Journal Name

ARTICLE TYPE

Cite this: DOI: 10.1039/xxxxxxxxxx

Thermodynamic and dynamical properties of the hard sphere system revisited by Molecular Dynamics simulation

Sławomir Pieprzyk,[a] Marcus N. Bannerman,[b] Arkadiusz C. Brańka,[a] Maciej Chudak,[c] and David M. Heyes[d]

Received Date
Accepted Date

DOI: 10.1039/xxxxxxxxxx

www.rsc.org/journalname

Revised thermodynamic and dynamical properties of the hard sphere (HS) system are obtained from extensive molecular dynamics calculations carried out with large system sizes (number of particles, N) and long times. Accurate formulas for the compressibility factor of the HS solid and fluid branch are proposed, which represents the metastable region and takes into account its divergence at close packing. Some basic second-order thermodynamic properties are obtained and a maximum in some of their derivatives in the metastable fluid region is found. The thermodynamic parameters associated with the melting-freezing transition have been determined to four digit accuracy which generates accurate new values for the coexistence properties of the HS system. For the self-diffusion coefficient, D, it is shown that relatively large systems ($N > 10^4$) are required to achieve an accurate linear extrapolation of D to the infinite size limit with a D *vs.* $N^{-1/3}$ plot. Moreover, it is found that there is a density dependence to the value of the slope in the linear regime. The density dependent correction becomes practically insignificant at higher densities and the hydrodynamic formula found in the literature is still accurate. However, with decreasing density the density dependence of the size correction cannot be neglected, which indicates that other sources of N-dependence, apart from those derived on purely hydrodynamic grounds, may also be important (and as yet unaccounted for). A detailed analytic representation of the density dependence of the HS self-diffusion coefficient and the HS viscosity, η, is given. It is shown that the HS viscosity near freezing and in the metastable region can be described well by the Krieger-Dougherty equation. Both D and η start to scale at high densities and in the metastable region in such a way that $D\eta^p = \text{const}$, where $p \simeq 0.97$, and $D \to 0$ and $\eta \to \infty$ at a packing fraction of 0.58 density which coincides with some previous predictions of the HS glass transition density.

1 Introduction

The hard sphere (HS) system has proved to be an invaluable model for simple and colloidal liquids, and granular media over many decades, in part because its simple analytic form lends itself to approximate analytic treatments of physical properties. It is also useful as a reference system for perturbation theories of systems interacting with more realistic potentials.[1] Despite the simplicity of the potential, accurate values of its basic fluid and solid properties are accessible only by computer simulations, which has a limited accuracy. Because of advances in computer software and hardware a very accurate representation of the HS properties is now feasible and the 'exact' or 'true' solution can be approached with minimal finite size effects. Some properties need high accuracy, for example the face centered cubic (fcc) crystalline arrangement is only more stable than the hexagonal close packed (hcp) form by a very small free energy difference. Values of, $0.005k_BT$,[2] and $0.001164(8)k_BT$,[3] where k_B is Boltzmann's constant and T is the temperature, have been calculated.

There are a number of HS properties that require long simulations with large systems to be calculated with sufficient precision to resolve certain issues. This study uses molecular dynamics (MD) simulation to perform this task. Improved statistics of the compressibility factor, Z, data at high density (including the metastable fluid region) enable the accuracy of a number of proposed analytic equations of state (EoS) to be tested and im-

[a] *Institute of Molecular Physics, Polish Academy of Sciences, M. Smoluchowskiego 17, 60-179 Poznań, Poland; E-mail: pieprzyk@ifmpan.poznan.pl, branka@ifmpan.poznan.pl*

[b] *School of Engineering, University of Aberdeen, AB24 3UE, United Kingdom; E-mail: m.campbellbannerman@abdn.ac.uk*

[c] *Department of Applied Physics, Eindhoven University of Technology, PO Box 513, 5600 MB Eindhoven, The Netherlands; E-mail: M.Chudak@tue.nl*

[d] *Department of Physics, Royal Holloway, University of London, Egham, Surrey TW20 0EX, United Kingdom; E-mail: david.heyes@rhul.ac.uk*

proved. A more accurate equation of state of the fluid and solid regions will also lead to improved values for the coexisting fluid-solid densities.

Many theories of the dynamics of the molecules in the liquid state also require accurate values of the fluid self-diffusion coefficient, D, and shear viscosity, η, in the thermodynamic limit, to test adequately (e.g., the applicability of the Stokes-Einstein relationship,[4]). It is known that D has a significant system size dependence which can be reduced significantly by taking about 1 million particles in a simulation. Using simulations with typically this number of hard spheres more accurate values of D and η for all densities in the thermodynamic limit have been obtained here. In Sec. 2 some technical simulation details and formulas defining basic calculated properties are given. In Sec. 3 the equations of state of the fluid and solid regions, other derived thermodynamic properties, and values of some properties at coexistence are presented. The self-diffusion coefficient and shear viscosity data are presented and discussed in Sec. 4. The main conclusions are summarized in Sec. 5.

2 Calculated quantities and simulation details

Most of the MD simulations carried out in this work were performed using the DYNAMO program.[5] In the study of the equation of state, the calculations covered essentially the entire density range from the very dilute fluid up to nearly the close packed fcc crystal value. Most of the calculations were made with $N = 1098500$ particles. Additionally, to establish the system size dependence calculations were also made employing, 2048, 8788, 16384, and 131072 particles. The system was equilibrated typically for 1.2×10^9 collisions, then the reported quantities were computed by averaging over a further 7.9×10^9 collisions. To improve and establish the statistical uncertainties at each density the simulation was carried out at least $10-100$ times starting from different random initial particle velocities. The compressibility factor, $Z = P/\rho\sigma^3 k_B T$ (where P is the pressure, $\rho = N/V$ is the number density, and V is the volume) was calculated from the MD simulations using the collision rate formula,[6]

$$Z_{MD} = 1 + \frac{\gamma(N)\sqrt{m\sigma^2\pi\beta}}{3N}\frac{C}{t}, \tag{1}$$

where m is the sphere mass, σ is the hard sphere diameter, $\beta = 1/k_B T$, C is the number of collisions between the spheres and t is the duration of the simulation. The quantity $\gamma(N)$ is a correction factor to convert the pressure calculated in the NVE**M** ensemble to the value for the NVT ensemble, and is given by

$$\gamma(N) = \frac{\Gamma\left[\frac{3(N-1)+1}{2}\right]}{\Gamma\left[\frac{3(N-1)}{2}\right]\left(\frac{3N}{2}\right)^{\frac{1}{2}}}, \tag{2}$$

where Γ is the gamma function. All quantities reported here are in the usual HS units of $k_B T$ for energy, σ for length.

The self-diffusion coefficient was obtained in two ways, by the velocity autocorrelation function (VACF) and the mean square displacement (MSD) routes. The differences between the value of D obtained with those two approaches was found to be less than 0.001. In the fluid phase the calculations were performed for the systems containing, $108 \leq N \leq 1098500$ particles.

The shear viscosity was computed using the Helfand-Einstein relation.[7,8] For discontinuous potentials such as the HS, it is more convenient to use this approach than the more conventional Green-Kubo (GK) route.[7,9] The viscosity calculations were carried out using three system sizes, $N = 16384$, 131072, and 1098500 particles in the simulation box. Although as the viscosity has a relatively weak N-dependence most calculations of this property were performed for a system of $N = 16384$ particles. In these simulations the system was equilibrated for 8.2×10^8 collisions, and then production data were collected over the following 8.7×10^9 collisions. Each simulation was conducted at least $20-40$ times to improve the statistics, and each density was started from different initial random velocities. Statistical errors in all simulation property averages were estimated by the block average method.[10]

The ensemble error in the compressibility factor is caused by suppressed fluctuations in the canonical ensemble. The magnitude of the required correction increases with density, but for a system of $N \sim 10^6$ particles, even at high densities (e.g., $\rho = 0.9$), the correction is about 100 times smaller than the usual statistical error or uncertainty from the system fluctuations.

The correction to the thermodynamic limit of the compressibility factor and shear viscosity due to the periodic boundary conditions becomes insignificant for systems of $N > 10^4$ particles. In the case of the self-diffusion coefficient the N-dependence is significant and requires a special treatment as will be discussed in Sec. 4.

3 Equation of state and thermodynamic properties

3.1 Fluid branch EoS

The compressibility factor for the HS fluid as a function of particle number density calculated by the present MD simulations is given in Table 1. The range of densities covered includes the metastable fluid region. The same fluid region has been the subject of intensive investigation over the years and many different equation of state formulas have been proposed to represent $Z(\rho)$.[11,12] A common feature of most of these is that the first few terms are taken to be the separately calculated virial coefficients which give very good agreement with the simulation data at low densities. However, their performance deteriorates substantially at higher densities near fluid-solid coexistence. Figure 1 shows the differences between the MD generated compressibility factors and several accurate EoS and the widely used simple Carnahan-Starling (CS) EoS. As may be seen the EoS of Kolafa, Labík and Malijevský (KLM),[13] reproduces the simulation data remarkably well across the entire fluid region (typically to an accuracy of $\sim 10^{-4}$). This unique feature of the KLM formula was noted previously by Bannerman *et al.*,[14] and is an exception to the usual trend. The KLM EoS is a semi-empirical equation fitted to the MD simulation data of $N = 13500$ spheres in which the appropriate error corrections are included.[13]

Table 1 The density dependence of the compressibility factor, Z, for the hard sphere fluid and metastable region obtained by MD simulations carried out in this work. Values in parenthesis are the standard deviations of the last two digits

ρ	Z	ρ	Z
0.050	1.11191724(54)	0.924	11.681268(35)
0.075	1.17368377(84)	0.926	11.762165(34)
0.100	1.2397199(12)	0.928	11.843840(38)
0.125	1.3103543(17)	0.930	11.926137(34)
0.150	1.3859479(23)	0.932	12.009235(33)
0.175	1.4668928(35)	0.934	12.092987(40)
0.200	1.5536043(45)	0.936	12.177520(41)
0.225	1.6465633(56)	0.938	12.262842(35)
0.250	1.7462744(47)	– metastable fluid below –	
0.275	1.8532923(48)	0.940	12.348898(37)
0.300	1.968231(11)	0.942	12.435706(38)
0.325	2.0917610(63)	0.944	12.523358(37)
0.350	2.2246014(90)	0.946	12.611809(42)
0.375	2.367604(11)	0.948	12.701034(37)
0.400	2.521620(13)	0.950	12.791119(40)
0.425	2.6876590(64)	0.952	12.882011(40)
0.450	2.866803(14)	0.954	12.973800(43)
0.475	3.060275(13)	0.956	13.066362(48)
0.500	3.269404(11)	0.958	13.159807(45)
0.525	3.495687(12)	0.960	13.254192(41)
0.550	3.740771(13)	0.962	13.349378(52)
0.575	4.006506(16)	0.964	13.445579(48)
0.600	4.294977(18)	0.966	13.542609(50)
0.625	4.608496(18)	0.968	13.640669(46)
0.650	4.949656(20)	0.970	13.739576(56)
0.675	5.321367(23)	0.972	13.839472(62)
0.700	5.726960(29)	0.974	13.940438(58)
0.725	6.170145(25)	0.976	14.042271(58)
0.750	6.655087(32)	0.978	14.145144(62)
0.775	7.186613(22)	0.980	14.249059(62)
0.800	7.770090(34)	0.982	14.35401(13)
0.825	8.411741(29)	0.984	14.459996(92)
0.850	9.118659(55)	0.986	14.56712(11)
0.875	9.899059(48)	0.988	14.67525(10)
0.890	10.406592(25)	0.990	14.78449(12)
0.892	10.476619(28)	0.992	14.89475(11)
0.894	10.547264(25)	0.994	15.00623(11)
0.896	10.618508(29)	0.996	15.11869(12)
0.898	10.690330(28)	0.998	15.23240(14)
0.900	10.762722(27)	1.000	15.34732(12)
0.902	10.835762(26)	1.002	15.463245(89)
0.904	10.909422(28)	1.004	15.580454(89)
0.906	10.983652(29)	1.006	15.698727(84)
0.908	11.058586(30)	1.008	15.818126(71)
0.910	11.134103(26)	1.010	15.938796(87)
0.912	11.210279(32)	1.012	16.060594(80)
0.914	11.287115(28)	1.014	16.18361(10)
0.916	11.364582(30)	1.016	16.30766(12)
0.918	11.442737(30)	1.018	16.43289(11)
0.920	11.521605(32)	1.020	16.55921(24)
0.922	11.601085(33)		

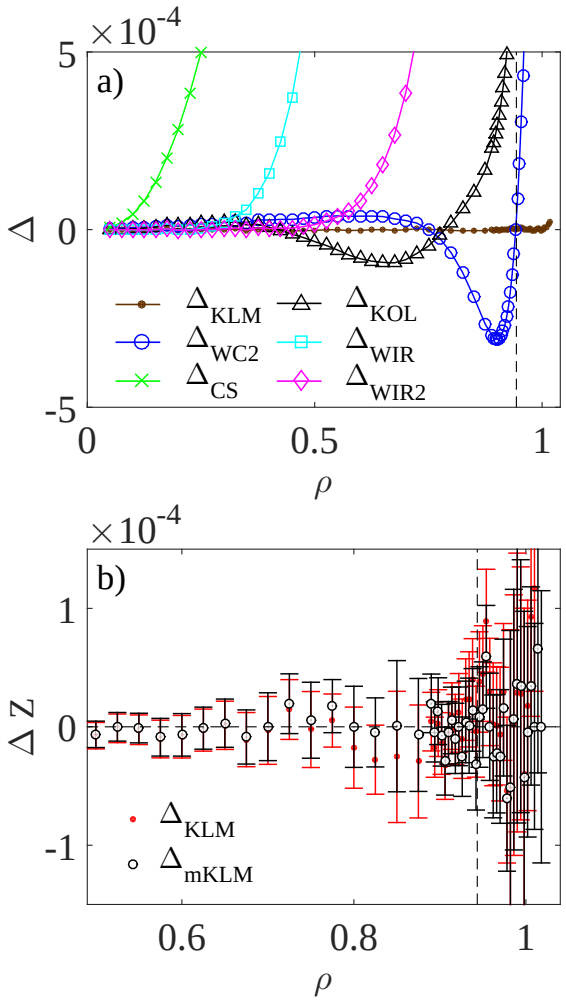

Fig. 1 (a) The difference between the simulation results for the compressibility factor from this work, and the predictions of various hard sphere fluid equations of state, defined as $\Delta = Z_{MD} - Z_{analytic}$. Key: Δ_{KLM} - Kolafa, Labík and Malijevský,[13]; Δ_{WC2} - Bannerman *et al.*,[14]; Δ_{CS} - Carnahan-Starling,[15]; Δ_{KOL} - Kolafa,[16,17]; Δ_{WIR} - virial equation (with B_8),[18]; Δ_{WIR2} - virial equation (with B_{12}),[19]. In the lower frame (b) are shown the deviations of: (i) the MD data of this work and the KLM EoS given in Eq. (3) which is labelled, $\Delta_{KLM} = Z_{MD} - Z_{KLM}$ (red solid dots), and (ii) the MD data of this work, and the mKLM formula from Eq. (4), which is given the symbol, $\Delta_{mKLM} = Z_{MD} - Z_{mKLM}$ (black open symbols). The standard error bars are given for the two cases. The vertical dashed line denotes the freezing density.

For the range of applicability, $\rho \leq 1.03$, this EoS has the following analytic form,

$$\begin{aligned} Z_{KLM} &= 1 + 4x + 6x^2 + 2.3647684x^3 - 0.8698551x^4 \\ &\quad + 1.1062803x^5 - 1.095049x^6 + 0.637614x^7 \\ &\quad - 0.2279397x^{10} + 0.1098948x^{14} - 0.00906797x^{22}, \end{aligned} \tag{3}$$

where $x = \zeta/(1-\zeta)$ and $\zeta = \pi\rho/6$ is the packing fraction. Using the MD data given in Table 1 obtained for $N \sim 10^6$ particles and a large number of state points the following slight modification of

the KLM EoS is,

$$\begin{aligned} Z_{mKLM} &= 1+4x+6x^2+2.3647684x^3-0.8698551x^4 \\ &+1.1062803x^5-1.1014221x^6+0.66605866x^7 \\ &-0.03633431x^8-0.20965164x^{10} \\ &+0.10555569x^{14}-0.00872380x^{22}, \end{aligned} \quad (4)$$

which gives an even better representation of the HS fluid EoS up to $\rho = 1.02$. In the above modified KLM equation (called 'mKLM'), the first six terms are exactly the same, one extra term is added and the values of the higher order term coefficients are changed. The performance of both equations is shown in the lower frame, (b) of Fig. 1 (note the small abscissa range). In what follows the mKLM expression in Eq. (4) is used as the default analytic EoS, unless stated. The accurate simulation data can provide insights into the behavior of the compressibility factor through its density derivatives. It is worth stressing that the second derivative calculated directly from MD data is captured by the mKLM formula. As the EoS has been determined accurately up to its second derivative any EoS feature of the fluid branch connected with the first or second derivatives of Z should be revealed with the mKLM formula given in Eq. (4). Also, this reparametrized EoS enables more conclusive statements on the behavior of other derived thermodynamic quantities to be made, which are covered in Sec. 3.4.

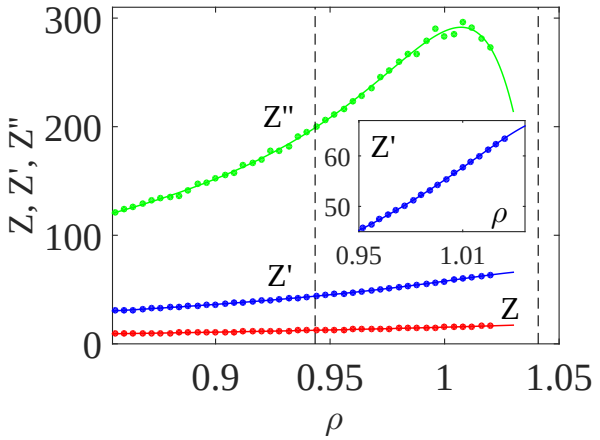

Fig. 2 The compressibility factor, Z and its first and second density derivatives obtained from the modified mKLM formula in Eq. (4) (solid lines) and by numerical differentiation of the original MD data (symbols). The two vertical dashed lines denote the approximate freezing and melting densities. In the inset the first density derivative is shown in the region close to the point of inflection.

Figure 2 presents the density dependence of Z, and its first, $Z' = dZ/d\rho$ and second, $Z'' = d^2Z/d\rho^2$ density derivatives. The mKLM EoS reproduces the Z'' maximum produced by the MD simulations. Up to the second derivative perhaps the only notable feature is in the metastable region which exhibits a point of inflection in the first derivative of Z. This is located at a density of, $\rho_I = 1.007$ ($\zeta = 0.527$) which corresponds to the limiting density of the metastable HS fluid $\rho \approx 1.012$ ($\zeta \approx 0.530$) recently reported by Isobe and Krauth,[20] above which in simulations the fluid started to form fcc nuclei. The present calculations with $N > 10^4$ particles indicate that just up to this density there is a high probability that the system will persist along the fluid branch. For $\rho > \rho_I$ this probability is substantially lower and decreases rapidly with increasing density. Nucleation is a process dominated by statistical fluctuations and even for smaller systems in some simulations it was found that the system could be taken gradually to the melting density without nucleating. Previous workers have also found that the probability of freezing in the metastable fluid region depends on the length of the simulation, the density and the system size.[13,21,22] The important new observation here is that the inflection point or maximum in the second derivative of Z in the metastable region can be assigned to a well defined process, i.e., the start of a high probability of nucleation of solid nuclei.

3.2 Solid branch EoS

In contrast to the fluid branch, the number of proposed equations of state for the HS solid is rather small,[23,24]. A notable one is the equation of state of Speedy,[25] referred to as Z_S here,

$$Z_S = \frac{3}{1-w} - \frac{A(w-B)}{w-C}, \quad (5)$$

where $w = \rho/\rho_{cp}$, $\rho_{cp} = 1.4142$ is the fcc close packed density, and $A = 0.5921$, $B = 0.7072$ and $C = 0.601$ are the equation constants obtained by fitting Eq. (5) to HS MD simulation data available at the time. The first term takes into account the divergence of Z at $\rho = \rho_{cp}$, which is represented in Eq. (5) as a pole at ρ_{cp} in accordance with an expression in the limiting free volume equation of state.[26] Figure 3 shows the difference between the MD generated Z of this work and the prediction of various analytic equations of state for the hard sphere solid. As may be seen in Fig. 3, the simple empirical equation of state in Eq. (5) represents the solid branch remarkably well. Using the MD results in Table 2 revised values for the Speedy EoS constants were calculated to be, $A = 0.5864$, $B = 0.7085$ and $C = 0.6055$. This reparametrization of the Speedy EoS is referred to as, Z_{Sr}, which is seen to represent the simulation data with an accuracy better than 10^{-3}, even in the metastable solid in the coexistence region, seen on the left hand side of Fig. 3. Note that, at high densities near to ρ_{cp}, the Speedy formula given in Eq. (5) in both its Z_S and Z_{Sr} forms gives a small but persistent deviation from the simulation results. This means that the second empirical term may not to be the most appropriate representation of the limiting behavior of the HS solid branch. The inset in the figure clearly indicates that in this limiting region an exponential (rather than a rational function) form would be a better choice. Moreover, the remaining tiny residual part (i.e., $Z - 3/(1-w) - A\exp(B(1-w)) - E$) is a straight line with a large slope on a log-lin scale, which means that two well separated exponential functions are required to represent satisfactorily the HS solid branch. Taking into account these observations, a revised formula for the compressibility factor, Z_{S1}, is,

$$Z_{S1} = \frac{3}{1-w} + Ae^{B(1-w)} + Ce^{D(1-w)} + E, \quad (6)$$

where the coefficients are, $A = 0.061622$, $B = 6.151$, $C = 3.8437 \times 10^{-5}$, $D = 27.72$ and $E = -0.49541$. As may be seen in Fig. 3, the performance of the above equation is slightly better than Eq. (5) and the limiting high density behavior is now much better accounted for. Recently it has been argued that the free-volume equation for the HS crystal should also contain an extra term,[26] which yields an additional free-volume contribution to the compressibility factor of the form, $1/(c_1 w + c_2)$ where c_1 and c_2 are constants. A modification of Z_{S1} for the solid branch incorporating this extra term, referred to as Z_{S2}, is

$$Z_{S2} = \frac{3}{1-w} + \frac{1}{c_1 w + c_2} + A e^{B(1-w)} + C e^{D(1-w)} + E. \qquad (7)$$

The equation Z_{S2} was found to give a better agreement with the MD data if $c_1 = 1$, $c_2 = 3/2$, $A = 0.025882$, $B = 8.689$, $C = 3.5433 \times 10^{-6}$, $D = 34.377$, and $E = -0.85973$ are used. The above equation with these parameters yields Z values that, as revealed in ΔZ, are practically indistinguishable from the simulation data. The expression Z_{S2} is used in the calculations of the free energy reported below.

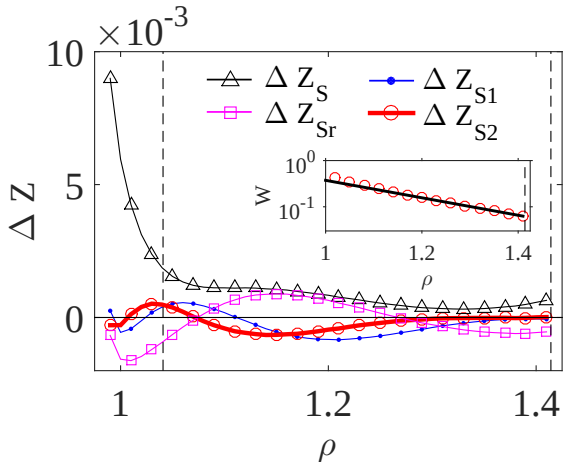

Fig. 3 The difference, ΔZ, between the compressibility factor Z from MD and the prediction of the equation of state for the hard sphere solid. The vertical dashed lines on left and right denote the approximate melting and close packing densities, respectively. The black open triangles (Z_S) are from Eq. (5) with the original coefficients of Speedy, the magenta open squares are from Eq. (5) with the new set of coefficients (i.e., Z_{Sr}). The blue solid points are Z_{S1} from Eq. (6), and the red open circles are Z_{S2} from Eq. (7). In the inset the convergence of the divergent term, $W = Z - 3/(1-w) - E$ derived from Eq. (7) to maximum close packing is shown.

3.3 Fluid-solid coexistence region

Establishing the fluid-solid coexistence boundary parameters of the HS system has been the subject of considerable interest since at least the 1950s, and the pioneering MD simulation work of Alder and Wainright.[27] The location of the transition pressure and densities is not a straightforward task however and consequently there is some scatter in the literature results even today. Finite size effects,[28,29] do influence significantly the parameters characterizing the fluid-solid boundary, which can be reduced using large N systems, as in this work. Thermodynamic integration,[30] has been used to establish the phase boundary parameters. Accurate representations of the compressibility factor of the HS fluid and solid branches in analytic form are presented in Eqs. (4) and (7), respectively. Consequently, the Helmholtz free energy, F, at any density on the phase diagram can be obtained directly by thermodynamic integration if the free energy of fluid and solid reference states are known. The chemical potential can be obtained from the free energy using, $\mu/k_BT = F/Nk_BT + Z$. Fluid-solid coexistence is determined using the simultaneous, P and μ equality condition.[30] For the fluid,

$$\frac{F(\rho)}{Nk_BT} = \frac{F(\rho)_{ideal}}{Nk_BT} + \int_0^\rho \frac{Z(\rho')-1}{\rho'} d\rho', \qquad (8)$$

where the mKLM EoS given in Eq. (4) was used for Z and the reference state used is the ideal gas,

$$\frac{F(\rho)_{ideal}}{Nk_BT} = \ln(\rho\Lambda^3) - 1 + \frac{\ln(2\pi N)}{2N}, \qquad (9)$$

where the Stirling formula is used for $N!$ and the thermal de Broglie wavelength term Λ is replaced by σ^3. For the solid phase the free energy was calculated using,

$$\frac{F(\rho)}{Nk_BT} = \frac{F(\rho_{ref})}{Nk_BT} + \int_{\rho_{ref}}^\rho \frac{Z(\rho')}{\rho'} d\rho', \qquad (10)$$

where ρ_{ref} is the reference state number density, $F(\rho_{ref})$ is the Helmholtz free energy at that density, and Z is Z_{S2} defined in Eq. (7). The Helmholtz free energy in the fcc crystal has been calculated from a number of reference densities. The free energy value for the density, $\rho = 1.04086$ has been determined in several works (mainly with the Frenkel and Ladd method,[31] or its

Table 2 The compressibility factor, Z, of the hard sphere solid and metastable region obtained in these MD simulations. Values in parenthesis are the standard deviations of the last two digits

ρ	Z	ρ	Z
0.99	10.053019(61)	1.20	19.4684310(43)
1.00	10.249112(19)	1.21	20.4311538(56)
1.01	10.463778(14)	1.22	21.4941538(53)
1.02	10.696236(11)	1.23	22.6736404(49)
1.03	10.946568(10)	1.24	23.9895836(52)
– metastable solid above –		1.25	25.4668429(49)
1.04	11.215150(15)	1.26	27.1367198(36)
1.05	11.502723(13)	1.27	29.0392033(48)
1.06	11.810217(10)	1.28	31.2262022(35)
1.07	12.138823(12)	1.29	33.7663683(39)
1.08	12.4899835(59)	1.30	36.7524001(63)
1.09	12.8654017(74)	1.31	40.3125530(32)
1.10	13.2670527(65)	1.32	44.6295693(22)
1.11	13.6972133(29)	1.33	49.9729990(38)
1.12	14.1585807(87)	1.34	56.7576817(32)
1.13	14.6541871(50)	1.35	65.6568583(32)
1.14	15.1875924(55)	1.36	77.8405325(38)
1.15	15.7629452(83)	1.37	95.5371959(43)
1.16	16.3850184(42)	1.38	123.580746(54)
1.17	17.0593924(56)	1.39	174.790587(78)
1.18	17.7926751(35)	1.40	298.062562(53)
1.19	18.5926092(48)	1.41	1006.468451(33)

modifications) giving an average value of, 4.9590 to an accuracy of 0.0002,[28,29,32,33]. Vega and Noya determined the free energy also at $\rho = 1.099975$ to be 5.631(1) and at a density of 1.15 to be 6.273(2).[28]

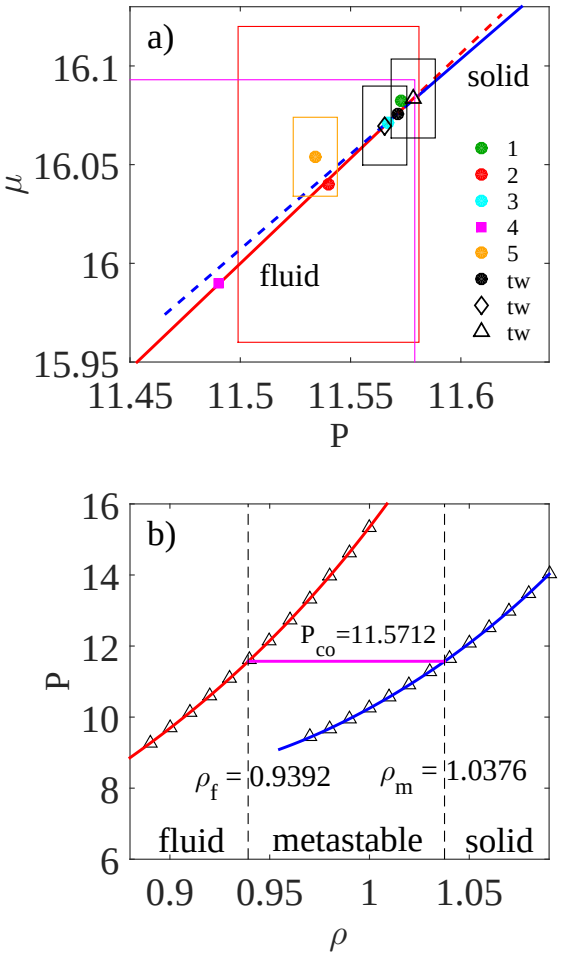

Fig. 4 (a) The chemical potential *vs.* pressure of the hard sphere system in the coexistence region. The crossing point of the red and blue lines specifies the coexistence pressure, P_{co} and chemical potential, μ_{co} and their dependence on the reference density of the solid in Eq. (10). Three reference points are described in the main text. The different colored points represent the coexistence pressure and chemical potential from the literature: 1 - Fernández et al.,[34]; 2 - Vega and Noya,[28]; 3 - Frenkel and Smit,[30]; 4 - Sweatman,[35]; 5 - Ustinov,[36] and tw - this work (for different reference points). The boxes indicate the statistical uncertainty of each estimate. The chemical potential value for Fernández *et al.* was estimated here with P_{co} obtained in that paper and using the mKLM EoS. Frame (b) shows the pressure as a function of density in the fluid-solid transition region.

The resulting chemical potential *vs.* pressure dependence for the two phases in the coexistence region is shown in Fig. 4(a) using data from this work. The intersection point gives the coexistence pressure, P_{co} and chemical potential, μ_{co} values. In addition different values for the intersection point from other works together with their statistical uncertainties are also given on the figure. As may be seen the crossover point determined here depends slightly on the reference density used in Eq. (10). In particular, $P = 11.5712(10)$ for $\rho_{ref} = 1.04086$ (filled black circle), $P = 11.566(10)$ for $\rho_{ref} = 1.099975$ (open diamond) and $P = 11.578(10)$ for $\rho_{ref} = 1.15$ (open triangle). These predictions mutually agree within their uncertainty limits. As the most firmly established reference free energy value is that for $\rho_{ref} = 1.04086$, our best estimate of the location of the fluid-crystal HS transition is where, $P_{co} = 11.5712(10)$, $\rho_f = 0.9392(1)$, $\rho_m = 1.0376(1)$, and $\mu_{co} = 16.0758(20)$. The main results established in this work for the coexistence region are shown graphically on the P against ρ diagram in Fig. 4(b). The obtained HS coexistence pressure and densities are in very close agreement with those reported by Frenkel and Smit ($P_{co} = 11.567$, $\rho_f = 0.9391$, $\rho_m = 1.0376$).[30]

Table 3 The coexistence values for various properties of the hard sphere system from this work. Values in parenthesis are the standard deviations of the last one or two digits

Property	Fluid	Solid
density	0.9392(1)	1.0376(1)
packing fraction	0.4918(1)	0.5433(1)
pressure	11.5712(10)	11.5712(10)
chemical potential	16.0758(20)	16.0758(20)
heat capacity	4.374(1)	
compressibility	0.01895(1)	

A summary of some of the coexistence property values is given in Table 3. To date, the most accurate estimate of the coexistence pressure from other work is $P_{co} = 11.5727(10)$,[34] which was determined using the tethered Monte Carlo approach. It is noteworthy that the calculations performed in this work by a different method yields practically the same value within the same level of accuracy. The very good agreement of P_{co} obtained by the two routes provides strong extra confirmation of the accuracy of the value of P_{co} derived here. Examples of different estimates of the location of the coexistence values (P_{co}, μ_{co}) are shown in Fig. 4(a). Figure 4(a) indicates that the location of coexistence point converges towards a relatively small (P, μ) region on the plot, which can be considered to be practically a point.

3.4 Thermodynamic properties

Basic second-order thermodynamic properties of the HS fluid can be obtained as the limiting case of the expressions for the soft sphere fluids given in Table I in Ref.[37]. All of these quantities are expressed in terms of Z and its first density derivative, $Z' = dZ/d\rho$ only. For example, the isobaric heat capacity, $C_P/Nk_B = 3/2 + Z^2/(Z + \rho Z')$, and the isobaric bulk modulus is, $B_S = B_T + 2\rho k_B T Z^2/3$, where $B_T/\rho k_B T = Z + \rho Z'$ is the isothermal bulk modulus. Thus, an accurate representation of the HS thermodynamic quantities such as the bulk moduli, heat capacity, volume expansion coefficient, isothermal pressure coefficient, Grüneisen parameter, Joule-Thompson coefficient, and sound velocity can be obtained directly from this work using the compressibility factor and its density derivatives. These data lead to the conclusion that up to the freezing density, ρ_f, all of these quantities evolve monotonically. As already noted, above the coexis-

tence density, the second derivative, Z'', has a maximum in the metastable region (see Fig. 2) and therefore some related features in the thermodynamic properties might be expected in the metastable region for the density range, $\rho > \rho_f$. Figure 5 shows that C_P and the isothermal compressibility, $\chi = 1/B_T$, display a point of inflection in their first derivative. The figure therefore indicates that just above the freezing density the HS system has a maximum in the second derivative of the specific isobaric heat capacity and in the compressibility.

It has been shown recently,[38] that for hard spheres, the expression $(\chi/2\rho k_B T)'$ represents exactly the thermodynamic curvature parameter, R, which was introduced by Ruppeiner as a basic invariant in thermodynamics.[39–41] This quantity is defined in terms of derivatives of the free energy, and has units of volume. It has been suggested that the magnitude of R could be a measure of the size of mesoscopic organized structures in the system.[39–42] The R parameter has been studied for different molecular systems and it has been put forward as a property that could be used to delineate various regions on the fluid part of the phase diagram.[43,44] In general, however, the R quantity is not easy to compute and its physical interpretation is still debated. In contrast, the compressibility is a more readily appreciated and accessible physical property than R. Thus, the established simple explicit expression for R of the HS system in terms of χ may also provide a useful reference system or surrogate for further studies of R in other model systems.

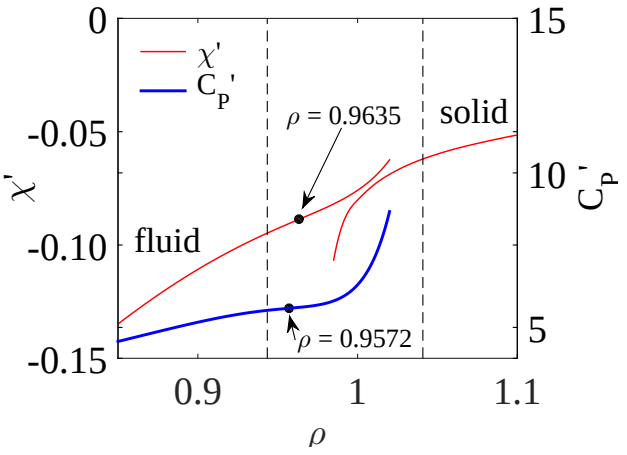

Fig. 5 The second order thermodynamic properties of the HS system in the high density range which includes the metastable region. The vertical dashed lines from left to right denote the approximate freezing and melting densities of the hard sphere system, respectively. In the figure the thin solid red line represents the first derivative of χ for the fluid branch calculated from the mKLM EoS in Eq. (4) and for the solid branch calculated from the Z_{S2} EoS given in Eq. (7). The bold blue line represents the first derivative of the isobaric heat capacity, C_P, calculated from the mKLM equation of state. The black solid points indicate the points of inflection in these two quantities.

4 Dynamical properties

The self-diffusion coefficient, D, and shear viscosity, η, are key quantities in understanding the dynamics and molecular-level relaxation processes in liquids. In the infinite dilution limit they are given by kinetic theory of the HS fluid,[45]

$$D_0 = \frac{3}{8\rho\sigma^2}\left(\frac{k_B T}{m\pi}\right)^{1/2}, \quad \eta_0 = \frac{5}{16\sigma^2}\left(\frac{mk_B T}{\pi}\right)^{1/2}. \tag{11}$$

The Enskog value,[45] for the self-diffusion coefficient is given by

$$D_E = 1.01896 D_0 \frac{b_2\rho}{Z-1}, \tag{12}$$

and the HS viscosity is,

$$\eta_E = \eta_0 b_2 \rho \left(\frac{1.016}{Z-1} + 0.8 + 0.7737(Z-1)\right), \tag{13}$$

where $b_2 = 2\pi\sigma^3/3$ is the second virial coefficient of the hard sphere fluid. In the following analysis the mKLM EoS was used for Z in Eqs. (12) and (13). For dilute systems the above formulas are accurate but the D and η values of model systems from these formulas increasingly deviate from the simulation values with increasing density. Any simulation is carried out for a finite number of particles and thus the thermodynamic limit is estimated by extrapolation of the data to infinite system size. For many physical quantities all finite-size corrections to leading order vary as $O(N^{-1})$ and thus an accurate estimate of their limiting thermodynamic value is straightforward for relatively modest system sizes.[10] However, in the case of the self-diffusion coefficient the size-dependence is much stronger ($\sim N^{-1/3}$, see below) and therefore more problematic to characterize. Much larger system sizes are required to perform a meaningful extrapolation to infinite system size. Even in the case of such an intensely studied system as the HS fluid, the forms of the N-dependent self-diffusion coefficient, $D_N(\rho)$ and the thermodynamic-limit values, $D_\infty(\rho)$, have still not been determined definitively. We first consider the self-diffusion coefficient below.

4.1 Self-diffusion coefficient

System size corrections must be applied to the calculated self-diffusion coefficients for all liquids owing to the significant difference between $D_{N\to\infty}(\rho)$ and the values of D obtained with values of $N \sim 10^4$. The N-dependence of the self-diffusion coefficient has been the subject of investigations for a range of liquids.[28,46–56]

From these studies, it has been firmly established by hydrodynamic theory and simulations that the effects of periodic images for sufficiently large N can be represented by the following analytic expression,

$$D_N = D_\infty - \frac{\lambda}{L} \tag{14}$$

where L is the cubic simulation cell side length, D_∞ is the self-diffusion coefficient of the infinitely sized system and λ is a constant which is independent of the system size. Thus, to leading order, the finite-size corrections are inversely proportional to the side length of the cubic simulation cell or equivalently $N^{-1/3}$. The key problem is that it is not known in advance what value of N or L can be considered to be 'sufficiently large' for this limiting linear regime to apply. Moreover, the value of N (or 'N_s') where the linear regime starts, usually cannot be established precisely

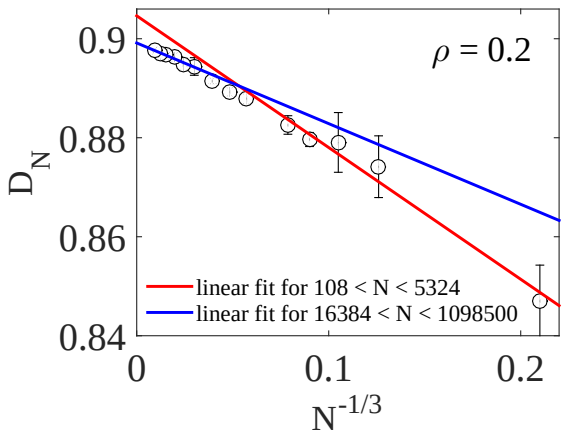

Fig. 6 The MD HS self-diffusion coefficient values, D_N, against $N^{-1/3}$ (black open circles) for the density $\rho = 0.2$. The red line is a linear fit to the MD data for systems where the number of particles is in the range, $108 < N < 5324$. The blue line represents a linear fit to the data for the larger systems systems in the range, $16384 < N < 1098500$. The D_N values shown on the figure are for systems composed of 108, 500, 864, 1372, 2048, 5324, 8788, 16384, 37044, 70304, 131072, 275684, 530604 and 1098500 hard spheres.

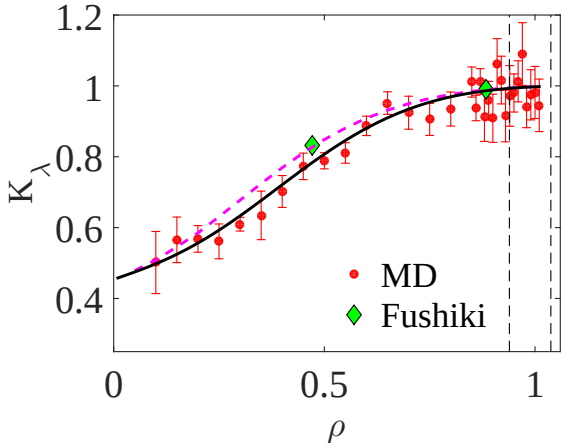

Fig. 7 Density dependence of the correction factor K_λ in Eq. (15). The K_λ-values (red solid points) are obtained from the linear fit to the D_N data with Eq. (14) for systems $16384 < N < 1098500$ at densities listed in Table 4. The solid black line is the function $K_\lambda = \nu/(\nu + D_\infty)$ described in main text. The dashed magenta line is the function, $K_F = \nu/(\nu + D_E)$. The two green diamond symbols represent Fushiki's correction,[46] and are explained in the main text.

and may depend on various factors such as the system density and the type of interparticle interaction.

Figure 6 for $\rho = 0.2$ shows that in the case of hard spheres the linear regime develops only for systems composed of more than $N_s \sim 10^4$ particles. This is rather unexpected as usually systems of a few thousands particles are considered to be sufficiently large to be essentially in the thermodynamic limit for other properties. The figure also demonstrates that by applying the formula in Eq. (14) to systems with $N < N_s$ a 'limiting value' value of D is predicted which in general is different from the true D_∞. In addition, such a 'limiting value' can depend on the number and range of points chosen for the extrapolation. This is not always the case, however, as for not too small systems and/or not too low densities the linear correction when applied to the range of states where $N < N_s$ can produce a limiting value for D which is close to the true D_∞ within the statistical uncertainty of the simulations. This indicates that the linear correction formula given in Eq. (14) can represent the trend in the data for a range of N values that is relatively small. To be certain one is in the correct linear regime is a very demanding task which is computationally impracticable in many cases, and it is tempting to estimate D_∞ from Eq. (14) using a range of N which is generally too small to give an accurate value for D_∞. It has also been pointed out recently,[52] that for certain types of real systems, such as ionic liquids, the size dependence has to be adequately addressed (the linear regime has to be reached) and simulations of large systems are necessary. Therefore, to summarize, the optimum criteria for extrapolating the $D_N(\rho)$ data to the infinite system size limit are still not firmly established.

The second issue that needs to be considered is what system parameters the coefficient λ in Eq. (14) depends on. From hydrodynamic analysis λ is equal to $k_B T\xi/6\pi\eta$ where the constant $\xi = 2.837297$ and η is the solvent viscosity.[49,56] We denote λ in the hydrodynamic approximation by λ_H. For large N the size dependence of η is irrelevant (as it is negligible anyway) and thus λ_H is independent of the system size. The data for D_N in the linear regime indicates that in general $\lambda \neq \lambda_H$, i.e., the slope does not agree with the hydrodynamic prediction. At higher densities (close to the freezing) $\lambda \approx \lambda_H$ but on decreasing density the difference increases considerably. We consider this to be an important finding as it demonstrates that other sources of N dependence, apart from those derived on purely hydrodynamic grounds may also be important. Such a possibility was noted by Heyes *et al.*[47] who treated λ as a density dependent empirical parameter instead of using λ_H. Also, a similar size dependence of the self-diffusion coefficient was observed by Fushiki,[46] for the HS fluid ($N \leq 16384$) at $\rho = 0.88$ and 0.47. This was attributed to aspects of the diffusion mechanism of a molecular scale solute particle, such as the long-time behavior of the velocity autocorrelation function, which are not included in the purely hydrodynamic treatment leading to λ_H. A corrected formula for the slope was proposed in Ref.[46] with the form, $K_F\lambda_H$ with the factor, $K_F = \nu/(\nu + D_E)$, where $\nu = \eta/\rho m$ is the kinematic viscosity and D_E is Enskog's value for D given in Eq. (12). Recently, deviations from the λ_H slope have also been reported for HS binary mixtures.[53]
The D_N data in the linear regime enable $\lambda(\rho)$ and D_∞ to be determined, as well as establishing the form of the correction factor, $K_\lambda = \lambda/\lambda_H$. This correction factor, shown in Fig. 7, is found to be well represented by a simple modification of the Fushiki's correction (the solid black line in the figure) which in turn gives the following explicit formula for the scaling constant in Eq. (14),

$$\lambda = K_\lambda \lambda_H = \left[\frac{\nu}{\nu + D_\infty}\right] \frac{\xi k_B T}{6\pi\eta}. \qquad (15)$$

Note that this is basically Fushiki's correction in which D_E has

been replaced by D_∞ (the two formulas converge when $\rho \to 0$). This provides a route to improve the hydrodynamic correction factor.

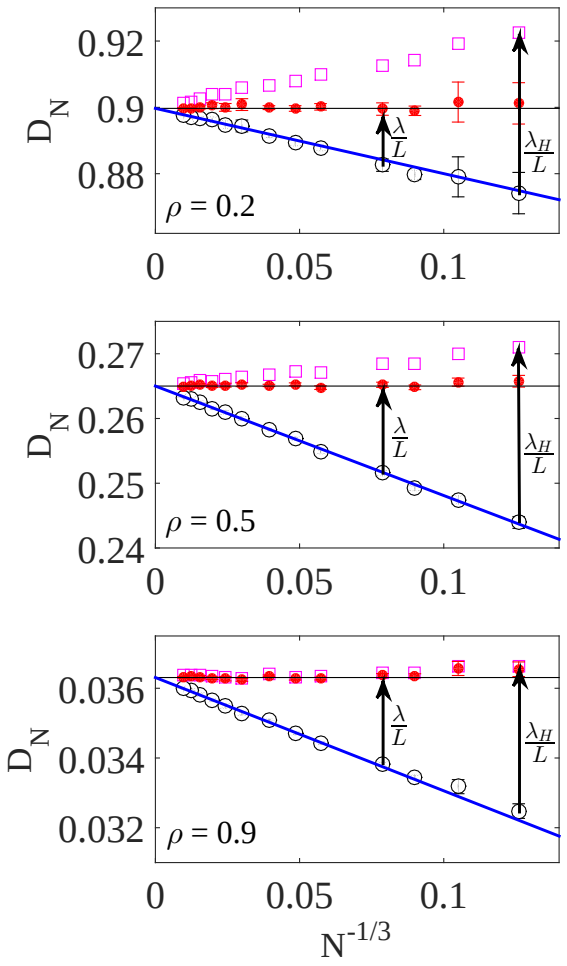

Fig. 8 The MD self-diffusion coefficient, D_N against $N^{-1/3}$ (black open circles). The frames from top to bottom are, respectively, for a dilute system ($\rho = 0.2$), an intermediate density ($\rho = 0.5$) and dense ($\rho = 0.9$) HS fluid. The blue line in each frame represents the linear fit with Eq. (14) for large systems composed of $16384 < N < 1098500$ particles. The horizontal solid black line represents the self-diffusion coefficient in thermodynamic limit, D_∞. The red solid points are the $D_\infty = D_N + \lambda L^{-1}$ with λ as given in Eq. (15) and the magenta open square symbols represent the $D_\infty = D_N + \lambda_H L^{-1}$. The MD data points considered are for 500, 864, 1372, 2048, 5324, 8788, 16384, 37044, 70304, 131072, 275684, 530604 and 1098500 particles.

The set of D_N corrected with the λ and λ_H factors is shown in Fig. 8 for densities in the dilute, intermediate and dense HS fluid ranges. Following on from Fig. 7, and as may be seen using λ_H in Eq. (14) the formula overestimates the limiting self-diffusion coefficients practically over the entire (ρ, N) domain of N values. At higher densities the correction factor becomes nearly unity and λ_H can be used for λ. As may be seen in Fig. 8 all reasonable values of λ will give the correct value of D_∞ for large enough N. The objective of using Eq. (14) is to obtain D_∞ from the smallest values of N as possible. The correction $\lambda = K_\lambda \lambda_H$ for low and intermediate densities gives the correct D_∞ for a range of smaller N than using the hydrodynamic treatment (i.e., $\lambda = \lambda_H$).

Table 4 Self-diffusion coefficient (in units of $\sigma(k_BT/m)^{1/2}$) and shear viscosity (in units of $\sigma^{-2}(mk_BT)^{1/2}$) of the hard sphere fluid and metastable region in the thermodynamic limit. Values in parenthesis are the standard deviations of the last two digits

ρ	D_∞	η
0.10	1.94786(56)	0.1950(10)
0.15	1.24636(43)	0.21076(85)
0.20	0.89989(26)	0.22989(96)
0.25	0.69308(35)	0.25837(31)
0.30	0.55523(13)	0.29385(41)
0.35	0.45527(42)	0.33893(67)
0.40	0.37845(25)	0.39527(73)
0.45	0.31663(18)	0.46684(91)
0.50	0.265009(96)	0.5573(11)
0.55	0.221105(98)	0.6708(13)
0.60	0.183259(83)	0.8198(18)
0.65	0.150317(81)	1.0094(25)
0.70	0.121081(97)	1.2674(46)
0.75	0.095425(75)	1.6280(66)
0.80	0.072934(61)	2.1577(70)
0.85	0.053298(40)	3.006(43)
0.86	0.049608(32)	3.224(48)
0.87	0.046203(30)	3.474(66)
0.88	0.042720(53)	3.742(67)
0.89	0.039478(39)	4.059(74)
0.90	0.036254(44)	4.468(65)
0.91	0.033286(43)	4.917(66)
0.92	0.030326(37)	5.336(55)
0.93	0.027419(36)	5.932(86)
0.94	0.024746(38)	6.603(76)
0.95	0.022134(22)	7.338(98)
0.96	0.019695(20)	8.50(11)
0.97	0.017357(27)	9.77(16)
0.98	0.015082(16)	11.00(17)
0.99	0.013013(17)	12.76(23)
1.00	0.011053(14)	15.38(33)
1.01	0.009222(12)	18.16(36)

The resulting D_∞ are given in Table 4 and are compared with the values from previous studies in Fig. 9. The data are plotted in a normalized form by dividing by the Enskog self-diffusion coefficient, D_E, calculated from Eq. (12) for each density. It was reported in the previous studies that the density dependence of the normalized self-diffusion coefficient can be well represented by a polynomial which in the present study is,

$$D_\infty/D_E = 1 + a_1\rho + a_2\rho^2 + a_3\rho^3 + a_4\rho^4 + a_5\rho^5 + a_6\rho^6, \qquad (16)$$

where $a_1 = -0.0118979$, $a_2 = 4.32042$, $a_3 = -10.9115$, $a_4 = 17.2687$, $a_5 = -18.0587$, $a_6 = 6.7426$. Erpenbeck and Wood,[57] fitted their MD hard sphere D simulation data when extrapolated to infinite N by this polynomial where the coefficients $a_1, \cdots, a_6$ are [0.038208, 3.182808, −3.868772, 0, 0, 0] and Heyes *et al.*[47] gave the set of values, [0.277645, 3.98964, 26.49881, −134.0015, 110.7344, 0]. As shown in the inset of Fig. 9, the overall agreement between these three explicit expressions is quite good, the

differences in D_∞/D_E in the fluid phase being less than 0.03. Moreover, considering only the data for higher densities ($\rho > 0.8$) the HS self-diffusion coefficient can be very well (see Fig.10a) represented by the expression, $D = A_d(\rho_d - \rho)^d$, where $A_d = 0.58057$, $\rho_d = 1.1011$ and $d = 1.7282$. We find that the Vogel-Fulcher functional form, $D \propto \exp(A_{vf}/(\rho - \rho_d))$, where A_{vf} is a constant, does not fit the data as well in the same density range.

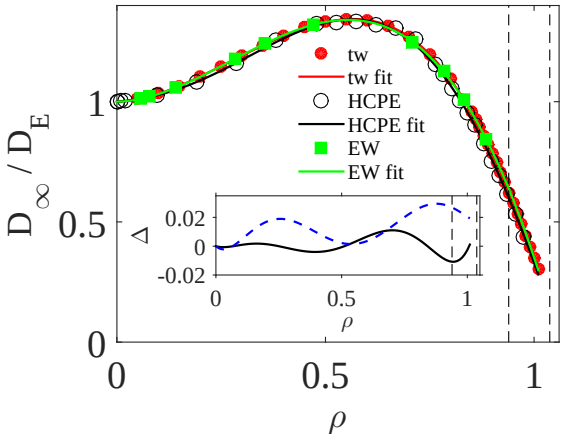

Fig. 9 The ratio, D_∞/D_E, from the MD simulations with D_∞ derived from Eq. (14) using the analytic formula for λ given in Eq. (15) (the red solid points) and the polynomial fit of Eq. (16) (solid red line). Key: EW simulation-derived data (square green symbol) and the polynomial fit (the solid green line) is taken from Erpenbeck and Wood.[57] HCPE, are extrapolated MD data of D (open black circles) and the polynomial fit (the solid black line) from Heyes *et al.*[47]. The inset shows the difference between D_∞/D_E of this work (tw) from the EW (solid black line) and HCPE (dashed blue line) data.

4.2 Viscosity

Studies on various fluid systems have verified that the shear viscosity is less dependent on system size than the self-diffusion coefficient. Nevertheless, it is natural to investigate quantitatively how significant is the N-dependence of the viscosity. This is an important question because η is a collective quantity and is therefore usually calculated with relatively small systems, for which the statistical uncertainties can still be restrictive. Recently this issue has been investigated by Kim *et. al.*[55] who performed detailed studies of Lennard-Jones (LJ) and star-polymer model systems at relatively high densities. They demonstrated that the shear viscosity in dense fluids exhibits strong and complex size effects in small systems, but without any overall systematic size dependence and that above a certain length scale, reliable shear viscosity values were obtained without any noticeable scaling behavior with system size.

Accurate MD HS shear viscosities carried out in this work also indicate that above a certain length scale there is no noticeable system size dependence and therefore systematic scaling behavior in the shear viscosity. Moreover this lack of the size dependence and size scaling is observed for large systems ($N > 10^4$) at all studied densities, and not only at high densities. The MD data also indicate that some size scaling occurs at low densities for small systems composed of less than about 5000 particles and this scaling appears to behave as, $\sim 1/N$.

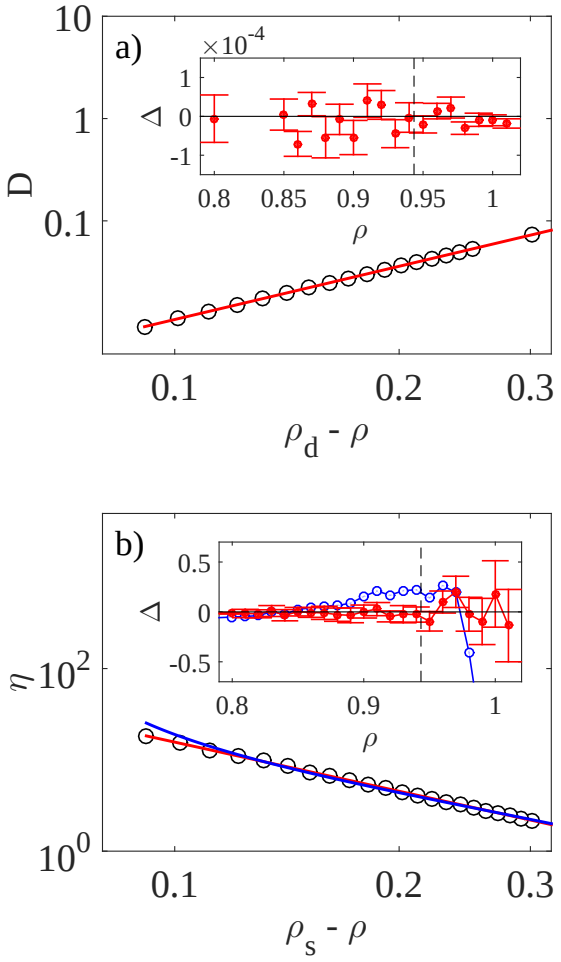

Fig. 10 (a) The HS self-diffusion coefficient, D_∞, (black open circles) as a function of $\rho_d - \rho$ near the freezing point in a log-log scale. The red line represents a fit to the MD data with the $D_\infty(\rho > 0.8) = A_d(\rho_d - \rho)^d$ formula. In the inset the difference between fit and D_∞ is shown on a linear scale (dashed line denotes the freezing density). (b) The HS fluid viscosity, η, as a function of $\rho_s - \rho$ near the freezing point is shown on a log-log scale. The open circles are the MD data for $N = 16384$ particles from this work. The red line represent the fit to the MD data with Eq. (17), and the blue line represents a fit to the MD data with Eq. (18). In the inset the differences between both fits and MD data are shown on a linear scale.

The calculated values of the HS shear viscosity are given in Table 4. With decreasing density the Enskog kinetic theory prediction for the shear viscosity, η_E, becomes increasingly more accurate. Therefore it is natural to express the viscosity normalized by the Enskog value, and the relationship, $\eta = \eta_E f(\rho)$ can therefore be used to express the actual viscosity in terms of the analytic Enskog expression at arbitrary fluid density. A possible empirical form of the correction factor, $f(\rho)$, is $f(\rho) = 1 + d_1\rho^{d_2}$, where $d_1 = 2588$ and $d_2 = 11.11$, was proposed in Refs.[47,48], which represents well the MD simulation data within the statistical uncertainties over nearly the entire fluid, except in the high density

region. We seek to improve the fitting of the viscosity data in this region too.

The following simple expression gives a quite accurate representation of the HS viscosity near freezing and in the metastable region

$$\eta_{KD} = \frac{A_s}{(\rho_s - \rho)^q}, \quad \text{for } \rho > 0.8. \tag{17}$$

An alternative formula, in which the inverse shear viscosity is linearly dependent on $1/\rho$, is

$$\frac{1}{\eta} = A_0 \left(\frac{1}{\rho_0} - \frac{1}{\rho} \right) \quad \text{for } \rho > 0.8, \tag{18}$$

with coefficients, $A_s = 0.25465$, $\rho_s = 1.1016$, $q = 1.7883$ and $A_0 = -1.5896$, $\rho_0 = 1.036$ obtained from a non-linear least squares fit to the MD viscosity values in Table 4. The obtained ρ_s from Eq. (17) is practically the same (within 0.001) as that obtained with the $D = A_d(\rho_d - \rho)^d$ formula for $D(\rho > 0.8)$.

The formula in Eq. (17) is of the Krieger-Dougherty (KD) equation form, which is widely used in the colloid literature to represent the packing fraction dependence of the shear viscosity and was used in Ref.[58] with q and ρ_s fitted to HS simulation data, and $A_s = \eta_0(6\rho_s/\pi)^q$. In that study the value of ρ_s was not taken to be the actual location of the singularity (which was difficult to estimate at the time) but adjusted to match the simulation data at lower packing fractions. Therefore the high density region still needs to be accounted for in any improved fitting to that expression.

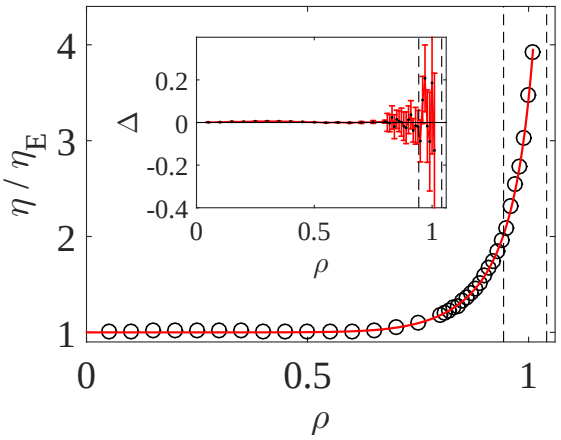

Fig. 11 The ratio η/η_E of the hard sphere system is shown. Figure shows MD data for $N = 16384$ particles (black open circles). The red line represent the analytic formula in Eq. (19). In the inset the difference between the analytic formula and MD data is shown.

The treatment in Eq. (18) for the viscosity dates back to Batschinski,[59] and Hildebrand,[60,61] and has also been exploited for soft sphere fluids.[62] As may be seen in Fig. 10 both of these simple expressions reproduce the simulation data of the dense HS fluid region quite well, apart from close to the melting line. The performance of Eq. (17) is better, probably in part because it has three disposable parameters rather than two. Nevertheless, the data on a log-log scale are consistent with a straight line which clearly indicates that Eq. (17) is fully consistent with the simulation data, and even extends into the metastable region (seen in the inset). It means also that the density $\rho_s = 1.1016$ may be considered to be the location of the singularity, which corresponds to $\zeta \approx 0.58$, the value close to that usually taken for the HS glass transition,[63–65] and it is also higher than the melting density (1.0376, see Table 3). It is always possible that more accurate viscosity values than we can practicably obtain in the range $\rho > 1.01$ would change the slope or even the form of the line in this limit. The red line in Fig. 10 gives us however no reason to expect such a deviation of the shown trend at higher densities.

The representations of the viscosity in the dilute (η_E) and dense regions (η_{KD}) can be combined with a 'bridging' function, $f_B(\rho)$ to represent accurately $\eta(\rho)$ over a wide density range,

$$\eta(\rho) = \eta_E f_B + (1 - f_B)\eta_{KD}. \tag{19}$$

The function, $f_B(\rho)$, must go from 0 for $\rho > 0.8$ and to 1 as $\rho \to 0$. An example of a function which satisfies these requirements is, $f_B(\rho) = 1/(1+\exp(-17.477+24.963\rho^{0.91283}))$. The inset in Fig. 11 shows that the formula in Eq. (19) represents the viscosity of HS fluid within the statistical uncertainties of the MD data. Also, as may be noticed in the inset, for $\rho < 0.4$, there is a small deviation in the form of a tiny regular 'hump'. This feature and its possible relation to the D/D_E behavior as a function of density shown in Fig. 9 may be worth further study.

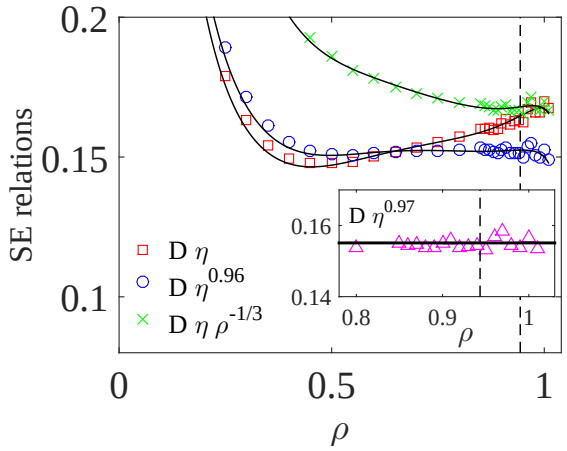

Fig. 12 Different forms of modified Stokes-Einstein relation for the equilibrium and metastable hard sphere fluid. The symbols are the MD data of this work and the solid lines are analytic functions for the self-diffusion coefficient in Eq. (16) and viscosity in Eq. (19). In the inset the fractional SE obtained with $p = d/q = 0.97$ (triangles) is shown and the solid line is the constant 0.155 obtained with the D and η formulas for $\rho > 0.8$.

The self-diffusion coefficient and viscosity data and their representations in Eqs. (16) and (19) can be used to test various semi-empirical relationships used in the literature to relate these two quantities, notably the Stokes-Einstein, SE, formula and more recent generalizations.[1,4] In Fig. 12 the original SE relation, $D \propto \eta^{-1}$ is shown along with the fractional SE,[66] $D \propto \eta^{-p}$ and the SE variant involving the density, $D \propto \rho^{1/3}\eta^{-1}$.[4,67] The metastable region up to $\rho_I \approx 1.01$ is considered. The best performance is obtained for the fractional SE,[68] with $p = 0.96$ i.e., the value close to that found previously by Heyes *et al.*[47]. The

product $D\eta^p$ is approximately constant, with a value of 0.152 for $\rho > 0.4$. Thus, the results indicate the fractional SE applies well also in the metastable region (at least up to the point of infection, ρ_I). Also, as $\rho_d \cong \rho_s$ the formulas for D and η for $\rho > 0.8$ yield exactly the fractional SE with $p = d/q = 0.97$ which is shown in the inset of Fig. 12. A value of $p = 0.975$ for HS data up to $\rho \simeq 0.955$ was obtained in Ref.[47] and it is shown here with more accurate data that this relation also applies to states in the metastable fluid region. Also, the results in Fig.12 clearly indicate that the product, $D\eta$, cannot be constant, which follows from isomorph theory and has been demonstrated in practice for the LJ system recently.[69] The SE modification with $\rho^{-1/3}$ replacing the diameter of the particle,[70] applies only approximately for the HS fluid and cannot be considered in general as a superior relation for hard sphere systems as suggested recently in Refs.[4,67].

5 Conclusions

Extensive new Molecular Dynamics calculations using large hard sphere systems performed to obtain accurate representations of thermodynamic and dynamical properties have aided in resolving a number of issues. A new formula for the compressibility factor, Z, of the HS solid branch is proposed, which represents the metastable region and takes into account its divergence at close packing. Also a reparameterization of the Kolafa, Labík and Malijevský fluid equation of state (EoS) is made. A maximum in the second derivative of Z, found in the metastable region correlates well with the start of a high probability of solid nucleation. A detailed study of basic second-order thermodynamic derivative properties of the HS fluid is performed. There is a maximum in the second derivative of the isobaric heat capacity and the compressibility in the metastable fluid region. Detailed knowledge of the HS compressibility throughout the whole fluid range obtained here could be useful in exploring Ruppeiner's thermodynamic curvature parameter,[39,40] as an exact relationship between these quantities has been pointed out here.

The thermodynamic parameters associated with the freezing-melting transition have been determined to four digit accuracy. Combined with the results obtained by Fernández *et al.*,[34] this work has generated accurate new values for the coexistence values of certain properties of the HS system which for the densities is in perfect agreement, to ± 1 in the last quoted digit (fourth or fifth significant figure), with the values given in Ref.[30]. The methodology in the two treatments was different as the tethered Monte Carlo approach and the thermodynamic integration method were applied. The mutual excellent agreement between the two routes provides additional confidence in the coexistence values determined in both studies. We consider that further improvement in these values should concentrate mainly on improving the accuracy of the reference point Helmholtz free energy of the HS crystal.

For the self-diffusion coefficient, D, it is shown that relatively large systems ($N > 10^4$) are required to achieve an accurate linear extrapolation of D to infinite size. This is a rather unexpected result as usually systems of a few thousands particles are considered to be sufficiently large to be in the linear regime for typical physical properties. This requirement is not entirely universal, however, as for not too small systems and not too low densities the linear correction outside the 'true' linear region can provide a quite satisfactory estimation of D in the infinite system size limit. In general, larger systems are required, as there is a density dependence of the value of the slope in the linear regime of $D(N^{-1/3})$ according to Eq. (15) of this work. The density dependent correction becomes practically insignificant at higher densities and the hydrodynamic formula found in the literature is still accurate. However, with decreasing density the density dependence of the size correction cannot be neglected, which indicates that other sources of N-dependence, apart from those derived on purely hydrodynamic grounds, may also be important. These trends would be worth further investigation in future work.

The established formula for the N-dependence of the self-diffusion coefficient in Eq. (15) provides a general form for the system size correction which is possible also applicable to other molecular different systems (not only to the HS), but this would need to be tested.

From the MD self-diffusion and viscosity data of this work, formulas representing their density dependence in the thermodynamic limit have been obtained. It has been shown here that the hard sphere inverse self-diffusion coefficient and shear viscosity near freezing and in the metastable fluid region can be described well by a Krieger-Dougherty type of equation with a limiting (singular) density in both cases which is statistically the same as the HS glass transition value, ρ_g. Combination of these two dependencies leads to the fractional Stokes-Einstein equation, which has gained increasing popularity since the 1980s,[68] and here has the specific form, $D\eta^p = 0.155$, where $p = 0.97$. This suggests that there is 'decoupling' of D and η even in the metastable fluid state for densities below ρ_g, and this generic relationship[71] and general trend[72,73] has been observed in experimental systems. In practice in simulations this limiting density is difficult to achieve without the system nucleating. Solid nucleation manifests a particularly extreme case of heterogeneous dynamics, HD, (highly cooperative and intermittent motion of the molecules) and different qualitative trends in the density dependence of the self-diffusion coefficient and shear viscosity. HD has been widely studied for highly dense metastable fluids (see e.g., Refs.[74,75]) although experimentally temperature is the usual independent variable rather than the density. In fact, one of the main advances made in the present work is that we have characterized and parametrized the metastable fluid region of hard spheres much better than in previous work, and to an extent where the fundamental underpinning characteristics of the metastable fluid state could be more rigorously investigated.

The hard sphere model was originally developed as a convenient and analytically tractable model and reference system for the static and dynamical properties of small molecule liquids. The equations and other new results presented in this report can be used for any molecular model which uses the HS system as reference system. Isomorphic theory has become a powerful complementary approach for small molecule liquids.[76] However for colloidal and granular systems the HS model still retains its central role, as often the details of the true interaction potential are not so important (at least for those systems where the attractive

interactions can be neglected).

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

Some of the MD calculations were performed at the Poznań Supercomputing and Networking Center (PCSS). DMH would like to thank Dr. T. Crane (Department of Physics, Royal Holloway, University of London, UK) for helpful software support.

References

1 J. P. Hansen and I. R. McDonald, *Theory of Simple Liquids: With Applications to Soft Matter*, Academic Press, New York, 4th edn, 2013.
2 L. V. Woodcock, *Nature*, 1997, **385**, 141.
3 E. G. Noya and N. G. Almarza, *Molecular Physics*, 2015, **113**, 1061.
4 N. Ohtori, H. Uchiyama and Y. Ishii, *The Journal of Chemical Physics*, 2018, **149**, 214501.
5 M. N. Bannerman, R. Sargant and L. Lue, *Journal of Computational Chemistry*, 2011, **32**, 3329.
6 J. Erpenbeck and W. Wood, *Molecular Dynamics Techniques for Hard Core Systems, in Modern Theoretical Chemistry, Vol. 6, Statistical Mechanics*, Plenum Press, New York, 1977.
7 B. J. Alder, D. M. Gass and T. E. Wainwright, *The Journal of Chemical Physics*, 1970, **53**, 3813.
8 J. M. Haile, *Molecular Dynamics Simulation: Elementary Methods*, Wiley, New York, 1992.
9 A. M. Puertas, C. De Michele, F. Sciortino, P. Tartaglia and E. Zaccarelli, *The Journal of Chemical Physics*, 2007, **127**, 144906.
10 M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids*, Oxford University Press, New York, 1989.
11 J. Tian, H. Jiang, Y. Gui and A. Mulero, *Phys. Chem. Chem. Phys.*, 2009, **11**, 11213.
12 M. Miandehy and H. Modarress, *Journal of Chemical Physics*, 2003, **119**, 2716.
13 J. Kolafa, S. Labík and A. Malijevský, *Phys. Chem. Chem. Phys.*, 2004, **6**, 2335.
14 M. N. Bannerman, L. Lue and L. V. Woodcock, *Journal of Chemical Physics*, 2010, **132**, 084507.
15 N. F. Carnahan and K. E. Starling, *The Journal of Chemical Physics*, 1969, **51**, 635.
16 T. Boublik and I. Nezbeda, *Collection of Czechoslovak Chemical Communications*, 1986, **51**, 2301.
17 D. A. de Lonngi and P. A. L. Villanueva, *Molecular Physics*, 1991, **73**, 763.
18 S. Labík, J. Kolafa and A. Malijevský, *Phys. Rev. E*, 2005, **71**, 021105.
19 A. J. Schultz and D. A. Kofke, *Phys. Rev. E*, 2014, **90**, 023301.
20 M. Isobe and W. Krauth, *The Journal of Chemical Physics*, 2015, **143**, 084509.
21 J. Kolafa, *Phys. Chem. Chem. Phys.*, 2006, **8**, 464.
22 R. D. Kamien and A. J. Liu, *Phys. Rev. Lett.*, 2007, **99**, 155501.
23 K. R. Hall, *The Journal of Chemical Physics*, 1972, **57**, 2252.
24 P. Tarazona, *Phys. Rev. Lett.*, 2000, **84**, 694.
25 R. J. Speedy, *Journal of Physics: Condensed Matter*, 1998, **10**, 4387.
26 S. K. Kwak, T. Park, Y.-J. Yoon and J.-M. Lee, *Molecular Simulation*, 2012, **38**, 16.
27 B. J. Alder and T. E. Wainwright, *The Journal of Chemical Physics*, 1957, **27**, 1208.
28 C. Vega and E. G. Noya, *The Journal of Chemical Physics*, 2007, **127**, 154113.
29 J. M. Polson, E. Trizac, S. Pronk and D. Frenkel, *The Journal of Chemical Physics*, 2000, **112**, 5339.
30 D. Frenkel and B. Smit, *Understanding Molecular Simulation*, Academic Press, New York, 2001.
31 D. Frenkel and A. J. C. Ladd, *The Journal of Chemical Physics*, 1984, **81**, 3188.
32 J. Chang and S. I. Sandler, *The Journal of Chemical Physics*, 2003, **118**, 8390.
33 N. G. Almarza, *The Journal of Chemical Physics*, 2007, **126**, 211103.
34 L. A. Fernández, V. Martín-Mayor, B. Seoane and P. Verrocchio, *Phys. Rev. Lett.*, 2012, **108**, 165701.
35 M. B. Sweatman, *Phys. Rev. E*, 2005, **72**, 016711.
36 E. A. Ustinov, *The Journal of Chemical Physics*, 2017, **146**, 034110.
37 S. Pieprzyk, D. M. Heyes and A. C. Brańka, *Phys. Rev. E*, 2014, **90**, 012106.
38 A. C. Brańka, S. Pieprzyk and D. M. Heyes, *Phys. Rev. E*, 2018, **97**, 022119.
39 G. Ruppeiner, *Rev. Mod. Phys.*, 1995, **67**, 605.
40 G. Ruppeiner, *Rev. Mod. Phys.*, 1996, **68**, 313.
41 G. Ruppeiner, *Journal of Low Temperature Physics*, 2016, **185**, 246.
42 G. Ruppeiner, *Phys. Rev. E*, 2012, **86**, 021130.
43 G. Ruppeiner, P. Mausbach and H.-O. May, *Physics Letters A*, 2015, **379**, 646.
44 G. Ruppeiner, N. Dyjack, A. McAloon and J. Stoops, *The Journal of Chemical Physics*, 2017, **146**, 224501.
45 S. Chapman, T. G. Cowling and D. Burnett, *The Mathematical Theory of Non-Uniform Gases*, Cambridge University Press, 1990.
46 M. Fushiki, *Phys. Rev. E*, 2003, **68**, 021203.
47 D. M. Heyes, M. J. Cass, J. G. Powles and W. A. B. Evans, *The Journal of Physical Chemistry B*, 2007, **111**, 1455.
48 D. M. Heyes, *Journal of Physics: Condensed Matter*, 2007, **19**, 376106.
49 I.-C. Yeh and G. Hummer, *The Journal of Physical Chemistry B*, 2004, **108**, 15873.
50 O. A. Moultos, Y. Zhang, I. N. Tsimpanogiannis, I. G. Economou and E. J. Maginn, *The Journal of Chemical Physics*, 2016, **145**, 074109.
51 R. O. Sokolovskii, M. Thachuk and G. N. Patey, *The Journal of Chemical Physics*, 2006, **125**, 204502.

52 S. Gabl, C. Schröder and O. Steinhauser, *The Journal of Chemical Physics*, 2012, **137**, 094501.
53 B. Charbonneau, P. Charbonneau, Y. Jin, G. Parisi and F. Zamponi, *The Journal of Chemical Physics*, 2013, **139**, 164502.
54 G. Kikugawa, T. Nakano and T. Ohara, *The Journal of Chemical Physics*, 2015, **143**, 024507.
55 K.-S. Kim, M. H. Han, C. Kim, Z. Li, G. E. Karniadakis and E. K. Lee, *The Journal of Chemical Physics*, 2018, **149**, 044510.
56 B. Dünweg and K. Kremer, *The Journal of Chemical Physics*, 1993, **99**, 6983.
57 J. J. Erpenbeck and W. W. Wood, *Phys. Rev. A*, 1991, **43**, 4254.
58 H. Sigurgeirsson and D. M. Heyes, *Molecular Physics*, 2003, **101**, 469.
59 A. Batschinski, *Z. Phys. Chem., Stoechiom. Verwandtschaftsl.*, 1913, **84**, 643.
60 J. H. Hildebrand, *Science*, 1971, **174**, 490.
61 J. Hildebrand, *Viscosity and Diffusivity: A Predictive Approach*, Wiley, New York, 1986.
62 D. M. Heyes and A. C. Brańka, *The Journal of Chemical Physics*, 2005, **122**, 234504.
63 R. J. Speedy, *Molecular Physics*, 1998, **95**, 169.
64 G. Parisi and F. Zamponi, *The Journal of Chemical Physics*, 2005, **123**, 144501.
65 G. L. Hunter and E. R. Weeks, *Reports on Progress in Physics*, 2012, **75**, 066501.
66 K. R. Harris, *The Journal of Chemical Physics*, 2009, **131**, 054503.
67 N. Ohtori and Y. Ishii, *Phys. Rev. E*, 2015, **91**, 012111.
68 R. Zwanzig and A. K. Harrison, *The Journal of Chemical Physics*, 1985, **83**, 5861.
69 L. Costigliola, D. M. Heyes, T. B. SchrÃÿder and J. C. Dyre, *The Journal of Chemical Physics*, 2019, **150**, 021101.
70 R. Zwanzig, *The Journal of Chemical Physics*, 1983, **79**, 4507.
71 M. D. Ediger and P. Harrowell, *The Journal of Chemical Physics*, 2012, **137**, 080901.
72 P. G. Debenedetti and F. H. Stillinger, *Nature*, 2001, **410**, 259.
73 K. Vollmayr-Lee, W. Kob, K. Binder and A. Zippelius, *The Journal of Chemical Physics*, 2002, **116**, 5158–5166.
74 C. Donati, J. F. Douglas, W. Kob, S. J. Plimpton, P. H. Poole and S. C. Glotzer, *Phys. Rev. Lett.*, 1998, **80**, 2338.
75 B. O'Malley and I. Snook, *The Journal of Chemical Physics*, 2005, **123**, 054511.
76 J. C. Dyre, *The Journal of Chemical Physics*, 2018, **149**, 210901.