Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Computational Physics 196 (2004) 173–183
JOURNAL OF COMPUTATIONAL PHYSICS
www.elsevier.com/locate/jcp

Measurement bias of fluid velocity in molecular simulations

Martin W. Tysanner *, Alejandro L. Garcia

Department of Physics, San Jose State University, San Jose, CA 95192-0106, USA

Received 31 July 2003; received in revised form 27 October 2003; accepted 29 October 2003

Abstract

In molecular simulations of fluid flow, the measurement of mean fluid velocity is considered to be a straightforward computation, yet there is some ambiguity in its definition. We show that in systems far from equilibrium, such as those with large temperature or velocity gradients, two commonly used definitions give slightly different results. Specifically, a bias can arise when computing the mean fluid velocity by measuring the mean particle velocity in a cell and averaging this mean over samples. We show that this bias comes from the correlation of momentum and density fluctuations in non-equilibrium fluids, obtain an analytical expression for predicting it, and discuss what system characteristics (e.g., number of particles per cell, temperature gradients) reduce or magnify the error. The bias has a physical origin so although we demonstrate it by direct simulation Monte Carlo (DSMC) computations, the same effect will be observed with other particle-based simulation methods, such as molecular dynamics and lattice gases.

PACS: 05.10.−a; 05.40.−a; 47.11.+j; 51.10.+y

1. Introduction

Since the discovery of long-time tails by molecular dynamics computations circa 1970 [1], it has been evident that molecular simulations of hydrodynamic phenomena allow insight into fluid behavior at microscopic scales. As computational power has increased, molecular computations have become increasingly useful tools for analyzing complex flow phenomena of both practical and theoretical interest, such as vortex shedding, thermal convection, Rayleigh–Taylor mixing, and other forms of instability [2]. More recently, new applications have arisen where the distances involved may be too small for hydrodynamic equations alone to adequately characterize the full range of fluid behavior, and augmentation with molecular simulations becomes a necessity. Nanoscale devices are one important example [3]; the study of fluctuations in fluids driving Feynman ratchets and Brownian motors are another [4].

* Corresponding author. Tel.: +1-408-927-7140; fax: +1-408-924-2917.
E-mail addresses: tysanner@pacbell.net (M.W. Tysanner), algarcia@algarcia.org (A.L. Garcia).
URL: http://www.algarcia.org.

0021-9991/$ - see front matter
doi:10.1016/j.jcp.2003.10.021

In simulating non-equilibrium systems, such as fluid flows, special care is often necessary when measuring hydrodynamic variables. For example, it is well known that subtle issues exist in measuring temperature, partly because temperature is unambiguously defined only at equilibrium [5,6]. Furthermore, defining temperature in terms of translational kinetic energy requires some care due to the instantaneous fluctuations of the center of mass velocity. Independence of variables must be considered; many hydrodynamic variables that are uncorrelated at thermodynamic equilibrium are correlated in non-equilibrium systems due to mode-coupling effects [7].

The measurement of fluid velocity is considered less controversial than that of temperature. Nonetheless, it is important to remember that fluid velocity is a hydrodynamic variable, rather than a mechanical variable like momentum density. In fact, as we demonstrate in this paper, a subtle measurement bias can appear when mean fluid velocity is computed in particle-based molecular simulations such as molecular dynamics [8], direct simulation Monte Carlo (DSMC) [9] and lattice gases [10]. This bias is present in even the simplest molecular simulations; we find it in measuring fluid velocity for a closed system consisting of a monatomic hard sphere gas contained between thermal walls. Furthermore, we show that this bias has its origin in the correlation of non-equilibrium fluctuations and is absent at thermodynamic equilibrium and also in the thermodynamic limit (i.e., number of particles going to infinity at fixed density).

We begin by defining the important quantities and measurement methods in Section 2. In Section 3 we present some illustrative examples of the bias in simulations of simple systems. We analyze the origin of the bias in Section 4 and confirm that our analysis agrees with the results shown in the previous section, and verify specific points with results from additional simulations. Section 5 gives a brief summary and some closing remarks, and indicates future work.

2. Measurement of fluid velocity

For particle-based simulations, there are two common, general methods for calculating fluid velocity. Because differences between these methods form an important part of the context of this paper, we will define each of them before proceeding to a discussion of the actual phenomenon.

We assume a simulation that partitions the system into K cells and gathers statistics separately for each cell. Cell k has volume V_k; during a time t_j it contains $N_k(t_j)$ particles, of mass m, so $M_k = mN_k$ is the fluid mass in the cell. The total number of particles in the system, $N_\Sigma(t_j) = \sum_{k=1}^{K} N_k(t_j)$, will, in general, vary with time, though for a closed system it is necessarily a conserved quantity.

The measurement of an intensive macroscopic variable like instantaneous fluid velocity $\mathbf{u}_k$ or temperature T_k will generally be an average over particles in that cell. Within a given cell, the velocity of particle i is $\mathbf{v}_i$ and its momentum is $m\mathbf{v}_i$; the total fluid momentum in cell k is $\mathbf{J}_k = \sum_{i\in k}^{N_k} m\mathbf{v}_i$. The notation $i \in k$ denotes the index i ranging over the particles within cell k. The instantaneous fluid velocity is the center of mass velocity for the particles in the cell, $\mathbf{u}_k = \mathbf{J}_k/M_k$. Since we take the particles to have equal mass, this fluid velocity equals the average particle velocity $\overline{\mathbf{v}}_k = (1/N_k)\sum_{i\in k}^{N_k} \mathbf{v}_i$. Note that the overbar indicates an average over all particles in a cell. We reserve angle brackets for averages over all samples, for example $\langle N_k \rangle$, as defined below.

The mean fluid velocity $\langle \mathbf{u}_k \rangle$ is an average over samples, which may be taken over time for steady state flows or over an ensemble of systems for time-dependent flows. In this paper we focus on the former although the bias described below is present in both scenarios. In evaluating $\langle \mathbf{u}_k \rangle$ from samples of particle velocities, there is some flexibility in computing the statistics. Two methods find common usage, which we now define.

The first, which we will call cumulative average measurement (CAM), sums the velocities of all particles in a cell over all samples, then divides this cumulative total by the cumulative total number of particles in

the cell to yield a mean velocity per particle for the cell. That is, for cell k over S samples and with t_j denoting a particular sample time, the mean fluid velocity by the CAM definition is

$$\langle \mathbf{u}_k \rangle_c = \frac{\sum_{j=1}^{S} \sum_{i\in k}^{N_k(t_j)} \mathbf{v}_i(t_j)}{\sum_{j=1}^{S} N_k(t_j)} = \frac{\langle \mathbf{J}_k \rangle}{\langle M_k \rangle}, \tag{1}$$

where

$$\langle M_k \rangle = m\langle N_k \rangle = \frac{m}{S} \sum_{j=1}^{S} N_k(t_j)$$

and

$$\langle \mathbf{J}_k \rangle = \frac{1}{S} \sum_{j=1}^{S} \mathbf{J}_k(t_j) = \frac{1}{S} \sum_{j=1}^{S} \sum_{i\in k}^{N_k(t_j)} m\mathbf{v}_i(t_j)$$

are the mean mass and momentum, respectively, in cell k. These averages are similarly defined for each component so $\langle u_k^x \rangle_c = \langle J_k^x \rangle / \langle M_k \rangle$, where the superscript indicates the x component of a vector.

The second method, which we call sample averaged measurement (SAM), obtains the average particle velocity in a cell at each sample, $\overline{\mathbf{v}}_k(t_j)$, then averages these means over all samples:

$$\langle \mathbf{u}_k \rangle_s = \frac{1}{S} \sum_{j=1}^{S} \overline{\mathbf{v}}_k(t_j) = \frac{1}{S} \sum_{j=1}^{S} \frac{1}{N_k(t_j)} \sum_{i\in k}^{N_k(t_j)} \mathbf{v}_i(t_j). \tag{2}$$

From the above definitions, $\langle \mathbf{u}_k \rangle_s = \langle \mathbf{J}_k / M_k \rangle$ which is evidently different from $\langle \mathbf{u}_k \rangle_c = \langle \mathbf{J}_k \rangle / \langle M_k \rangle$; compare Eqs. (1) and (2). The SAM definition is a natural choice when it is useful to measure the instantaneous fluid velocity, for example to monitor the relaxation towards a steady state, and these data are post-processed to obtain mean values.

Note that while $\langle N_k \rangle$ and $\langle \mathbf{J}_k \rangle$ are the same for both SAM and CAM, this is not true in general for $\langle \mathbf{u}_k \rangle$. Although $\langle \mathbf{u}_k \rangle_s$ and $\langle \mathbf{u}_k \rangle_c$ give the same results at thermodynamic equilibrium and in the limit $\langle N \rangle \to \infty$, they differ for finite, non-equilibrium systems. In fact, it is with sample averaged measurement that an interesting bias can occur, which we now illustrate for two similar systems at steady state but far from thermodynamic equilibrium.

3. Simulation results

Our simulations were of a simple closed system containing a monatomic hard sphere gas (particle mass $m = 1$ and diameter $d = 7.5 \times 10^{-2}$) bounded by thermal walls at each end in the x-direction and having periodic boundaries in the y- and z-directions. The total number of particles was $N_\Sigma = 2000$, except in one run where it was increased to $N_\Sigma = 5000$. The system was divided into $K = 20$ rectangular cells of equal size, distributed linearly between the two ends. Distance between the thermal walls was $L = 2.25$ and the cell volume was $V_k \cong 0.57$, giving a mean free path of $\lambda = L/10$ and mean time between collisions of 2.6×10^{-3} at thermal equilibrium when using a gas density of $\rho = M_k/V_k \cong 175$, which corresponds to a volume fraction of 3.4×10^{-2}. Boltzmann's constant was $k_B = 0.5$, so the most probable molecular speed, $\sqrt{2k_BT}$, is unity at the equilibrium reference state of $T = 1$. The sound speed at the reference state is $c = \sqrt{5k_BT/3m} \cong 0.91$.

Particles striking a thermal wall were reinjected into the system with a biased Maxwell–Boltzmann velocity distribution based on the wall's velocity and temperature. Specifically, particles striking the left wall had their velocity components reset, choosing new values distributed as [9]

$$P(v^x) = \frac{m}{k_B T_L} v^x e^{-m v^{x2}/2k_B T_L},$$

$$P(v^y) = \sqrt{\frac{m}{2\pi k_B T_L}} e^{-m(v^y - u_L^y)^2/2k_B T_L},$$

$$P(v^z) = \sqrt{\frac{m}{2\pi k_B T_L}} e^{-m v^{z2}/2k_B T_L},$$

and similarly for the right wall. The left and right walls had temperatures of T_L and T_R, respectively; their velocities in the y-direction were u_L^y and u_R^y. These wall parameters were varied to impose a temperature difference, $\Delta T = T_R - T_L$, or a shear velocity difference (i.e., Couette flow) $\Delta u^y = u_R^y - u_L^y$, across the system.

The simulation itself used DSMC; see Bird [9] for an overview of this method. With one exception, the simulations gathered statistics for each cell during each of $S = 5 \times 10^6$ successive DSMC time steps of duration 2.25×10^{-2}, following an initial system relaxation time of 5×10^6 time steps; the one exception extended S to 5×10^7. Though the initial conditions were close to the steady state, we imposed the relaxation period of half the total simulation time to eliminate the potential for the decaying transients to persist and corrupt the results. Estimating the viscous relaxation time as $\tau_v = L^2/\lambda c$ gives $\tau_v \approx 25$ at the reference state; that is, sampling began after about $2000\tau_v$ (approximately 10^8 collisions).

Fig. 1 shows the measured profile of the x-component of mean fluid velocity using SAM for various temperature gradients and stationary walls. What is immediately striking about the velocity profile is that it

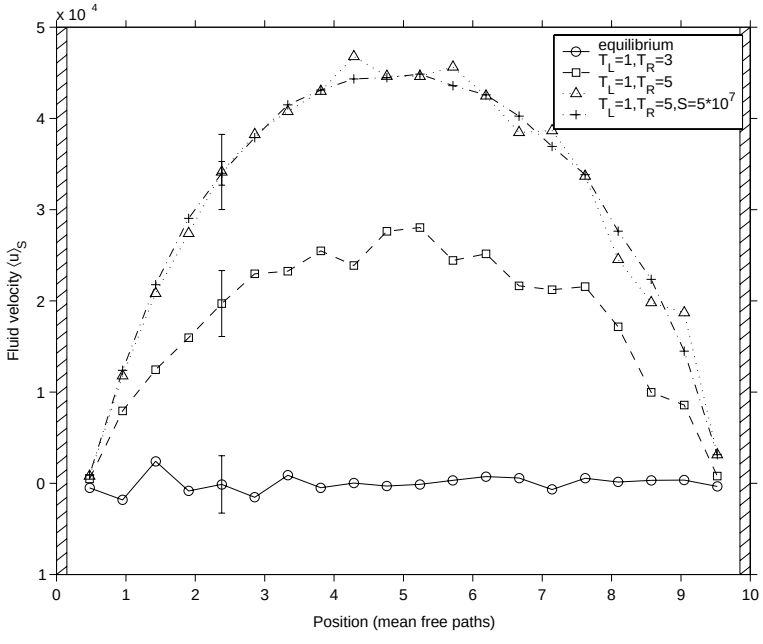

Fig. 1. Sample averaged measurement (SAM) of mean fluid velocity, $\langle u_k^x \rangle_s$, for different temperature gradients. Note the apparent net velocity, implying an anomalous flow from left to right. Plots are: $\Delta T = 0$ (equilibrium, circles), $\Delta T = 2$ (squares), and $\Delta T = 4$ (triangles), all with $S = 5 \times 10^6$ samples; $\Delta T = 4$ with $S = 5 \times 10^7$ (crosses). Evidently, increasing S results in a smoother plot but does not change the magnitude of the effect. The error bar is estimated as $\sigma_u = \sqrt{k_B \langle T_k \rangle / m \langle N_k \rangle S}$ for cell $k = 5$. Other simulation parameters were $N_\Sigma = 2000$, $\Delta u^y = 0$, and $T_L = 1$.

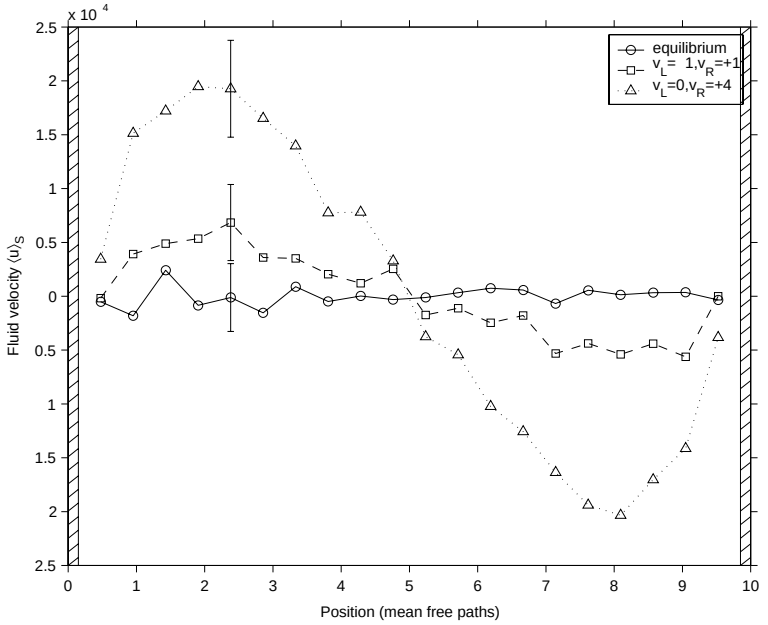

Fig. 2. Sample averaged measurement (SAM) of mean fluid velocity, $\langle u_k^x \rangle_s$, for Couette flow with different shear velocity gradients. Note the anomalous fluid flow from either side toward the center. Plots are: $\Delta u^y = 0$ (equilibrium, circles), $\Delta u^y = 2$ with $u_L^y = -1$ (squares), and $\Delta u^y = 4$ with $u_L^y = 0$ (triangles). Other parameters were $N_\Sigma = 2000$, $T_L = 1$, $\Delta T = 0$, and $S = 5 \times 10^6$ samples. See Fig. 1 caption for an explanation of the error bars.

does *not* have a mean of zero when the temperature gradient is non-zero. This is in contradiction with the expectation that the fluid is at rest in this closed system at steady state. Yet there is a systematic, positive velocity from left to right, suggesting a fluid flow from the cold wall to the hot wall. When we exchanged the temperatures of the walls, we again observed an apparent flow from cold to hot. Initializing the random deviate generator with a different seed had no meaningful effect.

Fig. 2 shows the measurement of $\langle u_k^x \rangle_s$ within a similar closed system, again using SAM, but this time for Couette flow. The walls were at equal temperatures, but their motion imposed a large shear velocity to the fluid. Now we see what appears to be a net fluid flow from left to right between $x = 0$ to $x = L/2$ and from right to left between $x = L/2$ to $x = L$, that is, a flow from the center toward the walls. The apparent net flow turns out to have a similar origin to the first system in that Couette flow creates a temperature gradient within the system due to viscous heating (see Fig. 3), the maximum temperature being in the center of the system. [1]

The error bars on the largest gradient plots of Figs. 1 and 2 show that the effect is statistically significant. These error bars were estimated using results from statistical mechanics [11]. Running the simulation longer and taking more samples gave a smoother plot but did not affect the magnitude of the effect, as Fig. 1 shows.

Fig. 4 shows the results for CAM, the other definition for the mean fluid velocity, for the same systems as in Figs. 1 and 2. In fact, the data for Fig. 4 were taken from the same simulation runs as their counterparts in Figs. 1 and 2. We now have a very different result, with it being apparent that actually there is no net fluid velocity; $\langle u_k^x \rangle_c$ fluctuates about a mean of zero, as expected. Thus the simulation measurements from both

[1] See Section 5 for a discussion on the measurement of temperature.

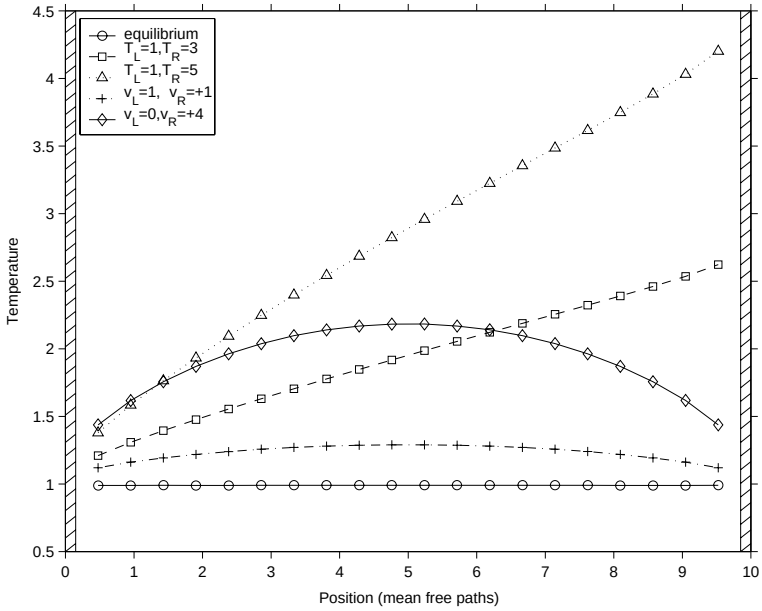

Fig. 3. Temperature profiles for the simulations of Fig. 1 (temperature gradients) and Fig. 2 (Couette flow): equilibrium (circles); $\Delta T = 2$ (squares); $\Delta T = 4$ (triangles); $\Delta u^y = 2$ (crosses); and $\Delta u^y = 4$ (diamonds). The error bar widths, $\sigma_T \approx 5 \times 10^{-4}$, are much smaller than the plot line widths.

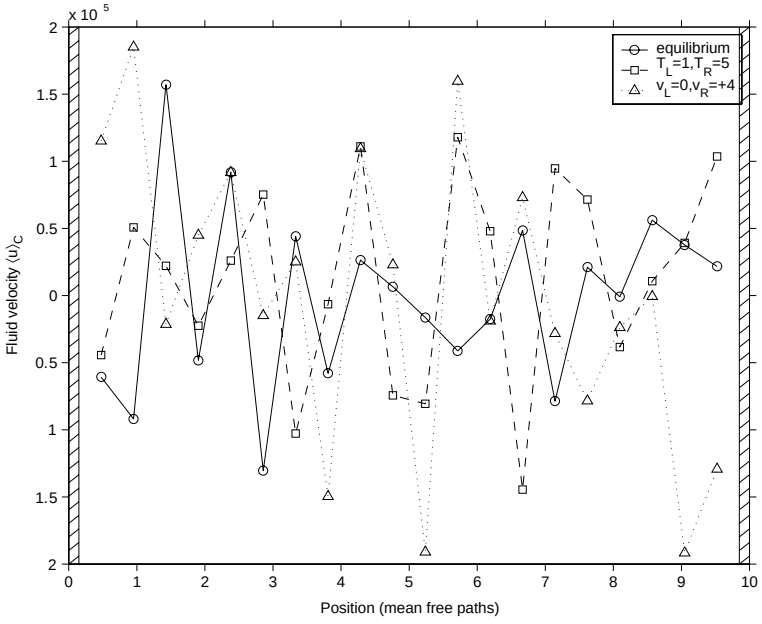

Fig. 4. Cumulative average measurement (CAM) of mean fluid velocity, $\langle u_k^x \rangle_c$; note the absence of net fluid flow. Plots are: $\Delta T = \Delta u^y = 0$ (equilibrium, circles); $\Delta T = 4$, $\Delta u^y = 0$ (squares); and $\Delta u^y = 4$, $\Delta T = 0$ (triangles). Other simulation parameters were $N_\Sigma = 2000$, $T_L = 1$, and $S = 5 \times 10^6$ samples. The error bar width, $\sigma_u \approx 4 \times 10^{-5}$, equals the vertical axis width. Values are computed from the same simulation data sets as in Figs. 1 and 2.

cases indicate that the CAM definition gives the correct result while the SAM definition contains a bias, as evident from the anomalous flow in the closed systems at steady state.

4. Physical explanation for the bias

In examining the origin of the anomalous fluid velocity obtained using SAM, two considerations are key. First, both Figs. 1 and 2 show measurements in systems far from equilibrium, where significant temperature gradients are present (see Fig. 3), although the temperature gradient in the Couette system arises indirectly from viscous heating. In the equilibrium plot of Fig. 1 (walls at equal temperature), the effect is entirely absent. Second, with $N_\Sigma = 2000$ particles and $K = 20$ cells, the mean number of simulation particles $\langle N \rangle$ in any cell was between 75 and 160, depending on the temperature. With so few particles the thermodynamic fluctuations are significant, the theoretical standard deviation of N_k being $\sqrt{\langle N_k \rangle}$, approximately 10% of the mean [12]. Fig. 5 illustrates how increasing the total number of particles to $N_\Sigma = 5000$, without changing any other physical parameters, decreased the measured SAM fluid velocity.

These clues indicate that non-equilibrium fluctuations may be causing the bias. Because a cell contains a finite number of moving particles, there will be statistical fluctuations about the mean of any variable that describes the macroscopic fluid behavior in a given cell. The fluctuation in the number of particles is $\delta N_k \equiv N_k - \langle N_k \rangle$, similarly $\delta \mathbf{J}_k$ is the fluctuation of the total momentum in a cell; by definition these fluctuations have zero mean (but non-zero variance). In the analysis below, we consider the variance, $\langle (\delta N_k)^2 \rangle$, as well as

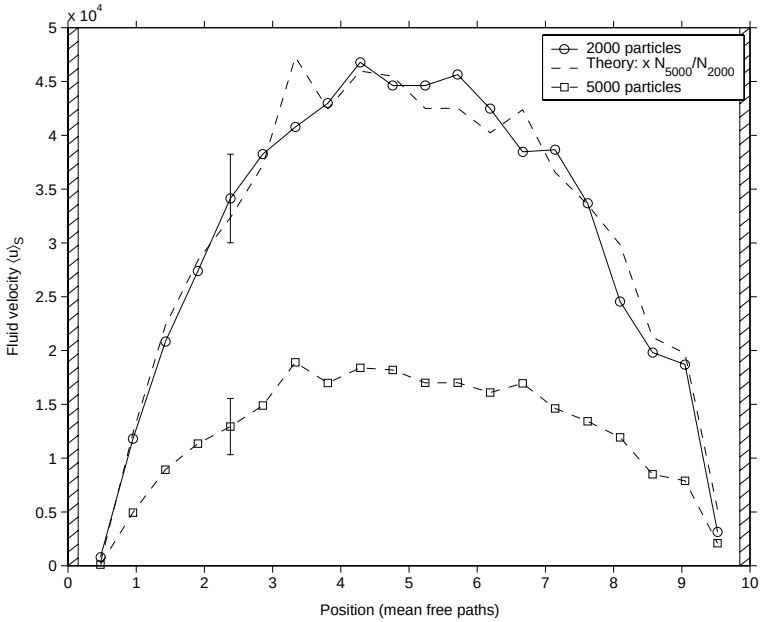

Fig. 5. Sample averaged measurement (SAM) of mean fluid velocity, $\langle u_k^x \rangle_s$, with different total particle numbers. Note that the measured anomalous fluid flow decreases as $\langle N_k \rangle$ increases. This closed system had a gradient $\Delta T = 4$ with: $N_\Sigma = 2000$ (circles) or $N_\Sigma = 5000$ particles (squares). The dashed line without symbols is the lower curve multiplied by 5/2 (see Eq. (7)). Other simulation parameters were $T_L = 1$ and $S = 5 \times 10^6$ samples.

$$\langle \delta \mathbf{J}_k \delta N_k \rangle = \left[\frac{1}{S} \sum_{j=1}^{S} \mathbf{J}_k(t_j) N_k(t_j) \right] - \langle \mathbf{J}_k \rangle \langle N_k \rangle, \tag{3}$$

which is the covariance of momentum and particle number fluctuations in cell k.

Now we analyze the effect quantitatively. CAM measures mean particle velocity according to Eq. (1). This is a straightforward measure of average fluid velocity in a cell and we will not elaborate on it further.

For SAM, which measures mean particle velocity at each sample and then averages these means over all samples (see Eq. (2)), we start by writing it as $\langle \mathbf{u}_k \rangle_s = \langle \mathbf{J}_k / M_k \rangle = \langle \mathbf{J}_k / m N_k \rangle$. Note that the instantaneous total momentum, $\mathbf{J}_k$, may be decomposed into a sum of the mean momentum $\langle \mathbf{J}_k \rangle$ and a fluctuation $\delta \mathbf{J}_k$ about the mean; similarly, $N_k = \langle N_k \rangle + \delta N_k$. Dropping the subscripts because cell k is implied and using the binomial expansion

$$\begin{aligned} \langle \mathbf{u} \rangle_s &= \left\langle \frac{\langle \mathbf{J} \rangle + \delta \mathbf{J}}{m(\langle N \rangle + \delta N)} \right\rangle = \left\langle \frac{\langle \mathbf{J} \rangle + \delta \mathbf{J}}{m \langle N \rangle} \left[1 - \frac{\delta N}{\langle N \rangle} + \left(\frac{\delta N}{\langle N \rangle} \right)^2 + \mathrm{O}(\delta N^3) \right] \right\rangle \\ &= \frac{\langle \mathbf{J} \rangle}{\langle M \rangle} - \frac{\langle \delta \mathbf{J} \delta N \rangle}{m \langle N \rangle^2} + \frac{\langle \mathbf{J} \rangle \langle (\delta N)^2 \rangle}{m \langle N \rangle^3} + \mathrm{O}(\delta X^3) = \langle \mathbf{u} \rangle_c \left(1 + \frac{\langle (\delta N)^2 \rangle}{\langle N \rangle^2} \right) - \frac{\langle \delta \mathbf{J} \delta N \rangle}{m \langle N \rangle^2} + \mathrm{O}(\delta X^3), \end{aligned} \tag{4}$$

where $\mathrm{O}(\delta X^3)$ indicates terms of cubic order in the fluctuations. This gives us an alternative expression for SAM that relates it to the CAM definition.

To relate the above to mode-coupling theory, it is useful to rewrite Eq. (4) in a slightly different form. Since $\langle \mathbf{J} \rangle = m \langle N \rangle \langle \mathbf{u} \rangle_c$, and thus $\delta \mathbf{J} = m \langle N \rangle \delta \mathbf{u} + m \langle \mathbf{u} \rangle_c \delta N$, we can substitute this expression for $\delta \mathbf{J}$ into the $\langle \delta \mathbf{J} \delta N \rangle$ term of Eq. (4) to obtain

$$\langle \mathbf{u} \rangle_s = \langle \mathbf{u} \rangle_c - \frac{\langle \delta \mathbf{u} \delta N \rangle}{\langle N \rangle} + \mathrm{O}(\delta X^3). \tag{5}$$

Note that the covariance term on the right hand side is zero at equilibrium and in the limit $\langle N \rangle \to \infty$ but non-zero for finite, non-equilibrium systems [12]. The $\langle \delta \mathbf{u} \delta N \rangle$ correlation is of theoretical interest since it is the origin of the asymmetric Brillouin peaks observed in light scattering for a fluid subjected to a temperature gradient [7,13].

Returning to Eq. (4), we see that SAM differs from CAM by introducing the fluctuation covariance for the cell, $\langle \delta \mathbf{J} \delta N \rangle$, and mean square fluctuation $\langle (\delta N)^2 \rangle$. For our closed system at steady state, the average total momentum $\langle J_k^x \rangle$ is zero; thus $\langle u_k^x \rangle_c = 0$ and Eq. (4) predicts

$$\langle u_k^x \rangle_s \approx - \frac{\langle \delta J_k^x \delta N_k \rangle}{m \langle N_k \rangle^2} \tag{6}$$

to quadratic order in the fluctuations. To verify this analysis, we measured the covariance $\langle \delta J_k^x \delta N_k \rangle$ (see Eq. (3)) and average particle number, $\langle N_k \rangle$, in the same simulations used for Figs. 1 and 2. Fig. 6 shows the resulting prediction from Eq. (6) along with the corresponding SAM measurement of mean fluid velocity. Agreement is excellent for both the temperature gradient and Couette flow scenarios.

Note that $\langle \mathbf{u}_k \rangle_s \propto 1 / \langle N_k \rangle$ since the variance and covariance of extensive quantities such as $\langle (\delta N_k)^2 \rangle$, $\langle (\delta \mathbf{J}_k)^2 \rangle$, and $\langle \delta \mathbf{J}_k \delta N_k \rangle$ are proportional to $\langle N_k \rangle$ [12]. As mentioned in the discussion for Fig. 5, we ran two similar simulations that differed only in the total number of particles. We expect that

$$\frac{\left(\langle u_k^x \rangle_s \right)_{2000}}{\left(\langle u_k^x \rangle_s \right)_{5000}} = \frac{\left(\langle N_k \rangle \right)_{5000}}{\left(\langle N_k \rangle \right)_{2000}} = \frac{5}{2}, \tag{7}$$

with the subscripts denoting the $N_\Sigma = 2000$ and $N_\Sigma = 5000$ simulations. Fig. 5 confirms this prediction.

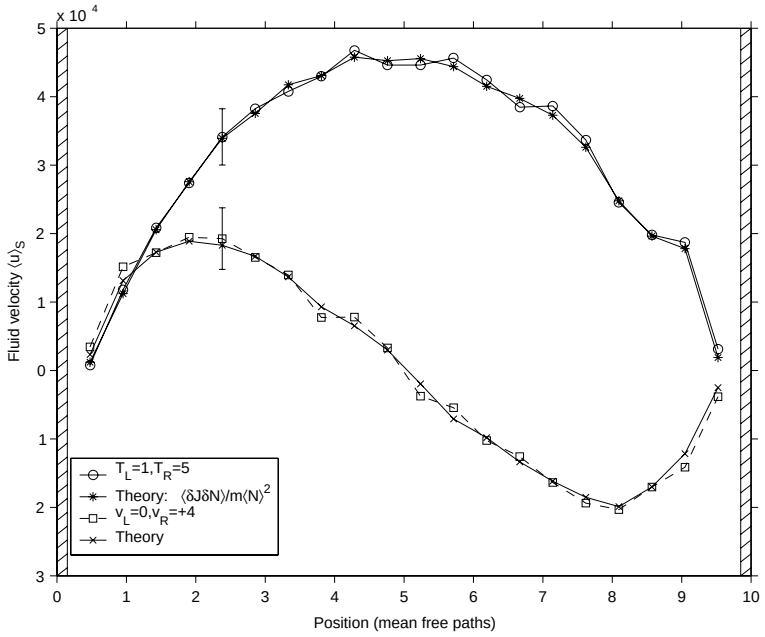

Fig. 6. Comparison between theoretical and measured mean fluid velocity for sample averaged measurement (SAM), $\langle u^x \rangle_s$. The data for the temperature gradient case (circles) are the same as in Fig. 1 for $\Delta T = 4$; the asterisks plot is computed from Eq. (6) by measuring $\langle \delta J_k^x \delta N_k \rangle$ and $\langle N_k \rangle$ in the simulation. Similarly, the data for the Couette flow case (squares) are the same as in Fig. 2 for $\Delta u^y = 4$ with the "×" plot given by Eq. (6).

Earlier theoretical work [13] predicts that the non-equilibrium correlation of fluctuations, which produces a contribution due to the $\langle \delta u_k^x \delta N_k \rangle$ term in Eq. (5), approximately goes as $x(L-x)\Delta T$. From this we expect the $\langle u_k^x \rangle_s$ profiles in Fig. 1 to be approximately quadratic in x, with a magnitude that is proportional to the temperature gradient. Our results are consistent with these predictions.

Finally, in an equilibrium system, Eq. (5) predicts $\langle \mathbf{u}_k \rangle_s = \langle \mathbf{u}_k \rangle_c$. We verified this in two simulations of an equilibrium system having the same physical parameters as that used for Fig. 1, but with periodic boundaries in all directions. In the first simulation the particles were initialized with a mean velocity of $u_0^x = 4.0$ and in the second there was no initial mean flow (i.e., $u_0^x = 0$). In both cases the measured velocity profiles gave $\langle u_k^x \rangle_s = \langle u_k^x \rangle_c = u_0^x$ to within statistical error.

5. Concluding remarks

The purpose of this paper is to demonstrate that the SAM of mean fluid velocity, as defined by Eq. (2), can have a bias due to the correlation of fluctuations in systems far from equilibrium. Not only is this bias observable as an anomalous flow in simulations of closed systems, it is accurately predicted by Eq. (6) as Fig. 6 confirms. On the other hand, the CAM, as defined by Eq. (1), is an unbiased definition of mean fluid velocity.

We can generalize the definitions of SAM and CAM to an abstract hydrodynamic variable, $\mathscr{H}(M, \mathbf{J}, E)$: $\langle \mathscr{H} \rangle_s \equiv \langle \mathscr{H}(M, \mathbf{J}, E) \rangle$ and $\langle \mathscr{H} \rangle_c \equiv \mathscr{H}(\langle M \rangle, \langle \mathbf{J} \rangle, \langle E \rangle)$, where $M, \mathbf{J}$ and E are mass, momentum, and energy, respectively. For example, from the equipartition theorem [12], the translational temperature may defined as,

$$T(M,\mathbf{J},E)=\frac{2m}{3k_{\mathrm{B}}M}\left(E-\frac{|\mathbf{J}|^2}{2M}\right). \tag{8}$$

It is well known that the CAM definition of mean temperature is unbiased while the SAM definition is biased due to the variance of equilibrium fluctuations, specifically $\langle|\mathbf{J}|^2\rangle \neq |\langle\mathbf{J}\rangle|^2$. In this paper we show that a similar bias for fluid velocity exists due to non-equilibrium fluctuations.

Average temperature is sometimes computed as

$$\langle T\rangle=\frac{m}{3k_{\mathrm{B}}}\left(\langle u^2\rangle-|\langle\mathbf{u}\rangle|^2\right). \tag{9}$$

While this definition is unbiased at equilibrium, it has a small bias from the correlation of non-equilibrium fluctuations. However, the temperature profiles shown in Fig. 3 are virtually identical whether one uses $\langle T\rangle_{\mathbf{c}}$ from Eq. (8) or $\langle T\rangle$ from Eq. (9) since the bias is small relative to the gradients. We expect the magnitude of the temperature bias in Eq. (9) to be proportional to $|\langle\mathbf{u}\rangle_{\mathrm{s}}-\langle\mathbf{u}\rangle_{\mathrm{c}}|^2$, which at $T=2$ is of order 10^{-7}, much less than the plot line widths for the cases presented. While it would be of interest to analyze this bias in temperature in more detail, the focus of this paper is the bias of velocity measurements.

Even for systems far from equilibrium, the magnitude of non-equilibrium fluctuations is quite small relative to equilibrium fluctuations. For example, in Fig. 1 for the moderate gradient case the temperature from left wall to right wall increases by a factor of three over a distance of only ten mean free paths. Even with this strong gradient, the anomalous fluid velocity from the SAM definition is small (Mach number of approximately 10^{-4}) and is noticeable above the statistical noise only when the number of samples in our simulations exceeds 10^4. For these reasons, the bias is not significant in many molecular simulations of fluid flows and thus, to our knowledge, has not been investigated.

While the bias, in most cases, is very small, for some problems it can be quite important. For example, in simulations of Brownian motors and related micro-mechanical systems the magnitude of fluctuations is large and the systems are very far from equilibrium. If the purpose of a simulation is to determine whether a Brownian motor will move or a nanoscale pump will create a flow, it is essential to correctly measure the fluid velocity. Other examples include simulations of rarefied gas flows, such as vapor deposition in near-vacuum conditions, aerospace flows in the upper atmosphere and astrophysical flows [14]; or flows with sharp interfaces, such as strong shock waves. Finally, computer simulation programs are usually tested on simple benchmark problems for which the flow properties are known; the temperature gradient and Couette flow scenarios considered in Section 3 are two common benchmarks. The anomalous flow measured by the SAM definition could be mistaken as indicating an error in the programming logic.

We want to emphasize that the correlation of non-equilibrium fluctuations that produces the bias is a general, physical phenomenon, and consequently the bias in SAM is not an artifact of a particular method of molecular simulation. Although we demonstrate the effect in simulations of a dilute gas using the DSMC method, similar correlations of non-equilibrium fluctuations are observed in other particle-based simulations, such as molecular dynamics [15] and lattice gases [16]. Furthermore, while DSMC is a stochastic algorithm, the hydrodynamic fluctuations in DSMC simulations have nothing to do with the Monte Carlo elements of the method, which involve the random selection of collisions among particles.

A technical point peculiar to DSMC simulations is relevant in light of the inversely proportional relationship between the particle number and the magnitude of the SAM bias, as illustrated in Fig. 5. By re-scaling, each DSMC particle may represent a large number of molecules in the physical system being modelled. The variance of fluctuations for extensive variables is proportional to the number of simulation particles, so this variance in a DSMC simulation differs from that in the physical system by the ratio of physical molecules per simulation particle. When this ratio is large, the effect of fluctuations in the DSMC simulation exceeds that of the physical system, typically by many orders of magnitude. When the ratio is one-to-one, as in the simulations we present here, then the fluctuations in the simulation have the same

magnitude as those in the physical system. For DSMC simulations, then, the bias in SAM depends on the number of simulation particles and not the number of physical particles they represent.

Although the CAM definition gives an unbiased measurement of mean fluid velocity there are situations in which the SAM definition is useful. For example, for transient flows, unless an ensemble of systems is considered, the measured fluid velocity will necessarily be the instantaneous velocity. Another example is in particle/continuum hybrids [17,18], where the coupling of the two algorithms requires the estimation of instantaneous hydrodynamic quantities. As we show in this paper, the average of instantaneous fluid velocity is biased; in many applications this small error is acceptable, but researchers should be aware of its existence.

To keep the presentation uncluttered and conceptually clear, in this paper we only consider a single species fluid, that is, all particles having the same mass. The measurement bias in SAM will likely be more pronounced in multi-species fluids, especially when there is a large disparity in the number densities and molecular masses of the species. We will provide a more general treatment in future work. Furthermore, in this paper we only consider closed systems, while in molecular simulations of fluid flow, open systems (''computational wind tunnels'') are common. Typically, simulations of such open systems use particle reservoirs at the system boundaries. The details of how these reservoirs are simulated can affect the correlation of fluctuations in non-equilibrium systems, an effect we also plan to present in a future study.

Acknowledgements

The authors thank M. Malek-Mansour for helpful comments.

References

[1] B. Alder, T. Wainwright, Phys. Rev. A 1 (1970) 18.
[2] J. Koplik, J. Banavar, Annu. Rev. Fluid Mech. 27 (1995) 257.
[3] C. Ho, Y. Tai, Annu. Rev. Fluid Mech. 30 (1998) 579.
[4] P. Reimann, Phys. Rep. 361 (2002) 57.
[5] L. Landau, E. Lifshitz, Fluid Mechanics, Pergamon Press, Oxford, 1959.
[6] L. Marino, Phys. Fluids 15 (2003) 662.
[7] R. Schmitz, Phys. Rep. 171 (1988) 1.
[8] D. Frenkel, B. Smit, Understanding Molecular Simulation, Academic Press, San Diego, 2002.
[9] G.A. Bird, Molecular Gas Dynamics and the Direct Simulation of Gas Flows, Clarendon Press, Oxford, 1994.
[10] J.-P. Rivet, J.P. Boon, Lattice Gas Hydrodynamics, Cambridge University Press, Cambridge, MA, 2001.
[11] N. Hadjiconstantinou, A. Garcia, M. Bazant, G. He, J. Comput. Phys. 187 (2003) 274.
[12] L. Landau, E. Lifshitz, Statistical Physics, Part 1, Butterworth–Heinemann, Oxford, 1980.
[13] M.M. Mansour, J.W. Turner, A.L. Garcia, J. Statist. Phys. 48 (1987) 1157.
[14] A. Ketsdever, E. Muntz, (Eds.), Proceedings of the 23rd International Symposium on Rarefied Gas Dynamics, AIP, College Park, MD, 2003.
[15] M. Mareschal, M.M. Mansour, G. Sonnino, E. Kestemont, Phys. Rev. A 45 (1992) 7180.
[16] A. Suarez, J. Boon, P. Grosfils, Phys. Rev. E 54 (1996) 1208.
[17] A. Garcia, J. Bell, W. Crutchfield, B. Alder, J. Comput. Phys. 154 (1999) 134.
[18] E. Flekkoy, G. Wagner, J. Feder, Europhys. Lett. 52 (2000) 271.