shown in figure 1. The volume of the shell is $V = 4\pi[r^3 - (2a)^3]/3$. We now divide this volume into N equal small volumes. Each of them has a volume $v = V/N$. The probability of finding a particle centre in such a small volume is $n_p v$, and the probability of no particle centre in it is $1 - n_p v$. For suspensions of randomly distributed particles with a small particle volume fraction θ_p, the probability of particle overlap is of $O(\theta_p^2)$ and can be neglected. Under this assumption, the probability of no particle in all the N small volumes (i.e. the entire shell) is $(1 - n_p v)^N$. As the number N of the small volumes increases, we have

$$\bar{h}_2 = \lim_{N \to \infty} (1 - n_p v)^N = \lim_{N \to \infty} \left(1 - \frac{n_p V}{N}\right)^N = e^{-n_p V}. \qquad (2.22)$$

Substituting this relation into (2.21) we find

$$P_{nst}^p(\boldsymbol{y}|\boldsymbol{x}) = \begin{cases} n_p \exp(-4\pi n_p[r^3 - (2a)^3]/3) & \text{if } r = |\boldsymbol{y} - \boldsymbol{x}| \geqslant 2a; \\ 0 & \text{if } r = |\boldsymbol{y} - \boldsymbol{x}| < 2a. \end{cases} \qquad (2.23)$$

The nearest-neighbour distribution function $H(r)$ of Torquato, Lu & Rubinstein (1990*b*) can be expressed as $4\pi r^2 P_{nst}^p(\boldsymbol{y}|\boldsymbol{x})$. With (2.23), such calculated $H(r)$ agrees with (7) of Torquato *et al.* (1990*b*) at the limit of a small particle volume fraction.

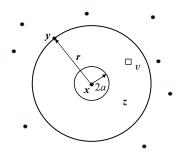FIGURE 1. Calculation of $\bar{h}_2(\boldsymbol{x}, \boldsymbol{y})$. Black dots are particle centres.

In cases where position $\boldsymbol{x}$ is occupied by the continuous phase instead of a particle, particle centres are allowed to be as close as one radius a away from $\boldsymbol{x}$. The radius of the inner spherical region in figure 1 becomes a, and the shell volume becomes $4\pi(r^3 - a^3)/3$. Similar to (2.23), we then have

$$P_{nst}^c(\boldsymbol{y}|\boldsymbol{x}) = \begin{cases} n_p \exp(-4\pi n_p(r^3 - a^3)/3) & \text{if } r = |\boldsymbol{y} - \boldsymbol{x}| \geqslant a; \\ 0 & \text{if } r = |\boldsymbol{y} - \boldsymbol{x}| < a. \end{cases} \tag{2.24}$$

It is easy to verify that probability densities (2.23) and (2.24) satisfy the normalization conditions (2.18) and (2.9).

A relation similar to (2.6) is (2.10) of Batchelor (1972), which can be written in our notations as

$$\langle q_c \rangle(\boldsymbol{x}) = \int_{|\boldsymbol{y} - \boldsymbol{x}| > a} q_c^1(\boldsymbol{x}, \boldsymbol{y}) n_p(\boldsymbol{y}) \, \mathrm{d}^3 y + O(\theta_p^2), \tag{2.25}$$

where q_c^1 is the contribution from a single particle at $\boldsymbol{y}$ to quantity q_c at position $\boldsymbol{x}$. This equation is valid if $q_c^1(\boldsymbol{x}, \boldsymbol{y})$ decays faster than $1/r^{3+\varepsilon}$ ($\varepsilon > 0$). There are conceptual differences between this relation and relation (2.6). In (2.25), $n_p(\boldsymbol{y}) \, \mathrm{d}^3 y$ is the probable number of particles in the volume element $\mathrm{d}^3 y$. The integral sums over individual contributions from all the particles in the flow. An implied assumption in the relation is that the average $\langle q_c \rangle$ can be calculated by adding contributions from all the particles in the flow; therefore, relation (2.25) is only valid for linear problems, if the integral converges. Since in (2.25) q_c^1 is the contribution from a single particle, to calculate the ensemble average, one cannot only account for the contribution from the nearest particle by simply replacing $n_p(\boldsymbol{y})$ with the nearest particle probability density $P_{nst}^c(\boldsymbol{y}|\boldsymbol{x})$. In contrast, the starting point of the nearest particle statistics is that the quantity q_c is well-defined in all the flows, including effects from all the particles and possible boundary and initial conditions. In (2.6) $\langle q_c \rangle_{nst}(\boldsymbol{x}, \boldsymbol{y})$ is the value of q_c conditionally averaged on $\boldsymbol{y}$ being the nearest particle to $\boldsymbol{x}$, as defined in (2.7). In (2.6), $P_{nst}^c \mathrm{d}^3 y$ is the probability of having a nearest particle to $\boldsymbol{x}$ in the volume element $\mathrm{d}^3 y$ around $\boldsymbol{y}$. The integral sums over all such probabilities, instead of over the individual particle contributions as in (2.25).

In the case of a dilute particle phase, particles are far apart. If q_c is the particle contribution to a continuous phase quantity, such as the fluid velocity caused by particle sedimentation, intuitively we have $\langle q_c \rangle_{nst} \approx q_c^1$, but the order of this approximation cannot be determined easily from the intuition. More rigorously, one can obtain $\langle q_c \rangle_{nst}$ by solving the equation conditionally averaged on the nearest particle as done by Hinch (1977) for the equation conditionally averaged on a particle, (not necessarily the nearest one), fixed at a

given position. It is shown in the following section that the solution to the conditionally averaged equation for the fluid velocity can indeed be approximated this way, resulting an error of $O(\theta_p^{1/3})$ in the ensemble average, because of the long-range hydrodynamic effect of the particles.

For short-range interactions, using $\langle q_c \rangle_{nst} = q_c^1 + O(n_p^2)$ in (2.6) and using (2.24), we find

$$\langle q_c \rangle(\boldsymbol{x}) = \int_{|\boldsymbol{y}-\boldsymbol{x}|>a} q_c^1(\boldsymbol{x}, \boldsymbol{y}) n_p(\boldsymbol{y}) \exp(-4\pi n_p (r^3 - a^3)/3) \, \mathrm{d}^3 y + O(\theta_p^2). \qquad (2.26)$$

For cases of small particle volume fractions, n_p is also small and $\exp(-4\pi n_p (r^3 - a^3)/3) \approx 1$. This relation reduces to (2.25), which is the result of Batchelor (1972). In other words, for short-range particle interactions, the new relation (2.6) yields the same first order (in θ_p) result as that from relation (2.10) of Batchelor (1972). For long-range particle interactions, the exponential decay in (2.26) becomes important to ensure the convergence of the integral, while the error from (2.26) is not necessarily $O(\theta_p^2)$. The calculation of the Stokes drag in the next section encounters such an example.

3. Drag in dilute Stokes suspension

Let us consider a dispersion of rigid spherical particles with radii a in a Stokes flow with the fluid viscosity μ. We assume the flow is statistically homogeneous with constant average fluid and particle velocities and with a random particle distribution. This system is studied by Batchelor (1972) using the renormalization technique.

Using the Faxén theorem, following the notation of Batchelor (1972), the drag on a sphere at $\boldsymbol{x}$ moving at velocity $\boldsymbol{v}_p$ in flow $\mathscr{F}$ can be calculated as

$$\boldsymbol{f}_p(\boldsymbol{x}, \mathscr{F}) = -6\pi a\mu[\boldsymbol{v}_p - \boldsymbol{V}'(\boldsymbol{x}, \mathscr{F}) - \boldsymbol{V}''(\boldsymbol{x}, \mathscr{F}) - \boldsymbol{W}(\boldsymbol{x}, \mathscr{F})], \qquad (3.1)$$

where $\boldsymbol{V}'(\boldsymbol{x}, \mathscr{F})$ is the fluid velocity $\boldsymbol{v}_c(\boldsymbol{x}, \mathscr{F}')$ in flow $\mathscr{F}'$ with the particle at $\boldsymbol{x}$ replaced by the fluid while keeping locations of other particles unaltered compared with flow $\mathscr{F}$,

$$\boldsymbol{V}'(\boldsymbol{x}, \mathscr{F}) = \boldsymbol{v}_c(\boldsymbol{x}, \mathscr{F}'), \qquad (3.2)$$

$$\boldsymbol{V}''(\boldsymbol{x}, \mathscr{F}) = \tfrac{1}{6}a^2 \nabla^2 \boldsymbol{V}'(\boldsymbol{x}, \mathscr{F}') = \tfrac{1}{6}a^2 \nabla^2 \boldsymbol{v}_c(\boldsymbol{x}, \mathscr{F}') \qquad (3.3)$$

and $\boldsymbol{W}$ accounts the reflection from the surrounding particles caused by the Stokeslet of the particle at $\boldsymbol{x}$.

The particle drag $\boldsymbol{f}_p$ and velocities in (3.1) can be considered as quantities pertaining to the particle at $\boldsymbol{x}$. By definition (2.11) the ensemble average force on the particle is calculated by averaging over all other particle positions given one of them at $\boldsymbol{x}$ as follows:

$$\bar{\boldsymbol{f}}_p(\boldsymbol{x}) = -6\pi a\mu(\bar{\boldsymbol{v}}_p - \overline{\boldsymbol{V}'} - \overline{\boldsymbol{V}''} - \bar{\boldsymbol{W}}). \qquad (3.4)$$

Using relation (2.15) between the ensemble phase average and the nearest particle statistics, we have

$$\overline{\boldsymbol{V}'} = \int \overline{\boldsymbol{V}'}_{nst}(\boldsymbol{x}, \boldsymbol{y}) P_{nst}^p(\boldsymbol{y}|\boldsymbol{x}) \, \mathrm{d}^3 y \qquad (3.5)$$

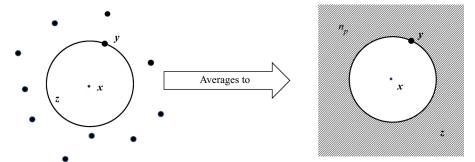and similar relations for velocities $\overline{\boldsymbol{V}''}$ and $\bar{\boldsymbol{W}}$.

FIGURE 2. Effective medium formed by particles outside the particle-free region.

Let $\boldsymbol{y}$ be the centre of the nearest particle to $\boldsymbol{x}$ in both $\mathscr{F}$ and $\mathscr{F}'$, then $|\boldsymbol{y} - \boldsymbol{x}| \geqslant 2a$ because $\boldsymbol{x}$ is occupied by a particle in flow $\mathscr{F}$. Under the assumption (Batchelor 1972) that for flows containing a large number of particles the difference by one particle has little effect on their averages, after averaging over positions of all other particles, we have

$$\overline{V}'_{nst}(\boldsymbol{x}, \boldsymbol{y}) = \langle \boldsymbol{v}_c \rangle_{nst}(\boldsymbol{x}, \boldsymbol{y}), \quad \overline{V}''_{nst}(\boldsymbol{x}, \boldsymbol{y}) = \frac{a^2}{6} \langle \nabla^2 \boldsymbol{v}_c \rangle_{nst}(\boldsymbol{x}, \boldsymbol{y}), \quad |\boldsymbol{y} - \boldsymbol{x}| \geqslant 2a. \tag{3.6a–c}$$

Here, we write $\langle \nabla^2 \boldsymbol{v}_c \rangle_{nst}$ instead of $\nabla^2 \langle \boldsymbol{v}_c \rangle_{nst}$ because the Laplacian is taken in (3.3) before the average. Generally for an ensemble phase average, the differentiation and the average operators do not commutate (Zhang & Prosperetti 1994; Zhang *et al.* 2007). The same is also true for quantities conditionally averaged on the nearest particle as (A 7) in appendix A.

The velocity $\langle \boldsymbol{v}_c \rangle_{nst}(\boldsymbol{x}, \boldsymbol{y})$ can be obtained by averaging over all the configurations as shown in figure 2, in which position $\boldsymbol{x}$ is occupied by the fluid, the nearest particle to $\boldsymbol{x}$ is centred at $\boldsymbol{y}$. After averaging over all such configurations, one finds an effective medium as shown on the right-hand side of the figure. In the effective medium, there is no particle inside the spherical region of radius $r = |\boldsymbol{y} - \boldsymbol{x}|$ centred at $\boldsymbol{x}$. To study the particle drag, an external force field is needed to drive the motion. We now assume the force is gravity. Other forces can be considered similarly. The density of the particle-free region is the fluid density ρ_c. Outside of the particle-free region, the mixture density becomes $\theta_p \rho_p + (1 - \theta_p) \rho_c$, where ρ_p is the particle density. This is similar to the problem of a lighter viscous droplet, the particle-free region, immersed in a heavier fluid outside (assuming $\rho_p > \rho_c$). Far away from the particle-free region, the gravity $[\theta_p \rho_p + (1 - \theta_p) \rho_c] \boldsymbol{g}$ is balanced by the pressure gradient. With the lighter particle-free region, the buoyancy causes a back (against the gravity) flow velocity $\boldsymbol{v}_b$. The total buoyancy is proportional to $\frac{4}{3} \pi r^3 \theta_p \Delta \rho \boldsymbol{g}$, where $\Delta \rho = \rho_p - \rho_c$; while the total drag on the surface of the particle-free region is proportional to $\mu r \boldsymbol{v}_b$. The resulting back flow velocity is $\boldsymbol{v}_b = O(\theta_p r^2 \Delta \rho \boldsymbol{g} / \mu)$. This back flow velocity can also be obtained by the more rigorous averaged equation approach in appendix B.

Let $\boldsymbol{v}_0$ be the velocity of the mixture far away from the spherical particle-free region. Correct to the zeroth order of the particle volume fraction θ_p, the presence of the nearest particle at $\boldsymbol{y}$ induces a fluid velocity

$$\boldsymbol{v}_c^1(\boldsymbol{x}, \boldsymbol{y}) = (\bar{\boldsymbol{v}}_p - \boldsymbol{v}_0 - \boldsymbol{v}_b') \left(\frac{3a}{4r} + \frac{a^3}{4r^3} \right) + \frac{\boldsymbol{r}}{r} \left[\frac{\boldsymbol{r}}{r} \cdot (\bar{\boldsymbol{v}}_p - \boldsymbol{v}_0 - \boldsymbol{v}_b') \right] \left(\frac{3a}{4r} - \frac{3a^3}{4r^3} \right), \tag{3.7}$$

at x, where $r = x - y$, and $v'_b = O(\theta_p r^2)$ is the back flow velocity at y, similar to v_b. Velocity (3.7) is the solution of a sphere in the Stokes flow with the background velocity $v_0 + v'_b$ without consideration of other particles. The fluid velocity at the centre of the particle-free region is then

$$\overline{V}'_{nst}(x, y) = \langle v_c \rangle_{nst}(x, y) = v_0 + v_c^1(x, y) + v_b(x, y). \tag{3.8}$$

Since v'_b appears with the factor a/r in (3.7), it causes an $O(\theta_p a r)$ effect on the velocity $\overline{V}'_{nst}$, which is small compared with $v_b = O(\theta_p r^2)$, and hence negligible as shown below.

A divergent integral is encountered if (3.7) (without v'_b) is used in (2.25) to calculate $\overline{V}'$. The first renormalization method is then used by Batchelor (1972). In the present method, instead of using (2.25), we substitute (3.8) into (2.15) or (3.5) and then use (2.23) to find

$$\overline{V}' = v_0 + \theta_p^{1/3} \Gamma_{inc} \left(\tfrac{2}{3}, 8\theta_p \right) [\bar{v}_p - v_0 + O(\theta_p^{1/3})] + \int v_b(x, y) P_{nst}^p(y|x) \, d^3 y, \tag{3.9}$$

where Γ_{inc} is the incomplete Γ-function, defined as

$$\Gamma_{inc}(s, \eta) = \int_\eta^\infty t^{s-1} e^{-t} \, dt, \tag{3.10}$$

and the $O(\theta_p^{1/3})$ term results from v'_b. By noting $v_b = O(\theta_p r^2)$, and $\int \theta_p r^2 P_{nst}^p \, d^3 y = O(\theta_p^{1/3})$, the last term in (3.9) is of $O(\theta_p^{1/3})$. At this point, it might seem that $\theta_p^{1/3}$ is the leading-order effect on the drag if the nearest particle statistics is used, in contradiction to the result of Batchelor (1972). This issue is resolved by noting that v_0 is not the average fluid velocity. Using (3.8) and (2.24) in (2.6) one finds the average fluid velocity

$$\langle v_c \rangle = v_0 + \theta_p^{1/3} \Gamma_{inc} \left(\tfrac{2}{3}, \theta_p \right) [\bar{v}_p - v_0 + O(\theta_p^{1/3})] + \int v_b(x, y) P_{nst}^c(y|x) \, d^3 y. \tag{3.11}$$

Similar to (3.9), the $O(\theta_p^{1/3})$ term comes from v'_b, and the last term above is of $O(\theta_p^{1/3})$. Subtracting (3.11) from (3.9), we find

$$\begin{aligned} \overline{V}' &= \langle v_c \rangle - \theta_p^{1/3} \left[\Gamma_{inc} \left(\tfrac{2}{3}, \theta_p \right) - \Gamma_{inc} \left(\tfrac{2}{3}, 8\theta_p \right) \right] [\bar{v}_p - v_0 + O(\theta_p^{1/3})] \\ &\quad + \int v_b (P_{nst}^p - P_{nst}^c) \, d^3 y \\ &= \langle v_c \rangle - \tfrac{9}{2} \theta_p (\bar{v}_p - \langle v_c \rangle) + O(\theta_p^{4/3}). \end{aligned} \tag{3.12}$$

The last identity holds because $\Gamma_{inc}(\tfrac{2}{3}, \theta_p) - \Gamma_{inc}(\tfrac{2}{3}, 8\theta_p) \approx \tfrac{9}{2}\theta_p^{2/3}$, and v_0 can be replaced by $\langle v_c \rangle$ after using (3.11). The integral can be calculated in regions with $r = |y - x| < 2a$ and $r > 2a$. In the finite region ($r < 2a$) using (2.24) and (2.23) $P_{nst}^p = 0$ and $P_{nst}^c = O(\theta_p)$. The integral in this region is then of $O(\theta_p^2 a^2)$, since $v_b = O(n_p r^2)$. In the infinite region ($r > 2a$) with (2.24) and (2.23), we find $P_{nst}^p - P_{nst}^c = (e^{7\theta_p} - 1) P_{nst}^c$, and then the integral is of $O(\theta_p^{4/3})$, after noting $\int_{r>2a}^\infty n_p r^2 P_{nst}^c \, d^3 y = O(\theta_p^{1/3})$.

This calculation of $\overline{V}'$ avoids the need for the first renormalization of Batchelor (1972). The leading-order effect of the particle phase is $\theta_p^{1/3}$ relative to the reference velocity v_0. This also explains the reason for the divergence difficulty in using (2.25). Had the integral converged, the effect of the particle phase relative to v_0 would be of $O(\theta_p)$. The particle

phase also causes an $O(\theta_p^{1/3})$ change in the average fluid velocity $\langle \boldsymbol{v}_c \rangle$. These two $O(\theta_p^{1/3})$ changes cancel each other when the velocity $\overline{V'}$ is expressed in terms of the physically more meaningful average velocity $\langle \boldsymbol{v}_c \rangle$ of the fluid phase, instead of $\boldsymbol{v}_0$. The effect of a particle at $\boldsymbol{y}$ to the particle at $\boldsymbol{x}$ is either considered explicitly through velocity $\boldsymbol{v}_c^1$, if the particle is the nearest to $\boldsymbol{x}$, or considered as a part contributing to the background field $\boldsymbol{v}_b$. As the distance to the nearest particle increases, the probability for the particle at $\boldsymbol{y}$ needing the explicit consideration is reduced because of the rapid decay of the nearest particle density. This shields the calculation from the slow decay of the Stokes interaction and avoids the <u>need</u> for renormalization.

To calculate $\overline{V''}$, Batchelor (1972) used

$$\nabla^2 \boldsymbol{v}_c^1 = (\bar{\boldsymbol{v}}_p - \boldsymbol{v}_0)\frac{3a}{2r^3} - (\bar{\boldsymbol{v}}_p - \boldsymbol{v}_0) \cdot \frac{\boldsymbol{r}\,\boldsymbol{r}}{r\,r}\frac{9a}{2r^3}. \tag{3.13}$$

The integral converges only conditionally by using this in (2.25). The second renormalization is then used. Without the renormalization in the present method, $\overline{V''}$ is calculated by using the new relation between the ensemble average and the nearest particle statistics. Using (3.6b) in (3.5) (with V' replaced by V''), one finds

$$\overline{V''} = \frac{a^2}{6} \int \langle \nabla^2 \boldsymbol{v}_c \rangle_{nst}(\boldsymbol{x}, \boldsymbol{y}) P_{nst}^p(\boldsymbol{y}|\boldsymbol{x})\, \mathrm{d}^3 y. \tag{3.14}$$

With this relation, there is not convergence difficulty, but we cannot calculate $\langle \nabla^2 \boldsymbol{v}_c \rangle_{nst}$ by directly taking the Laplacian of (3.8) and then using (3.13) because $\langle \nabla^2 \boldsymbol{v}_c \rangle_{nst} \neq \nabla^2 \langle \boldsymbol{v}_c \rangle_{nst}$. To calculate $\overline{V''}$, we note the similarity of the $\overline{V''}$ and the corresponding average $\langle \nabla^2 \boldsymbol{v}_c \rangle$ for the continuous phase. By adding and subtracting $a^2/6\langle \nabla^2 \boldsymbol{v}_c \rangle$ to (3.14) and then using (2.6) with $q_c = \nabla^2 \boldsymbol{v}_c$ we have

$$\overline{V''} = \frac{a^2}{6} \int \langle \nabla^2 \boldsymbol{v}_c \rangle_{nst}(\boldsymbol{x}, \boldsymbol{y})[P_{nst}^p(\boldsymbol{y}|\boldsymbol{x}) - P_{nst}^c(\boldsymbol{y}|\boldsymbol{x})]\, \mathrm{d}^3 y + \frac{a^2}{6}\langle \nabla^2 \boldsymbol{v}_c \rangle(\boldsymbol{x}). \tag{3.15}$$

Similar to the discussion after (3.12) we have $P_{nst}^p(\boldsymbol{y}|\boldsymbol{x}) - P_{nst}^c(\boldsymbol{y}|\boldsymbol{x}) = (\mathrm{e}^{7\theta_p} - 1)P_{nst}^c$ for $|\boldsymbol{y} - \boldsymbol{x}| > 2a$. The integral in (3.15) converges absolutely and can be integrated in any order over the solid angle or radial direction. Applying (A 7) in appendix A twice for the Laplacian and then using (3.8) and (3.7), we have

$$\langle \nabla^2 \boldsymbol{v}_c \rangle_{nst} = \nabla^2 \langle \boldsymbol{v}_c \rangle_{nst} + O(n_p r^2) = (\bar{\boldsymbol{v}}_p - \boldsymbol{v}_0)\frac{3a}{2r^3} - (\bar{\boldsymbol{v}}_p - \boldsymbol{v}_0) \cdot \frac{\boldsymbol{r}\,\boldsymbol{r}}{r\,r}\frac{9a}{2r^3} + O(n_p r^2). \tag{3.16}$$

After integrating (3.16) over the solid angle we find that the first two terms integrate to zero. Relation (3.15) then becomes

$$\overline{V''} = \tfrac{1}{6}a^2 \langle \nabla^2 \boldsymbol{v}_c \rangle + O(\theta_p^{4/3}), \tag{3.17}$$

where the $O(\theta_p^{4/3})$ term comes from the last term of (3.16) after its integration in (3.15). Using relation (2.13) of Zhang & Prosperetti (1997) with $f_c = \nabla \boldsymbol{v}_c$, one finds

$$\overline{V''} = \frac{1}{6}a^2 \nabla \cdot \langle \nabla \boldsymbol{v}_c \rangle - \frac{n_p a^2}{6\theta_c} \int_{|\boldsymbol{x}-\boldsymbol{z}|=a} [\langle \nabla \boldsymbol{v}_c \rangle_1(\boldsymbol{z}|\boldsymbol{x}) - \langle \nabla \boldsymbol{v}_c \rangle(\boldsymbol{z})] \cdot \boldsymbol{n}\, \mathrm{d}S_z + O(\theta_p^{4/3}), \tag{3.18}$$

where $\langle \nabla \boldsymbol{v}_c \rangle_1(\boldsymbol{z}|\boldsymbol{x})$ is the average of $\nabla \boldsymbol{v}_c$ at $\boldsymbol{z}$ on the surface of the particle centred at $\boldsymbol{x}$. The last term of (2.13) of Zhang & Prosperetti (1997) vanishes because the field is uniform.

For the same reason, the first term of (3.18) is also zero, and the last term in the square bracket, $\langle \nabla v_c \rangle(z)$, is a constant (independent of z) and integrates to zero. We then have

$$\overline{V''} = -\frac{n_p a^2}{6\theta_c} \int_{|x-y|=a} \langle \nabla v_c \rangle_1 (y|x) \cdot n \, \mathrm{d}S_y + O(\theta_p^{4/3}). \tag{3.19}$$

Noting that $\langle \nabla v_c \rangle_1 = \nabla \langle v_c \rangle_1 + O(\theta_p)$ (Zhang & Prosperetti 1997) and that $\langle v_c \rangle_1$ can be approximated by $v_c^1 + v_0$ with v_b' set to zero and $r = y - x$ in (3.7), with an error of $O(\theta_p^{4/3})$, one can then calculate the integral directly to find

$$\overline{V''} = -\frac{n_p a^2}{6\theta_c} \int_{|x-y|=a} \nabla v_c^1 \cdot n \, \mathrm{d}S_y = -\frac{n_p a^2}{6\theta_c} \int_{r=a} \frac{\partial v_c^1}{\partial r} \, \mathrm{d}S_y = \frac{\theta_p}{2}(\bar{v}_p - v_0). \tag{3.20}$$

This calculation of $\overline{V''}$ shows again that the ensemble average calculated using the nearest particle statistic includes the effects from particles other than the nearest one; otherwise, we would have $\langle \nabla^2 v \rangle_{nst} = \nabla^2 v_c^1$, the Laplacian of the velocity around a single particle, and erroneously find $\overline{V''} = 0$ by directly using (3.14).

To calculate $\bar{W}$, Batchelor (1972) linearly superposed the solution of Stimson & Jeffery (1926) for two spheres aligned in the flow direction and the solution of Goldman, Cox & Breebber (1966) for two spheres aligned perpendicular to the flow. It is found that $\bar{W}_{nst}(x, y)$ decays as $1/|y - x|^4$ in the far field, and can be calculated using (2.25) without the convergence difficulty. In our new method, using (3.5) (with $\overline{V'}$ replaced by W) and (2.23) we have

$$\bar{W} = \int \bar{W}_{nst}(x, y) P_{nst}^p(y|x) \, \mathrm{d}^3 y = n_p \int_{r>2a} \bar{W}_{nst}(x, x+r) \exp(-4\pi n_p[r^3 - (2a)^3]/3) \, \mathrm{d}^3 r. \tag{3.21}$$

As mentioned in §2.3, for a fast-decaying $\bar{W}_{nst}$, the exponential function can be approximated as one. The integral can then be calculated using (2.25), which is (5.7) of Batchelor (1972). The integration is performed by dividing the region into a near-field (lubrication) part with $2 \leqslant r/a \leqslant 8$ and a far-field part with $r/a > 8$. The result is

$$\bar{W} = -1.55\theta_p(\bar{v}_p - v_0) + O(\theta_p^2). \tag{3.22}$$

With (3.4), (3.12), (3.20) and (3.22), the average Stokes drag can be expressed in terms of relative velocity between the phases as follows:

$$\bar{f}_p = -6\pi a \mu (1 + 5.55\theta_p)(\bar{v}_p - \langle v_c \rangle) + O(\theta_p^{4/3}). \tag{3.23}$$

In the calculation above, the procedures of Batchelor (1972) are followed to compare our new method with the renormalization method. Since the renormalization technique is based on the zero mixture velocity in the particle sedimentation problem, the use of the mixture velocity is avoided intentionally here to show no renormalization is needed in this new calculation. To calculate the average particle sedimentation velocity in the frame of reference in which the mixture velocity $\theta_c \langle v_c \rangle + \theta_p \bar{v}_p = 0$, we set $\bar{f}_p = 4\pi a^3(\rho_p - \rho_c)g/3$, the difference between the gravity and buoyancy in (3.23) and find the same result as Batchelor (1972), namely

$$\bar{v}_p = (1 - 6.55\theta_p)\frac{2a^2}{9\mu}(\rho_p - \rho_c)g. \tag{3.24}$$

In this calculation, it is assumed (Batchelor 1972) that the particle distribution is random and uniform. Experiments of Ham & Homsy (1988) suggest that at the stationary state

particles form microstructures. This subject is further studied by Cichocki & Sadlej (2005) and Felderhof (2008). The presence of particle microstructures affects the nearest particle probability densities. Relations (2.23) and (2.24) need to be modified to consider them. Since the main objective of this work is not to calculate the sedimentation velocity, but to study the average method using the nearest particle statistics and to compare it with the renormalization technique, Batchelor's assumption is adapted above. Since the time is absent in the momentum and continuity equations for an incompressible Stokes flow, the instantaneous average particle sedimentation velocity can only depend on the particle distribution and the boundary conditions. Despite the transient nature of a uniform and random particle distribution in the sedimentation process, the average sedimentation velocity under this particle distribution calculated by Batchelor (1972) is a well-defined and important quantity. It could be interesting to study the average velocity change caused by the microstructure evolution during the sedimentation process.

4. Particle–fluid–particle stress

In addition to obtaining the classical result of sedimentation velocity in the last section, we now show that the nearest particle statistics introduces a new method to explore particle–fluid and particle–fluid–particle (PFP), or three-way, interactions (Fox 2012), in a multiphase flow. A new PFP stress is introduced.

Let $q_{p,nst}(\boldsymbol{x}, \boldsymbol{y})$ in (2.15) be the fluid force $\bar{\boldsymbol{f}}_{nst}(\boldsymbol{x}, \boldsymbol{y})$, such as the drag, or the added mass force, etc., acting on the particle at $\boldsymbol{x}$ conditional on the nearest particle at $\boldsymbol{y}$. Using (2.15) and (2.21), we can calculate the ensemble average of the fluid force $\boldsymbol{f}_p$ on the particle at $\boldsymbol{x}$ as

$$n_p(\boldsymbol{x})\bar{\boldsymbol{f}}_p(\boldsymbol{x}) = \int \bar{\boldsymbol{f}}_{nst}(\boldsymbol{x}, \boldsymbol{y})\bar{h}_2(\boldsymbol{x}, \boldsymbol{y})P_2(\boldsymbol{x}, \boldsymbol{y})\, \mathrm{d}^3y = \int F(\boldsymbol{x}, \boldsymbol{r})\, \mathrm{d}^3r, \qquad (4.1)$$

where $\boldsymbol{r} = \boldsymbol{y} - \boldsymbol{x}$, and $F(\boldsymbol{x}, \boldsymbol{r}) = \bar{\boldsymbol{f}}_{nst}(\boldsymbol{x}, \boldsymbol{x} + \boldsymbol{r})h_2(\boldsymbol{x}, \boldsymbol{x} + \boldsymbol{r})P_2(\boldsymbol{x}, \boldsymbol{x} + \boldsymbol{r})$ is the force density in the six-dimensional space $(\boldsymbol{x}, \boldsymbol{r})$. In this form, $\boldsymbol{x}$ is the centre of the particle on which the fluid force is applied, while its nearest neighbour is at distance $\boldsymbol{r}$ away. To explore the statistical effect of the nearest particle on the force, it is natural to study the Pearson correlation between the nearest particle distance $\boldsymbol{r}$ and the force. It turns out that the Pearson product moment

$$\boldsymbol{\Sigma}_{pfp}(\boldsymbol{x}) = \frac{1}{2} \int \boldsymbol{r}F(\boldsymbol{x}, \boldsymbol{r})\, \mathrm{d}^3r = \frac{n_p(\boldsymbol{x})}{2} \int (\boldsymbol{y} - \boldsymbol{x})\bar{\boldsymbol{f}}_{nst}(\boldsymbol{x}, \boldsymbol{y})P_{nst}^p(\boldsymbol{y}|\boldsymbol{x})\, \mathrm{d}^3y, \qquad (4.2)$$

has a significant meaning. In this definition the factor $1/2$ is for convenience that will be soon clear. In cases where $\bar{\boldsymbol{f}}_{nst}$ is independent of the nearest particle position $\boldsymbol{y}$, for an isotropic particle distribution this quantity is zero. This quantity contains information about particle interaction and distribution around $\boldsymbol{x}$. According to Torquato, Lu & Rubinstein (1990a) and Torquato *et al.* (1990b), for systems with finite particle volume fractions, the nearest particle probability density $P_{nst}^p(\boldsymbol{y}|\boldsymbol{x})$ also decays as $\exp(-4\pi n_p r^3/3)$ in the far field; the integral in (4.2) then converges absolutely. If the nearest probability density P_{nst}^p in (4.2) is replaced with $P(\boldsymbol{y}|\boldsymbol{x})$, the probability density of having a particle (not necessarily the nearest) at $\boldsymbol{y}$ conditional on a particle at $\boldsymbol{x}$, the integral diverges strongly for long-range interaction forces; therefore, a similar quantity cannot be defined using the classical kinetic theory for multiphase flows. The quantity defined in (4.2) has the dimension of stress, and the second identity is similar to the definition of the potential part of the virial stress in molecular systems (Irving & Kirkwood 1950). We now call it

the PFP stress. These observations motivate the following Taylor expansion of the force density $F(x, r)$:

$$F\left(x - \frac{r}{2}, r\right) = F(x, r) - \frac{r}{2} \cdot \nabla_x F(x, r) + O(\ell_p^2/L^2), \qquad (4.3)$$

where ℓ_p is the typical particle separation and L is the length scale of the physical problem. The subscript x in ∇ above emphasizes that gradient is taken on the first variable x, not on the second variable r. Using this in (4.1), we have

$$n_p(x)\bar{f}_p(x) = \int F\left(x - \frac{r}{2}, r\right) d^3r + \nabla \cdot \Sigma_{pfp}(x) + O(\ell_p^2/L^2), \qquad (4.4)$$

after exchanging the integration and the gradient operator in the second term. This exchange is legitimate because the integration variable is r, while the gradient is on x. By changing integration variables $r \to -r'$, the integral in (4.4) can also be calculated as $\int F(x + r'/2, -r') d^3r'$. Noting that the integral is independent of the integration variables, we can rewrite the first term in (4.4) as

$$n_p(x)f_{pm}(x) = \int \frac{F\left(x - \frac{r}{2}, r\right) + F\left(x + \frac{r}{2}, -r\right)}{2} d^3r \qquad (4.5)$$

and then write (4.4) as

$$n_p(x)\bar{f}_p(x) = n_p(x)f_{pm}(x) + \nabla \cdot \Sigma_{pfp}(x) + O(\ell_p^2/L^2). \qquad (4.6)$$

In the first term of the numerator in (4.5), the fluid force is on the particle centred at $x - r/2$ with the nearest particle at $x + r/2$, while in the second term, the roles of these two particles are reversed. The integrand of (4.5) represents the pair-mean field interaction for a pair of particles on the both sides ($\pm r/2$) of x. The force f_{pm} in (4.5) is then the average particle-mean field interaction force. The forces resulting from PFP interactions, such as the repulsion and attraction forces between the pair, (e.g. Bernoulli effects), have no contribution to this particle-mean field interaction force. These types of forces are antisymmetric (equal in magnitude and opposite in direction on the particles) and are the only contribution in the PFP stress. By decomposing $F(x, r)$ into the symmetric part $F_s(x, r)$, satisfying $F_s(x, -r) = F_s(x, r)$, and the antisymmetric part, $F_a(x, r)$, satisfying $F_a(x, -r) = -F_a(x, r)$, from definition (4.2), one can see that only the antisymmetric part contributes to the stress Σ_{pfp}.

Since the stress appears in (4.6) under the divergence operator, these types of PFP interactions only contribute to the particle–fluid force in statistically inhomogeneous systems. In statistically homogeneous systems, $f_{pm} = \bar{f}_p$. The force has been studied extensively (e.g. Patankar & Joseph 2001; Zhang & Prosperetti 2005; Wang *et al.* 2016; Subramanian & Balachandar 2017). Since most current numerical simulations are performed in homogeneous flows, this new PFP stress has not been studied. Clearly, more work is needed.

If we introduce a concept similar to the local thermal dynamic equilibrium in a molecular system, for cases where $L \gg \ell_p$, the PFP stress can be calculated in statistically homogeneous systems as a good approximation to a statistically inhomogeneous system. In other words, models for this stress can be developed using the studies of homogeneous systems. By adding the divergence of this stress to the well-studied phase interaction force

f_{pm}, the force density calculated using (4.6) could be a good approximation to multiphase flows with gradients of mean fields.

Suppose we have results from particle-scale-resolved numerical simulations for a homogeneous system. Let N_p be the number of particles in the system and f_i be the force on particle i. The volume average of the PFP stress can be calculated as

$$\boldsymbol{\Sigma}_{pfp} = \frac{1}{2V} \sum_{i=1}^{N_p} \frac{1}{N_i} \sum_{j=1}^{N_i} \boldsymbol{r}_{ij} \boldsymbol{f}_i, \qquad (4.7)$$

where V is the volume of the computational domain, N_i is the number of the nearest neighbours of particle i, and $\boldsymbol{r}_{ij}$ is the vector distance from particle i to its j-th nearest neighbour. In most of the cases with moving (or randomly placed) particles, each particle has only one nearest neighbour ($N_i = 1$). Time average can also be performed for this stress if the system is statistically steady.

A decomposition of the particle–fluid force similar to (4.6) and a particle phase stress has been found by Nott, Guazzelli & Pouliquen (2011) for Stokesian suspensions using the volume averaging method. Both the particle phase stress of Nott *et al.* (2011) and the PFP stress defined here describe the particle interactions at the interparticle length scale. In the definition of the particle phase stress, the effect of multi-particle interactions is expressed explicitly by the summation over particle pairs as in (66) of Nott *et al.* (2011), taking advantage of the linearity of the Stokes flow. In the PFP stress definition (4.2), the multi-particle interactions are considered in the force $\bar{f}_{nst}(\boldsymbol{x}, \boldsymbol{y})$, which is the average force acting on the particle at $\boldsymbol{x}$ conditional on the nearest particle at $\boldsymbol{y}$. This force includes nonlinear effects and contributions from all particles, including, but not only the nearest pair. This important point is also contained in (4.7) for the volume average of the PFP stress. The force f_i on particle i is a function of all particles. In the case of a Stokesian suspension, f_i can be calculated using (58) and (59) of Nott *et al.* (2011). However, the force decomposition ((58) of Nott *et al.* (2011)) is generally not available for nonlinear cases as pointed out by the authors. Clearly, the PFP stress defined here and the particle phase stress are closely related, at least for Stokesian suspensions. Their relations need to be studied more carefully.

5. Conclusion

A new relation between the ensemble average and the average conditional on the nearest particle is derived rigorously. This relation is exact and is valid for linear and nonlinear problems with finite particle volume fractions. Because of the rapid far-field decay of the nearest particle probability density function, the integral for calculating the ensemble average converges absolutely and can be used to study long-range particle interactions.

As an example of its application, the Stokes drag in a dilute particle dispersion is calculated following the steps of Batchelor (1972), but without using the difficult renormalizations. The first renormalization is avoided by using the new relation to directly calculate the would-be fluid velocity at the particle centre if the particle is replaced by the same volume fluid. The difference between this velocity and the average fluid velocity is found to be of the first order in the particle volume fraction and is proportional to the relative velocity between the two phases, after noting that the motion of the particles causes a change in the average fluid velocity. The second renormalization is bypassed because the average of the velocity Laplacian is not the same as the Laplacian of the

average velocity due to the effects from other particles. The commutator of the ensemble phase average operator and the differentiation operator from Zhang & Prosperetti (1997) is used to calculate the average of the velocity Laplacian. The Stokes drag from these calculations is then used to calculate the particle sedimentation velocity. The result agrees with that of Batchelor (1972). This example shows that the nearest particle statistics considers multi-particle effects correctly.

As an application of the nearest particle statistics, a new concept of PFP stress is developed for general multiphase flows. With the stress, results from numerical simulations with particle-scale resolved can be further interrogated to extract physics other than traditional particle–fluid interaction forces. This stress could have an interesting implication for hyperbolicity of the averaged equations fundamentally important for multiphase flow theories (Ndjinga 2007; Fox 2012; Theofanous & Chang 2018; Fox 2019).

Acknowledgements

Declaration of interests

The author reports no conflict of interest.

Appendix A

We calculate the gradient of an average conditional on the nearest particle.

A.1. *For a continuous phase quantity*

Let $r_i = \xi_i(\mathscr{F}) - x$, $r_i = |\xi_i - x|$, then $\nabla r_i = -\hat{r}_i$, where $\hat{r}_i = r_i/r_i$. We calculate

$$\nabla \prod_j H(|\xi_j - x| - |\xi_i - x|)$$

$$= \sum_k \delta(r_k - r_i)(\hat{r}_i - \hat{r}_k) \prod_{j \neq k} H(r_j - r_i)$$

$$= \sum_k \delta(r_k - r_i)(\hat{r}_i - \hat{r}_k) \prod_{j \neq k} H(r_j - r_i) \prod_{j \neq i} H(r_j - r_k), \quad (A\,1)$$

where the last identity comes from the fact that $\prod_{j \neq i} H(r_j - r_k)$ and $\prod_{j \neq i} H(r_j - r_i)$ are either 0 or 1, and $\prod_{j \neq i} H(r_j - r_k) = \prod_{j \neq k} H(r_j - r_i)$ for given i and k, since $r_i = r_k$ as required by the δ-function. Using (2.4), we have

$$\nabla N_x(x, \mathscr{F}) = \sum_i \sum_k (\hat{r}_i - \hat{r}_k)\delta(r_k - r_i) \prod_{j \neq k} H(r_j - r_i) \prod_{j \neq i} H(r_j - r_k) = 0, \quad (A\,2)$$

the last identity comes from the exchange of indices i and k. This relation is in confirmation that $N_x(x, \mathscr{F}) = 1$ almost everywhere (with the exception only in a set of

zero probability). Using the first line of (A 1), the gradient of h_i defined in (2.3) is

$$\nabla h_i(\boldsymbol{x}, \mathscr{F}) = \sum_k \delta(r_k - r_i)(\hat{\boldsymbol{r}}_i - \hat{\boldsymbol{r}}_k) \cdot \prod_{j \neq k} H(r_j - r_i). \tag{A 3}$$

Using (2.8) we have

$$\nabla[\theta_c(\boldsymbol{x})P_{nst}^c(\boldsymbol{y}|\boldsymbol{x})] = \int \nabla \left\{ \chi_c(\boldsymbol{x}, \mathscr{F}) \sum_i h_i(\boldsymbol{x}, \mathscr{F}) \delta[\boldsymbol{\xi}_i(\mathscr{F}) - \boldsymbol{y}] \right\} d\mathscr{P}. \tag{A 4}$$

Using (2.7) we have

$$\langle q_c \rangle_{nst}(\boldsymbol{x}, \boldsymbol{y}) \nabla[\theta_c(\boldsymbol{x})P_{nst}^c(\boldsymbol{y}|\boldsymbol{x})] + \theta_c(\boldsymbol{x})P_{nst}^c(\boldsymbol{y}|\boldsymbol{x}) \nabla \langle q_c \rangle_{nst}(\boldsymbol{x}, \boldsymbol{y})$$

$$= \int \nabla q_c(\boldsymbol{x}, \mathscr{F}) \chi_c(\boldsymbol{x}, \mathscr{F}) \sum_i h_i(\boldsymbol{x}, \mathscr{F}) \delta(\boldsymbol{\xi}_i - \boldsymbol{y}) d\mathscr{P}$$

$$+ \int q_c(\boldsymbol{x}, \mathscr{F}) \nabla \left[\chi_c(\boldsymbol{x}, \mathscr{F}) \sum_i h_i(\boldsymbol{x}, \mathscr{F}) \delta(\boldsymbol{\xi}_i - \boldsymbol{y}) \right] d\mathscr{P}. \tag{A 5}$$

Using (A 3), (A 4) and definition (2.7), one finds

$$\theta_c(\boldsymbol{x})P_{nst}^c(\boldsymbol{y}|\boldsymbol{x}) \nabla \langle q_c \rangle_{nst}(\boldsymbol{x}, \boldsymbol{y}) = \theta_c(\boldsymbol{x})P_{nst}^c(\boldsymbol{y}|\boldsymbol{x}) \langle \nabla q_c \rangle_{nst}(\boldsymbol{x}, \boldsymbol{y})$$

$$+ \int [q_c(\boldsymbol{x}, \mathscr{F}) - \langle q_c \rangle_{nst}(\boldsymbol{x}, \boldsymbol{y})] \sum_i h_i(\boldsymbol{x}, \mathscr{F}) \delta(\boldsymbol{\xi}_i - \boldsymbol{y}) \nabla \chi_c(\boldsymbol{x}, \mathscr{F}) d\mathscr{P}$$

$$+ \int [q_c(\boldsymbol{x}, \mathscr{F}) - \langle q_c \rangle_{nst}(\boldsymbol{x}, \boldsymbol{y})] \chi_c(\boldsymbol{x}, \mathscr{F}) \sum_i \delta(\boldsymbol{\xi}_i - \boldsymbol{y})$$

$$\sum_k (\hat{\boldsymbol{r}}_i - \hat{\boldsymbol{r}}_k) \delta(r_k - r_i) \prod_{j \neq k} H(r_j - r_i) d\mathscr{P}. \tag{A 6}$$

For the last term to be non-zero, a particle (k) other than the one at $\boldsymbol{y}$ has to be at the same distance $r_i = |\boldsymbol{y} - \boldsymbol{x}|$ away from $\boldsymbol{x}$. The last term is then caused by multi-particle interactions. For randomly distributed particles with a small particle volume fraction, the particle locations are independent of each other. The probability density of having both particles i and k as the nearest particles to $\boldsymbol{x}$ at distance $r_k = r_i$ is proportional to $4\pi r_k^2 P_{nst}^c(\boldsymbol{y}|\boldsymbol{x}) n_p(\boldsymbol{x} + \boldsymbol{r}_k)$. Furthermore, for a dispersion of identical spherical particles with radii a, when $|\boldsymbol{y} - \boldsymbol{x}| > a$, $\chi_c(\boldsymbol{x}, \mathscr{F}) = 1$, $\nabla \chi_c(\boldsymbol{x}, \mathscr{F}) = 0$, since $\boldsymbol{y}$ is the nearest particle centre to position $\boldsymbol{x}$. The second term on the right-hand side then vanishes. Under these conditions, using (2.24) we have

$$\nabla \langle q_c \rangle_{nst} = \langle \nabla q_c \rangle_{nst} + O(q_c' 4\pi r^2 n_p), \quad \text{when } r = |\boldsymbol{y} - \boldsymbol{x}| > a, \tag{A 7}$$

where $q_c' = q_c(\boldsymbol{x}, \mathscr{F}) - \langle q_c \rangle_{nst}(\boldsymbol{x}, \boldsymbol{y})$ is the fluctuation component of q_c.

A.2. *For a particle quantity*

Using (2.17) and noting that probability $\mathscr{P}$ is a function of $\boldsymbol{\xi}_i$ as mentioned in § 2.1, one finds

$$
\begin{aligned}
&\boldsymbol{\nabla}[n_p(\boldsymbol{x})P_{nst}^p(\boldsymbol{y}|\boldsymbol{x})\bar{q}_{p,nst}(\boldsymbol{x},\boldsymbol{y})] \\
&= \int \sum_i q_i(\mathscr{F})\boldsymbol{\nabla}\delta[\boldsymbol{\xi}_i(\mathscr{F})-\boldsymbol{x}]\sum_{j\neq i}h_{ij}(\mathscr{F})\delta[\boldsymbol{\xi}_j(\mathscr{F})-\boldsymbol{y}]\,\mathrm{d}\mathscr{P} \\
&= -\int \sum_i q_i(\mathscr{F})\boldsymbol{\nabla}_{\xi_i}\delta[\boldsymbol{\xi}_i(\mathscr{F})-\boldsymbol{x}]\sum_{j\neq i}h_{ij}(\mathscr{F})\delta[\boldsymbol{\xi}_j(\mathscr{F})-\boldsymbol{y}]\,\mathrm{d}\mathscr{P} \\
&= \int \sum_i [\boldsymbol{\nabla}_{\xi_i}q_i(\mathscr{F})]\delta[\boldsymbol{\xi}_i(\mathscr{F})-\boldsymbol{x}]\sum_{j\neq i}h_{ij}(\mathscr{F})\delta[\boldsymbol{\xi}_j(\mathscr{F})-\boldsymbol{y}]\,\mathrm{d}\mathscr{P} \\
&\quad + \int \sum_i q_i(\mathscr{F})\delta[\boldsymbol{\xi}_i(\mathscr{F})-\boldsymbol{x}]\boldsymbol{\nabla}_{\xi_i}\left\{\sum_{j\neq i}h_{ij}(\mathscr{F})\delta[\boldsymbol{\xi}_j(\mathscr{F})-\boldsymbol{y}]\,\mathrm{d}\mathscr{P}\right\}. \quad (A\,8)
\end{aligned}
$$

Setting $q_i = 1$, then $\bar{q}_{p,nst}(\boldsymbol{x},\boldsymbol{y}) = 1$, we have

$$
\boldsymbol{\nabla}[n_p(\boldsymbol{x})P_{nst}^p(\boldsymbol{y}|\boldsymbol{x})] = \int \sum_i \delta(\boldsymbol{\xi}_i-\boldsymbol{x})\boldsymbol{\nabla}_{\xi_i}\left\{\sum_{j\neq i}h_{ij}(\mathscr{F})\delta(\boldsymbol{\xi}_j-\boldsymbol{y})\,\mathrm{d}\mathscr{P}\right\}. \quad (A\,9)
$$

Multiplying this equation by $\bar{q}_{p,nst}(\boldsymbol{x},\boldsymbol{y})$ and subtracting the result from (A 8), after using (2.17) we find

$$
\begin{aligned}
n_p(\boldsymbol{x})P_{nst}^p(\boldsymbol{y}|\boldsymbol{x})\boldsymbol{\nabla}\bar{q}_{p,nst}(\boldsymbol{x},\boldsymbol{y}) &= n_p(\boldsymbol{x})P_{nst}^p(\boldsymbol{y}|\boldsymbol{x})\overline{\boldsymbol{\nabla}_{\xi_i}q_i}(\boldsymbol{x},\boldsymbol{y}) \\
&\quad + \int \sum_i [q_i(\mathscr{F})-\bar{q}_{p,nst}(\boldsymbol{x},\boldsymbol{y})]\delta(\boldsymbol{\xi}_i-\boldsymbol{x})\sum_{j\neq i}h_{ij}(\mathscr{F})\delta(\boldsymbol{\xi}_j-\boldsymbol{y})\boldsymbol{\nabla}_{\xi_i}\,\mathrm{d}\mathscr{P} \\
&\quad + \int \sum_i [q_i(\mathscr{F})-\bar{q}_{p,nst}(\boldsymbol{x},\boldsymbol{y})]\delta(\boldsymbol{\xi}_i-\boldsymbol{x})\delta(\boldsymbol{\xi}_j-\boldsymbol{y})\boldsymbol{\nabla}_{\xi_i}\sum_{j\neq i}h_{ij}(\mathscr{F})\,\mathrm{d}\mathscr{P}. \quad (A\,10)
\end{aligned}
$$

For finite size particles, the second term $\delta(\boldsymbol{\xi}_j-\boldsymbol{y})\boldsymbol{\nabla}_{\xi_i}\,\mathrm{d}\mathscr{P}$ represents the effect of particle i position on particle j position. For a dilute particle dispersion this terms is of order n_p^2. Similar to (A 8) the last term represents multi-particle effects, and is also of order n_p^2 in a dilute particle dispersion. Noting that on the left-hand side $n_p P_{nst}^p$ is also of order n_p^2; therefore, for a particle quantity, we have $\boldsymbol{\nabla}\bar{q}_{p,nst} - \overline{\boldsymbol{\nabla}_{\xi_i}q_i} = O(1)$, in contrast to (A 7) for a continuous phase quantity.

Appendix B

In this appendix, we consider Stokes flows in particle suspensions driven by body forces acting on the particles and the fluid. Since Stokes flow equations are linear, different types of forces can be considered separately. Their combined effect can be obtained easily by superposition. Without loss of generality, we can only consider the gravity.

We calculate the conditionally averaged velocity $\langle \boldsymbol{v}_c\rangle_{nst}(\boldsymbol{x},\boldsymbol{y})$ in the effective medium shown on the right-hand side of figure 2. According to definition (2.7), $\langle \boldsymbol{v}_c\rangle_{nst}(\boldsymbol{x},\boldsymbol{y})$ is

the velocity of the continuous phase at x averaged under the condition that x is in the fluid phase and that y is the nearest particle centre to x. To calculate the velocity, we now introduce the pressure $\langle p \rangle_{nst}(z|x, y)$ and the mixture velocity $\langle v_m \rangle_{nst}(z|x, y)$ at a generic point z, as shown on the right-hand side of figure 2, averaged under the condition that x is occupied by the fluid and that y is the centre of the nearest particle to x. Clearly, $\langle v_c \rangle_{nst}(x, y) = \langle v_m \rangle_{nst}(x|x, y)$, because x is in the particle-free region.

Let ρ_p and ρ_c be densities of the particle and the fluid. Outside the particle-free region ($|z - x| > R = |y - x|$), the medium has an effective density $\theta_p \rho_p + (1 - \theta_p)\rho_c$. In this medium, particle velocity fluctuations (Guazzelli & Hinch 2011) are not restricted by external mechanisms. The Stokes equation for the mixture becomes

$$- \boldsymbol{\nabla} \langle p \rangle_{nst}(z|x, y) + \mu \nabla^2 \langle v_m \rangle_{nst}(z|x, y)$$
$$= H(R - |z - x|)\theta_p(\rho_p - \rho_c)\boldsymbol{g} - [\theta_p \rho_p + (1 - \theta_p)\rho_c]\boldsymbol{g}, \tag{B 1}$$

where $\boldsymbol{g}$ is the acceleration due to gravity. In this equation, we have only considered force monopoles and neglected other $O(\theta_p)$ terms because these force monopoles are the leading contributors of the long-range effects. This equation is (7.4) of Hinch (1977) with only the first three terms on the left-hand side kept, and $3a$ in the fourth terms replaced by R. This difference in terms only cause an $O(\theta_p)$ correction to the velocity solution (Hinch 1977), while for our purpose, as shown in the main text, only the velocity correct to $O(1)$ is needed. We consider these $O(\theta_p)$ terms in (B 1) to ensure they do not cause a larger $O(1)$ effect on the averaged velocities V' and $\langle v_c \rangle$, since an $O(\theta_p)$ term in the equation does not necessarily result in the same order contributions to the average velocities. As shown in the main text, the contributions to the velocities from these terms are $O(\theta_p^{1/3})$.

The last term on the right-hand side of (B 1) can be accommodated by a constant pressure gradient without a consequence for the velocity. The velocity contribution from the first term on the right-hand side of (B 1) can be calculated as

$$v_b(z|x, y) = \theta_p \Delta \rho \boldsymbol{g} \cdot \int_{|x'-x|<|y-x|} S(r') \, \mathrm{d}^3 x', \quad r' = x' - z, \tag{B 2}$$

where $\Delta \rho = \rho_p - \rho_c$, the integration volume is the particle-free region centred at x as shown in figure 2, and $S(r)$ is the Stokeslet,

$$S(r) = -\frac{1}{8\pi\mu}\left(\frac{I}{r} + \frac{rr}{r^3}\right). \tag{B 3}$$

For a specified y at a large $r = |z - x| \gg R = |y - x|$, $v_b(z|x, y)$ decays as $O(n_p R^3/r)$. At the centre of the particle free region, with $z = x$, $v_b(x, y) = v_b(x|x, y) = -(R^2/3\mu)\theta_p \Delta \rho \boldsymbol{g}$.

With solution (B 2), the approximation of the conditionally averaged velocity for a generic position z can be constructed as

$$\langle v_m \rangle_{nst}(z|x, y) = v_0 + \left[\bar{v}_p - v_0 - v_b(y|x, y)\right]\left(\frac{3a}{4r} + \frac{a^3}{4r^3}\right)$$
$$+ \frac{r}{r}\left\{\frac{r}{r} \cdot \left[\bar{v}_p - v_0 - v_b(y|x, y)\right]\right\}\left(\frac{3a}{4r} - \frac{3a^3}{4r^3}\right)$$
$$+ v_b(z|x, y), \quad r = z - y, \tag{B 4}$$

since the first three terms with the corresponding solution for the pressure satisfy the Stokes equation, ((B 1) with the right-hand side set to zero), and $v_b(z|x, y)$ is obtained

from (B 2) satisfying (B 1) because of the Stokeslet. At the centre of the particle free region, $z = x$, this relation becomes (3.8) with $v_b(x, y) = v_b(x|x, y)$ and $v'_b = v_b(y|x, y)$ in (3.7).

This approximation implies the nearest particle at y experiences the background fluid velocity $v_0 + v_b(y|x, y)$ while satisfying the boundary condition $\langle v_m \rangle_{nst}(\infty|x, y) = v_0$. On the surface $z = y + an$ of the particle at y, where n is the unit outward normal, there is an error $v_b(y + an|x, y) - v_b(y|x, y) = O(a\nabla v_b(y|x, y)) = O(n_p aR)$ in satisfying the no-slip boundary condition, since ∇ is of $O(1/R)$. This error is small compared with $v_b(y|x, y) = O(n_p R^2)$, since $a < R$.

Although only the gravity is considered above, the calculation can be generalized if other uniform external force fields are used to maintain a specified average particle velocity. For instance, for electric fields, densities ρ_p and ρ_c above can be regarded as the charge densities, and g can be replaced with the electric field E. In this case, solution (B 4) remains valid; therefore, the force (3.23) in the main text is independent of the external force acting on the dilute dispersion.

REFERENCES

ALLEN, M. P. & TILDESLEY, D. J. 2017 *Computer Similation of Liquids*. Oxford University Press.
ASH, R. B. 1972 *Real Analysis and Probability*. Academic Press.
BATCHELOR, G. K. 1972 Sedimentation in a dilute dispersion of spheres. *J. Fluid Mech.* **52** (2), 245–268.
BRADY, J. F. & BOSSIS, G. 1988 Stokesian dynamics. *Annu. Rev. Fluid Mech.* **20**, 111–157.
CHANDRASEKHAR, S. 1943 Stochastic problems in physics and astronomy. *Rev. Mod. Phys.* **15** (1), 1–89.
CICHOCKI, B. & SADLEJ, K. 2005 Steady-state particle distribution of a dilute sedimenting suspension. *Europhys. Lett.* **72** (6), 93–42.
DREW, D. A. & MANDYAM, H. 1997 Effective media theory using nearest neighbor pair distribution. In *Particle Flow: Process and Rheology*. Springer.
FELDERHOF, B. U. 2008 Hydrodynamic screening and axisymmetric turbulence in the transient sedimentation of a dilute suspension at small Reynolds number. *Physica* A **387** (24), 5999–6012.
FIORE, A. M., WANG, G. & SWAN, J. W. 2018 From hindered to promoted settling in dispersions of attractive colloids: simulation, modeling, and application to macromolecular characterization. *Phys. Rev. Fluids* **3**, 063302.
FOX, R. O. 2019 A kinetic-based hyperbolic two-fluid model for binary hard-sphere mixtures. *J. Fluid Mech.* **877**, 28–29.
FOX, R. O. 2012 Large-eddy-simulation tools for multiphase flows. *Annu. Rev. Fluid. Mech.* **44** (1), 4–6.
GOLDMAN, A. J., COX, R. G. & BREEBBER, H. 1966 The slow motion of two identical arbitrarly oriented spheres through a viscous fluid. *Chem. Engng Sci.* **21**, 1151.
GUAZZELLI, È. & HINCH, J. 2011 Fluctuations and instability in sedimentation. *Annu. Rev. Fluid. Mech.* **43** (1), 9–16.
HAM, J. M. & HOMSY, G. M. 1988 Hindered settling and hydrodynamic dispersion in quiescent sedimenting suspensions. *Intl J. Multiphase Flow* **14** (5), 533–546.
HERTZ, P. 1909 Über den gegenseitigen durchschnittlichen abstand von punkten, die mit bekannter mittlerer dichte im raume angeordnet sind. *Math. Ann.* **67** (3), 38–98.
HINCH, E. J. 1977 An averaged-equation approach to particle interactions in fluid suspension. *J. Fluid Mech.* **83**, 695–720.
IRVING, J. H. & KIRKWOOD, I. G. 1950 The statistical theory of transport processes. IV. The equations of hydrodynamics. *J. Chem. Phys.* **18**, 81–29.
NDJINGA, M. 2007 Influence of interfacial pressure on the hyperbolicity of the two-fluid model. *C. R. Acad. Sci. Paris* **344**, 40–12.
NESTLER, F., PIPPIG, M. & POTT, D. 2015 Fast Ewald summation based on NFFT with mixed periodicity. *J. Comput. Phys.* **285**, 28–15.
NOTT, P. R., GUAZZELLI, E. & POULIQUEN, O. 2011 The suspension balance model revisited. *Phys. Fluids* **23** (4), 043304.

O'Brien, R. W. 1979 A method for the calculation of the effective transport properties of suspensions of interacting particles. *J. Fluid Mech.* **56**, 401–427.

Patankar, N. A. & Joseph, D. D. 2001 Modeling and numerical simulation of particulate flows by the Eulerian–Lagrangian approach. *Intl J. Multiphase Flow* **27**, 165–684.

Piazza, R. 2014 Settled and unsettled issues in particle settling. *Rep. Prog. Phys.* **77** (5), 056602.

Rackers, J. A., Liu, C., Ren, P. & Ponder, J. W. 2018 A physically grounded damped dispersion model with particle mesh Ewald summation. *J. Chem. Phys.* **149** (8), 084115.

Stimson, M. & Jeffery, G. B. 1926 The motion of two spheres in a viscous fluid. *Proc. R. Soc. Lond.* A **111**, 110.

Subramanian, A. & Balachandar, S. 2017 Faxán form of time-domain force on a sphere in unsteady spatially varying viscous compressible flows. *J. Fluid Mech.* **816**, 38–11.

Theofanous, T. G. & Chang, C. H. 2018 Shock dispersal of dilute particle clouds. *J. Fluid Mech.* **841**, 73–45.

Torquato, S., Lu, B. & Rubinstein, J. 1990a Nearest-neighbor function in many-body systems. *Phys. Rev.* A **41** (4), 2059–2075.

Torquato, S., Lu, B. & Rubinstein, J. 1990b Nearest-neighbour distribution function for systems of interacting particles. *J. Phys. A: Math. Gen.* **23**, L103–L107.

Wang, L., Peng, C., Guo, Z. & Yu, Z. 2016 Lattice Boltzmann simulation of particle-laden turbulent channel flow. *Comput. Fluids* **124**, 226–236.

Yao, Y. & Capecelatro, J. 2018 Competition between drag and coulomb interactions in turbulent particle-laden flows using a coupled-fluid-Ewald-summation based approach. *Phys. Rev. Fluids* **3**, 034301.

Zhang, D. Z. & Prosperetti, A. 1994 Averaged equations for inviscid disperse two-phase flow. *J. Fluid Mech.* **267**, 18–19.

Zhang, D. Z. & Prosperetti, A. 1997 Momentum and energy equations for disperse two-phase flows and their closure for dilute suspensions. *Intl J. Multiphase Flow* **23**, 42–53.

Zhang, D. Z., Ma, X. & Rauenzahn, R. M. 2006 Interspecies stress in momentum equations for dense binary particulate systems. *Phys. Rev. Lett.* **97** (4), 048301.

Zhang, D. Z., VanderHeyden, W. B., Zou, Q. & Padial-Collins, N. T. 2007 Pressure calculations in disperse and continuous multiphase flows. *Intl J. Multiphase Flow* **33**, 86–100.

Zhang, Z. & Prosperetti, A. 2005 A second-order method for three-dimensional particle simulation. *J. Comput. Phys.* **210** (1), 292–324.