

# On standardised moments of force distribution in simple liquids

Jonathan Utterson<sup>1, a)</sup> and Radek Erban<sup>1, b)</sup>

*Mathematical Institute, University of Oxford, Radcliffe Observatory Quarter, Woodstock Road, Oxford, OX2 6GG, United Kingdom*

(Dated: 20 December 2021)

The force distribution of a tagged atom in a Lennard-Jones fluid in the canonical ensemble is studied with a focus on its dependence on inherent physical parameters: number density ( $n$ ) and temperature ( $T$ ). Utilising structural information from molecular dynamics simulations of the Lennard-Jones fluid, explicit analytical expressions for the dependence of standardised force moments on  $n$  and  $T$  are derived. Leading order behaviour of standardised moments of the force distribution are obtained in the limiting cases of small density ( $n \rightarrow 0$ ) and low temperature ( $T \rightarrow 0$ ), while the variations in the standardised moments are probed for general  $n$  and  $T$  using molecular dynamics simulations. Clustering effects are seen in molecular dynamics simulations and their effect on these standardised moments is discussed.

## I. INTRODUCTION

Understanding the moments and measures of a distribution for a fully atomistic molecular dynamics (MD) simulation allow us to better fit coarser models that reproduce these<sup>1–3</sup>. It is often the case in model coarse graining that we wish to directly reconcile the energy landscape of the fully atomistic system to a more basic representation that allows us to maintain as many physical properties of the system of interest, with as little computational cost as possible<sup>4</sup>. Though, it is also natural to match forces between the high and low resolution systems in an effort to reproduce the force distribution which will inherently give rise to the energy landscape<sup>5–10</sup>.

Let  $\mathbf{F} = [F_1, F_2, F_3]$  denote a force on a tagged atom in a liquid. Depending on the relative positions of other atoms, force  $\mathbf{F}$  can vary over a range of values and a detailed information on  $\mathbf{F}$  can be obtained by calculating properties of its equilibrium distribution, which we will call force distribution in this manuscript. Considering an isotropic system, the equilibrium distribution of each force coordinate is the same. We define the standardised moment of the force distribution by averaging over the  $k$ -th power of its first coordinate as

$$\alpha_k = \frac{\langle F_1^k \rangle}{\langle F_1^2 \rangle^{k/2}}, \quad (1)$$

where  $\langle F_1^k \rangle$  is the  $k$ -th moment of the force distribution and  $\alpha_k$  standardises the  $k$ -th moment by scaling it with the  $k$ -th power of the standard deviation of the force distribution. In a simple homogeneous fluid with radially symmetric interactions between particles, the force distribution will exhibit symmetry around the origin and thus all odd standardised moments vanish, i.e.  $0 = \alpha_1 = \alpha_3 = \alpha_5 = \dots$ . As  $\alpha_2 \equiv 1$  by definition (1), the first non-trivial standardised moment is kurtosis, denoted  $\alpha_4$ , which provides a measure of spread that details how tailed the force distribution is relative to a normal distribution<sup>11</sup>. In this paper, we study how the force distribution depends on the number density of a homogeneous many-body system, and the temperature of the same system in a canonical ensemble. We

will do this by studying the behaviour of the second moment of the force distribution  $\langle F_1^2 \rangle$  and standardised even moments  $\alpha_4, \alpha_6, \alpha_8, \dots$ . If the force distribution was Gaussian, then the even standardised moments would be

$$\alpha_k = (k-1)!! = \prod_{i=1}^{k/2} (2i-1), \quad \text{for } k = 2, 4, 6, 8, 10, \dots, \quad (2)$$

and the second moment  $\langle F_1^2 \rangle$  would be sufficient to parametrize the force distribution. However, the force distributions in simple liquids have been reported to deviate from Gaussian distribution<sup>12–14</sup>. In particular, by comparing the results of our analysis with Gaussian moments in equation (2), we can also quantify how non-Gaussian the real force distribution is.

Much work has been done in the area of force distributions of many-body systems: with seminal work from Chandrasekhar<sup>15</sup> that employed Markov's theory of random flights to give an expression for the force distribution of a many-body system interacting through a  $1/r$  gravitational potential. More recent work has been done with the help of MD by Gabrielli et al<sup>16</sup>, who derived an expression for the kurtosis of the force distribution for a lattice system of atoms interacting through the gravitational potential. Further, using the classical density functional theory, an expression for the probability distribution of force for a system interacting through an arbitrary weakly repulsive potential was derived by Rickayzen et al<sup>17,18</sup>.

In this paper, we study the number density and temperature dependence of the force distribution for a many-body system interacting through a Lennard-Jones 12-6 potential<sup>19,20</sup>, which is ubiquitously used and has been shown to model homogeneous systems of interacting (Argon) atoms well<sup>21–23</sup>.

In Section III, an in depth investigation is given to the simple two-body system in one spatial dimension, which provides the ideal platform to illustrate the underlying methods while retaining interesting dynamical behaviour. From first principles we derive first-order partial differential equations (PDEs) describing the dependence of the standardised moments of the force distribution has on parameters. In doing so we further derive an analytic expression for the partition function of a two-body system that depends solely on the standardised moments of the force distribution whereupon the expression is exact in an asymptotic limit of the density going to zero

<sup>a)</sup>Electronic mail: utterson@maths.ox.ac.uk

<sup>b)</sup>Electronic mail: erban@maths.ox.ac.uk

( $n \rightarrow 0$ ). Similarly, an expression is derived relating the average energy of the system to standardised moments of force from the temperature dependent PDE. In parameter regimes where long-range forces between atoms dominate, we use a truncated Taylor series expansion to derive the leading order behaviour of the kurtosis of the force distribution in the limit  $n \rightarrow 0$ . Finally, we utilise a Laplace integral approximation to ascertain the leading order behaviour of the standardised moments of force at low temperatures ( $T \rightarrow 0$ ). Results from simple MD simulation are presented to provide evidence for the efficacy of these methods and underlying assumptions.

This is followed by Section IV, where the natural idea that long range force calculations dictate asymptotic behaviour is extended from the 1D model to many-body systems of arbitrary size in three spatial dimensions. These systems exhibit the physical properties of standard MD simulations: *i.e.* cubic geometry with periodic boundary conditions that employ the minimum image convention. In particular, we can analyze the system by performing calculations on a central cubic cell. In Section IV B, MD results are displayed for many-body systems. We present the dependence of the standardised force moments on density,  $n$ , and temperature,  $T$ , and discuss the parameters and integrator schemes utilised in producing the results of MD simulations.

## II. NOTATION

We consider a system of  $N$  identical atoms interacting via the Lennard-Jones 12-6 potential<sup>19</sup>. This is a ubiquitous interatomic pairwise potential; here the potential between atoms labelled  $i, j = 1, 2, \dots, N$  positioned at  $\mathbf{q}_i, \mathbf{q}_j \in \mathbb{R}^3$  is given (in reduced units<sup>24</sup>) by the expression

$$U_{ij}(r_{ij}) = 4 \left( \frac{1}{r_{ij}^{12}} - \frac{1}{r_{ij}^6} \right), \quad (3)$$

where  $r_{ij} = |\mathbf{q}_i - \mathbf{q}_j|$  is the distance between atoms. The Lennard-Jones potential (3) between two atoms has a unique minima obtained at  $r_{ij} = r_* = 2^{1/6}$ .

We employ the framework of statistical mechanics for this closed many-body system and describe atom  $i = 1, 2, \dots, N$  by phase space coordinates  $\{\mathbf{q}_i, \mathbf{p}_i\} \in \mathbb{R}^6$ , where  $\mathbf{p}_i$  denotes the momentum of the  $i$ -th atom. We work in the canonical ensemble with temperature  $T$ ; the partition function therefore becomes

$$\mathcal{Z}_N(T, V) = \frac{1}{h^{3N} N!} \iint_{\Omega_{\mathbf{q}} \times \Omega_{\mathbf{p}}} \exp[-\beta H(\mathbf{q}, \mathbf{p})] d^3 \mathbf{q} d^3 \mathbf{p}, \quad (4)$$

where  $V$  is the volume of our closed system, and  $\mathbf{q} = (\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_N)^T$  and  $\mathbf{p} = (\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_N)^T$  are vectors containing the positions and momenta of all atoms in the system. Our integration domain is given by  $\Omega_{\mathbf{q}} \times \Omega_{\mathbf{p}} \subset \mathbb{R}^{3N} \times \mathbb{R}^{3N}$ . This denotes the phase space of our system. For systems of interest  $\Omega_{\mathbf{p}} \equiv \mathbb{R}^{3N}$ . The underlying geometry of the system (and principle simulation cell) is a cubic box of size  $L > 0$ ,

therefore  $\Omega_{\mathbf{q}} \equiv (-L/2, L/2] \times \dots \times (-L/2, L/2]$ . The phase space volume elements in equation (4) are denoted by

$$d^3 \mathbf{q} = \prod_{i=1}^N d^3 \mathbf{q}_i \quad \text{and} \quad d^3 \mathbf{p} = \prod_{i=1}^N d^3 \mathbf{p}_i. \quad (5)$$

Throughout this work we make use of reduced units<sup>24</sup>, utilising Argon parameters<sup>25</sup>. In particular, all instances of  $T$  in this work can be translated back to SI units with the transformation  $T \rightarrow k_B T$  where  $k_B$  is the Boltzmann factor. Therefore, in the partition function (4), we have  $\beta = 1/T$  and  $h$  is the Planck constant ( $\approx 0.186$  in reduced units). Finally,  $H(\mathbf{q}, \mathbf{p})$  is the classical Hamiltonian  $H(\mathbf{q}, \mathbf{p}) = K(\mathbf{p}) + U(\mathbf{q})$  with kinetic energy  $K(\mathbf{p}) = |\mathbf{p}|^2/2$  (where the usual factor of mass is unity under reduced units) and a general potential  $U(\mathbf{q})$ . The statistical average of a quantity  $X$  for this  $N$ -body system is given by

$$\langle X \rangle = \frac{1}{\mathcal{Z}_N h^{3N} N!} \iint_{\Omega_{\mathbf{q}} \times \Omega_{\mathbf{p}}} X \exp[-\beta H(\mathbf{q}, \mathbf{p})] d^3 \mathbf{q} d^3 \mathbf{p}, \quad (6)$$

where the Boltzmann factor acts as a statistical weighting for a configuration  $\{\mathbf{q}, \mathbf{p}\} \in \mathbb{R}^{6N}$ , normalised such that  $\langle 1 \rangle = 1$ .

We label atoms so that the first one is the tagged atom. Denoting the force on the tagged atom produced from the  $j$ -th atom by  $\mathbf{F}_j = [F_{j,1}, F_{j,2}, F_{j,3}] \in \mathbb{R}^3$ , for  $j = 2, 3, \dots, N$ , the total force  $\mathbf{F} = [F_1, F_2, F_3]$  on the tagged atom is

$$\mathbf{F} = \sum_{j=2}^N \mathbf{F}_j.$$

We define

$$f_k = \int_{\Omega_{\mathbf{q}}} \left( \sum_{j=2}^N F_{j,1} \right)^k \exp[-\beta U(\mathbf{q})] d^3 \mathbf{q} \quad (7)$$

for  $k = 0, 1, 2, \dots$ . Then we have

$$\frac{f_k}{f_0} = \left\langle \left( \sum_{j=2}^N F_{j,1} \right)^k \right\rangle = \langle F_1^k \rangle.$$

Then the  $k$ -th standardised moment (1) is given by

$$\alpha_k = \frac{f_0^{k/2-1} f_k}{f_2^{k/2}}, \quad (8)$$

where we are interested in cases  $k = 4, 6, 8, \dots$ .

In order to study how the force distribution depends on the physical parameters of interest it is useful to identify how changes in these parameters will manifest themselves in the system. Indeed, we choose to work in the canonical ensemble with a target temperature of  $T$ : this is accomplished with the use of a thermostat which is discussed further in Section IV B and Appendix B. It is more illuminating to see that if we have a system with a fixed number of free interacting atoms  $N$  in a cubic box of side  $L$ ; the (reduced) number density is given by  $n = N/L^3$ . Therefore the approach we employ in this paper

to ascertain how values of standardised moments depend on number density, will be to keep the number of atoms fixed but vary the box width  $L$  - this will manifest as a change in density  $n$ . Similarly one could keep the volume of the cubic box the same and vary the number of atoms though this is a point of discussion in Section IV B.

For the remainder of the paper we will study systems with different spatial dimensions. The size of the system varies by changing the number of particles  $N$ ; we will use equation (8) as a crucial initial point in each calculation. We will naturally proceed by investigating systems of increasing complexity; starting from a cartoon one-dimensional model and culminating to a general many-body system of arbitrary size in three spatial dimensions.

### III. ONE ATOM IN A POTENTIAL WELL

We now go on to illustrate three approaches to obtain the dependence of the force distribution on parameters  $n$  and  $T$ . It is useful to note that, as we are now working in one spatial dimension, density  $n$  is proportional to  $1/L$ , *i.e.* we have  $n \propto 1/L$ . We will consider a simple system in one spatial dimension consisting of two atoms interacting through the Lennard-Jones potential (3) in interval  $[0, L]$  with periodic boundary conditions. One of the atoms is considered to be fixed at position  $q_0 = L/2 \in [0, L]$  and the other atom is free to move, therefore, we have  $N = 1$  free atom. Its position is denoted  $x \in [0, L]$ . Therefore, the inter-atomic distance is  $r = |x - q_0|$ . Using our simplified one-dimensional set up,  $F_1 = F$  and  $\Omega_{\mathbf{q}} = (0, L)$ , equation (7) reduces to

$$f_k(L) = \int_0^L F^k(|x - q_0|) \exp[-\beta U(|x - q_0|)] dx, \quad (9)$$

which is the marginalised expected value of the  $k$ -th moment of force  $F(x) = -dU/dx$ , where we have dropped subscripts in the Lennard-Jones potential (3) and we write it as  $U(z) = 4(z^{-12} - z^{-6})$ . Utilising the symmetry of the potential (and therefore the force) we are left with

$$f_k(L) = 2 \int_0^{L/2} F^k(r) \exp[-\beta U(r)] dr. \quad (10)$$

In what follows, we will assume that we are in a regime where the box width  $L$  satisfies  $L \gg r_*$ , where  $r_* = 2^{1/6}$  minimizes the Lennard-Jones potential  $U$ .

#### A. Differential equation for standardised moments

We consider a perturbation of the form  $L \rightarrow L + \delta L$ . Using equation (10) and considering terms to the order  $O(\delta L)$ , we obtain

$$\begin{aligned} f_k(L + \delta L) &= f_k(L) + f'_k(L) \delta L + O(\delta L^2) \\ &= f_k(L) + F^k(L/2) \exp[-\beta U(L/2)] \delta L + O(\delta L^2). \end{aligned}$$

Using equation (8), we approximate  $\alpha_k(L + \delta L)$  by

$$\alpha_k(L) + \alpha_k(L) v_k(L) \exp[-\beta U(L/2)] \delta L + O(\delta L^2),$$

where our notation  $\alpha_k(L)$  highlights the dependence of the standardised moments of force,  $\alpha_k$ , on  $L$ , and function  $v_k(L)$  is given by

$$v_k(L) = \frac{k-2}{2f_0(L)} + \frac{F^k(L/2)}{f_k(L)} - \frac{kF^2(L/2)}{2f_2(L)}. \quad (11)$$

Taking the limit  $\delta L \rightarrow 0$ , we obtain the derivative of the  $k$ -th standardised moment of force, with respect to  $L$ , as

$$\frac{\partial \alpha_k}{\partial L}(L) = v_k(L) \exp[-\beta U(L/2)] \alpha_k(L), \quad (12)$$

where  $v_k(L)$  are expressed in terms of integrals (10) as given by equation (11).

#### B. Far-field integral approximation

To further analyze integrals (10), we introduce a cutoff  $c$ , which satisfies that  $r_* < c < L/2$ , where  $r_* = 2^{1/6}$  is a unique maximum of  $\exp[-\beta U(z)]$ , which can be Taylor expanded as  $\beta(1 + 4z^{-6} + 4z^{-12} - 16/3z^{-18} + 8z^{-24} \dots)$ . Considering sufficiently large  $L$ , we can choose the cutoff  $c$ , so that

$$\left| f_0(L) - 2 \left( \int_0^c \exp[-\beta U(r)] dr + \beta \int_c^{L/2} 1 + \frac{4}{r^6} dr \right) \right| \leq \varepsilon, \quad (13)$$

where tolerance  $\varepsilon$  is chosen to be  $10^{-4}$  in our illustrative computations. This splitting allows us to numerically calculate the bulk of the integral (10) as a constant independent of  $L$  and then use the second term to give an analytic expression for  $\alpha_k$  with dependence on  $L$ , and ultimately on  $n$ .

The range of values of  $T$  that are of typical use are chosen in order to maintain the liquid state of Argon during simulation. These are approximately temperatures in the interval  $0.70 < T < 0.73$  under ambient conditions<sup>26</sup>. Therefore, as volume is varied we are in a regime where  $\beta = O(1)$ , for convenience we set  $\beta = 1$ . Though given that the density of our system changes between each simulation some systems will be in a liquid phase and others in a gaseous phase, this is a point of discussion in Section IV B.

Splitting the integration domain  $[0, L/2]$  of integral (10) into  $[0, c]$  and  $[c, L/2]$ , we use the exact form of the integrand in  $[0, c]$  to obtain a ‘near-field’ contribution. Utilising an approximate form for the integrand given by the truncated Taylor expansion  $f(z)$  in the domain  $[c, L/2]$  gives rise to a density dependent ‘far-field’ contribution. Combining these we arrive at the approximate form for  $f_0(L)$ . Using cutoff  $c = 2$ , equation (13) is satisfied with  $\varepsilon = 10^{-4}$ . Therefore, upon numerically calculating the bulk contribution for the integral with domain  $[0, 2]$ , we get

$$f_0(L) = 2 \int_0^{L/2} \exp[-\beta U(r)] dr = b_0 + L + O(L^{-6}) \quad (14)$$

with  $b_0 = -0.71832$ , which depends on our choice of cutoff  $c = 2$ . Similarly, we can calculate far-field integral approximations of integrals (10) for general values of  $k = 2, 4, 6, 8, 10, 12$ . The integrand  $F^k(r) \exp[-\beta U(r)]$  has maxima when  $r = r_* = 2^{1/6}$  or when  $kU''(r) = \beta(U'(r))^2$ . This forms a cubic in  $r^6$  that can be solved. For the values of  $k$  used in this work, this sometimes results in a global maximum, that always lies at a distance less than  $r < r_*$  from the origin. Therefore  $r_* = 2^{1/6}$  is the furthest maximum of the integrand from the origin.

Splitting integral (10) into a near-field and far-field contribution, using the general cutoff  $c = 2$ , we find

$$f_k(L) = b_k + O(L^{-7k}) \quad (15)$$

The near-field contributions,  $b_k$ , generally increase vastly if we increase the value of  $k$ , for example

$$b_0 = -0.71832, \quad b_2 = 130.64 \quad \text{and} \quad b_4 = 2.5727 \times 10^5, \quad (16)$$

while the dependence on  $L$  decreases more rapidly for larger values of  $k$ . Therefore, the non-negligible density contributions to  $\alpha_k(L)$  in the low density limit come exclusively from the normalisation  $f_0(L)$  given by (14).

Substituting equations (14) and (15) in equation (8), we obtain an expression for the general  $k$ -th standardised moment of force

$$\alpha_k(L) = \frac{b_0^{k/2-1} b_k}{b_2^{k/2}} \left( 1 + \frac{L}{b_0} + O(L^{-6}) \right)^{k/2-1}. \quad (17)$$

Using the values of  $b_0$ ,  $b_2$  and  $b_4$  given by (16), we obtain the dependence of the kurtosis of the force distribution on the reduced number density  $n = 1/L$  in the dilute limit  $n \rightarrow 0$  as  $\alpha_4 = -10.828 + 15.074n^{-1} + O(n^6)$ . Figure 1 compares this result with the results obtained by MD simulation of the one atom system. We observe that MD is in good agreement with the results obtained by formula (17).

### C. Leading order behaviour for differential equation (12)

Since  $L/2 > r_*$ , the force  $F(L/2)$  monotonically decreases as a function of  $L$ . When looking at leading order approximations in the low density limit  $n \rightarrow 0$  (equivalent to limit  $L \rightarrow \infty$ ) to equation (12), we need to analyse  $v_k(L)$ . The second and third term in equation (11) converge to zero more rapidly than the first term as  $L \rightarrow \infty$ , therefore the leading order behaviour is given by the first term

$$v_k(L) \sim \frac{k-2}{2f_0(L)} \quad \text{as} \quad L \rightarrow \infty. \quad (18)$$

By utilising the far field integral approximation (14), we arrive at  $f_0(L) \sim (b_0 + L)$ , where  $b_0 = b_0(c)$  is a constant term that depends on cutoff parameter  $c$ . With this, our leading order approximation of the  $k$ -th standardised moment,  $\alpha_k^0$ , obeys

$$\frac{\partial \alpha_k^0}{\partial L}(L) = \frac{k-2}{2(b_0 + L)} \alpha_k^0(L).$$

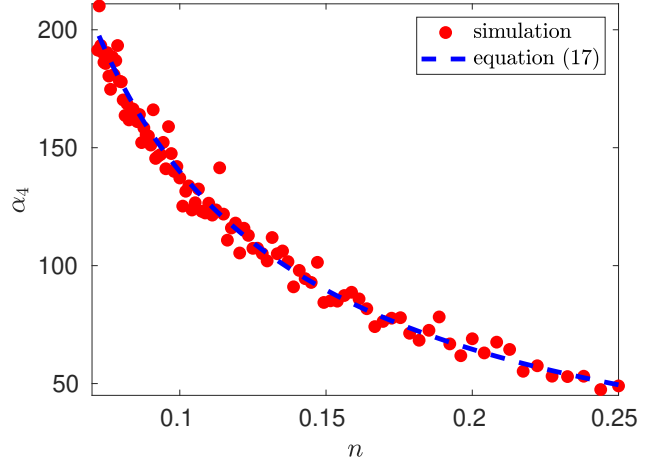


FIG. 1. Plot of  $\alpha_4$  as a function of  $n = 1/L$  for the illustrative one-atom system. Results of MD simulations are compared with  $\alpha_4 = -10.828 + 15.074n^{-1}$  obtained by using equation (17) with  $b_0$ ,  $b_2$  and  $b_4$  given by (16) (blue dashed line). MD simulation results for temperature  $T = 1$  utilising Langevin dynamics<sup>27</sup> described in equation (B1), with friction parameter  $\gamma = 0.1$ , are represented by red dots. The MD simulation length was a total of  $1.1 \times 10^8$  time steps with the first  $10^7$  time steps used for initialisation.

Finally this gives us that

$$\alpha_k^0(L) = C_k (b_0 + L)^{k/2-1} = C_k (b_0 + n^{-1})^{k/2-1}, \quad (19)$$

where  $n = 1/L$  is the reduced number density and  $C_k$  is a constant. Equation (19) gives the same leading order behaviour  $n^{1-k/2}$  in the limit  $n \rightarrow 0$  as equation (17): the same behaviour is also seen for the Lennard-Jones fluid in Section IV. Though the method above is more generally applicable to include potentials that monotonically decay as  $r^{-a}$  as  $r \rightarrow \infty$  for  $a > 0$ . We next make the observation that equation (4) in 1D can be written as:

$$\mathcal{Z}_1(T, V) = \frac{1}{h} \int_0^L \exp[-\beta U(q)] dq \int_{-\infty}^{\infty} \exp\left[-\frac{\beta p^2}{2}\right] dp, \quad (20)$$

where the Planck factor of  $1/h$  arises instead of  $1/h^3$  due to the fact that we are in one-dimensional physical space. Using (10), we obtain

$$f_0(L) = h \mathcal{Z}_1(T, V) \sqrt{\frac{\beta}{2\pi}}. \quad (21)$$

Considering the low density limit  $n \rightarrow 0$  (i.e.  $L \rightarrow \infty$ ) in equation (12) and using (18) and (21), we obtain

$$\mathcal{Z}_1(T, V) \sim \frac{(k-2)\sqrt{2\pi}}{\sqrt{h^2\beta}} \left[ \alpha_k(L) \left( \frac{\partial \alpha_k}{\partial L}(L) \right)^{-1} \right], \quad (22)$$

as  $L \rightarrow \infty$ . In particular, we can obtain the partition function (20) in the dilute (low density) limit by using information



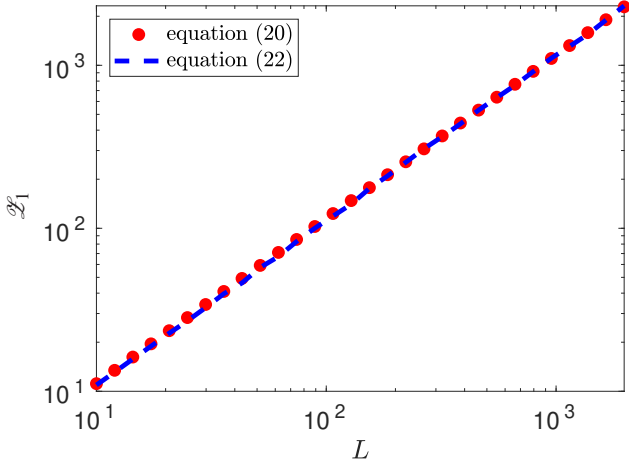


FIG. 2. Approximation of the partition function  $\mathcal{Z}_1(T, V)$  obtained using the right hand side of equation (22) with  $k = 4$  and values of kurtosis ( $\alpha_4$ ) estimated from MD simulation (blue dashed line). The exact values obtained by (20) are plotted as the red dots.

about the moments of the force distribution. The accuracy of equation (22) is illustrated in Figure 2, where we use  $k = 4$ . We use MD simulations of a single atom, using a range of simulation box widths  $L$ . We estimate the values of kurtosis of the force distribution, its derivative with respect of  $L$  and use the right hand side of equation (22) to estimate the  $\mathcal{Z}_1(T, V)$ . Considering  $L \geq 10$ , the result is within 5% error when compared with the exact result (20), while for larger values of box width  $L$  the error decreases to around 1%, confirming that the formula (22) is valid in the asymptotic limit  $L \rightarrow \infty$ .

#### D. Temperature dependence of standardised moments

One can perform a similar analysis as in Section III A, viewing the moments  $\alpha_k = \alpha_k(T)$  as a function of temperature  $T = 1/\beta$ . To do that, we consider the moment definition (10) as a function of temperature  $T$ , namely, we define

$$f_k(T) = 2 \int_0^{L/2} F^k(r) \exp\left[-\frac{U(r)}{T}\right] dr. \quad (23)$$

Considering small perturbations of these functions with respect to  $T \rightarrow T + \delta T$ , while fixing the domain length  $L$ , and collecting terms up to first order in  $\delta T$ , we obtain

$$\frac{\partial \alpha_k}{\partial T}(T) = v_k(T) \alpha_k(T), \quad (24)$$

where

$$v_k(T) = \left(\frac{k}{2} - 1\right) \frac{f'_0(T)}{f_0(T)} + \frac{f'_k(T)}{f_k(T)} - \left(\frac{k}{2}\right) \frac{f'_2(T)}{f_2(T)}. \quad (25)$$

Combining equations (24) and (25) with equation (21) where  $\beta = 1/T$ , we obtain

$$\frac{\partial}{\partial T} \ln\left(\frac{\alpha_k^2 f_2^k}{f_k^2}\right) = (k-2) \left(\frac{\partial}{\partial T} \ln(\mathcal{Z}_1) - \frac{1}{2T}\right).$$

Since  $-\partial/\partial\beta(\ln \mathcal{Z}_1)$  is equal to the average energy of the system,  $\langle E \rangle$ , we have

$$\langle E \rangle = \frac{T}{2} + \frac{T^2}{k-2} \frac{\partial}{\partial T} \ln\left(\frac{\alpha_k^2 f_2^k}{f_k^2}\right), \quad (26)$$

where the first term on the right hand side of equation (26) is the average kinetic energy of our one-atom system. Substituting equation (8) into the second term on the right hand side, it can be rewritten as  $T^2 \partial(\ln f_0)/\partial T$ . Thus, using equation (6), we confirm that the second term on the right hand side of equation (26) is the average potential energy.

#### E. Low temperature limit

Next, we consider the behaviour of the  $k$ -th standardised moment of force,  $\alpha_k(T)$ , given by equation (8), in the low temperature limit,  $T \rightarrow 0$ , which is equivalent to the limit  $\beta \rightarrow \infty$ . Since the inter-atomic potential  $U(r)$  has a global minimum at  $r = r_*$  in interval  $[0, L/2]$ , integrals of the form (10) and (23) can be approximated by Laplace's method in the limit  $\beta \rightarrow \infty$  and  $T \rightarrow 0$ , respectively. A general discussion of Laplace's method is given in Chapter 6 of the book by Bender and Orszag<sup>28</sup>. We calculate the asymptotic expansion of  $f_0(T)$  by applying Laplace's method to integral (23) for  $k = 0$ . We approximate the integration limits of integral (23) to lie within the domain  $r \in (r_* - \varepsilon, r_* + \varepsilon)$ , where  $\varepsilon \ll 1$ , and we Taylor expand  $U(r)$  at  $r = r_*$ . Using  $U'(r_*) = 0$ , we have

$$U(r) \approx U(r_*) + (r - r_*)^2 U''(r_*)/2 + (r - r_*)^3 U^{(3)}(r_*)/6 + (r - r_*)^4 U^{(4)}(r_*)/24,$$

where we denote the  $m^{\text{th}}$  derivative of  $U$  as  $U^{(m)}$  for  $m \geq 3$ . Substituting into integral (23), we arrive at the asymptotic expansion

$$f_0(T) \sim \frac{\sqrt{\pi T} \exp[-U(r_*)/T]}{\sqrt{2U''(r_*)}} \left[1 + B_0 T + O(T^2)\right], \quad (27)$$

as  $T \rightarrow 0$ , where constant  $B_0$  is given by<sup>28</sup>

$$B_0 = \frac{5(U^{(3)}(r_*))^2}{24(U''(r_*))^3} - \frac{U^{(4)}(r_*)}{8(U''(r_*))^2}. \quad (28)$$

To apply Laplace's method to integral (23) for  $k = 2, 4, 6, \dots$ , we note that  $F^k(r) = (U'(r))^k$  for even values of  $k$ . Using the truncated Taylor expansion around  $r = r_*$  and noting that  $U'(r_*) = 0$ , we have

$$F^k(r) \approx (r - r_*)^k \left( (U''(r_*))^k + (r - r_*) C_{k,1} + (r - r_*)^2 C_{k,2} \right), \quad (29)$$

where  $C_{k,1}$  and  $C_{k,2}$  are constants, which can be expressed in terms of the derivatives of potential  $U(r)$  at  $r = r_*$  (see equa-

tions (A1) and (A2) in Appendix A). This gives the asymptotic expansion

$$f_k(T) \sim \frac{\sqrt{\pi T} \exp[-U(r_*)/T]}{\sqrt{2U''(r_*)}} \times A_k \left[ T^{k/2} + B_k T^{k/2+1} + O(T^{k/2+2}) \right]. \quad (30)$$

as  $T \rightarrow 0$ , where constants  $A_k$  and  $B_k$  are given by

$$A_k = (U''(r_*))^{k/2} (k-1)!!$$

and

$$B_k = \frac{(4k-15)(k^2-1)(U^{(3)}(r_*))^2}{72(U''(r_*))^3} + \frac{(k^2-1)U^{(4)}(r_*)}{8(U''(r_*))^2},$$

where the last formula reduces to equation (28) for  $k=0$ . Substituting (27) and (30) into (8) gives the following expression in the limit  $T \rightarrow 0$ :

$$\alpha_k \sim (k-1)!! \left( 1 + \frac{(k-2)B_0 + 2B_k - kB_2}{2} T + O(T^2) \right).$$

In particular, we have  $\alpha_2 \sim 1 + O(T^2)$  and

$$\begin{aligned} \alpha_4 &\sim 3 + 3 \left( \frac{(U^{(3)}(r_*))^2}{(U''(r_*))^3} + \frac{U^{(4)}(r_*)}{(U''(r_*))^2} \right) T + O(T^2) \\ &= 3 + \frac{203}{6} T + O(T^2). \end{aligned} \quad (31)$$

Therefore, Laplace's method predicts that the standardised moments of the force distribution,  $\alpha_k(T)$ , tend to the values given in equation (2) for Gaussian moments in the low temperature limit. This limiting behaviour is to be expected as during the Laplace approximation we use a Gaussian distribution to approximate the Boltzmann factor. We can interpret this approach as approximating the force distribution as Gaussian and perturbations of the system around small temperatures give rise to non-Gaussian contributions to the standardised moments.

Results from MD simulation are illustrated in Figure 3 over the range of values of temperature  $T$ . We see that the behaviour of kurtosis,  $\alpha_4$ , is well approximated by the linear approximation  $3 + 203T/6$  given in equation (31) for the temperature values satisfying  $T \leq 0.1$ , though this agreement diverges as temperature  $T$  increases and higher order terms,  $O(T^2)$  in equation (31), become significant. In Figure 3, we fix the box width as  $L = 10$ . Increasing the box width much further would take us to a regime where the particle is essentially free and the approximation calculated by the Laplace method around the potential minimum would lose validity.

#### IV. MANY-BODY SYSTEMS

In this section we employ the far field approximation approach introduced in Section III B and we will vary the number density of the system by changing the size  $L$  of the integration domain, which will be given as the three-dimensional

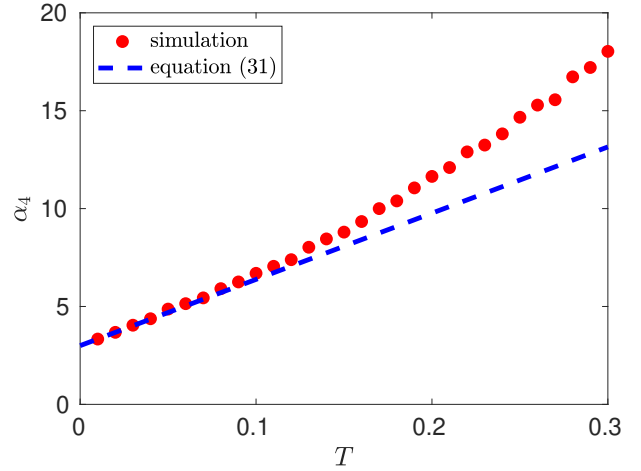


FIG. 3. Kurtosis,  $\alpha_4$ , as a function of temperature,  $T$ , for  $T \leq 0.3$ . The linear behaviour is estimated as  $\alpha_4(T) \sim 2.9388 + 37.002T$  for  $T \in (0.01, 0.10)$  (using the MD computed data, with density  $n = 0.1$ , visualized as red dots). We compare this to the theoretical linear result  $3 + 203T/6$  predicted by equation (31) (illustrated by the blue dashed line).

cube  $[0, L]^3$ . Using notation introduced in Section II, the distance between atoms labelled  $i, j = 1, 2, \dots, N$  positioned at  $\mathbf{q}_i, \mathbf{q}_j \in \mathbb{R}^3$  is denoted by  $r_{ij} = |\mathbf{q}_i - \mathbf{q}_j|$ . Taking into account the periodic boundary conditions, the distance  $|\mathbf{q}_i - \mathbf{q}_j|$  is the minimum image inter-atomic distance given by

$$|\mathbf{q}_i - \mathbf{q}_j| = \left( \overline{(q_i^x - q_j^x)}^2 + \overline{(q_i^y - q_j^y)}^2 + \overline{(q_i^z - q_j^z)}^2 \right)^{1/2}, \quad (32)$$

where the overline denotes  $\overline{\zeta} = \zeta - L[\zeta/L]$  for  $\zeta \in \mathbb{R}$  and  $[\cdot]$  rounds a real number to the nearest integer. For an interacting  $N$ -body system the dimensionality of the integral given by equation (7) is  $3N$ . We first present an illustrative calculation with  $N = 2$  interacting atoms in Section IV A and then we study systems with larger values of  $N$  in Section IV B.

##### A. Dependence of $\alpha_k$ on density for $N = 2$ interacting atoms

In Section III, we have considered two atoms in the one-dimensional spatial domain, where one atom was fixed at position  $q_0$ , i.e. we have effectively studied a single atom in a one-dimensional potential well. Here, we will consider  $N = 2$  interacting atoms in the three-dimensional cubic domain  $[0, L]^3$  with periodic boundary conditions. We calculate the  $k$ -th standardised moment of force according to equation (8). To do so, we consider equation (7), where we have  $d^3 \mathbf{q} = d^3 \mathbf{q}_1 d^3 \mathbf{q}_2$ ,  $U(\mathbf{q}) = U(r_{12})$ ,  $F_1(\mathbf{q}) = F_1(r_{12})$  and we integrate over the domain  $\Omega = [0, L]^3 \times [0, L]^3$  to get

$$f_k = \int_{\Omega} F_1^k(r_{12}) \exp[-\beta U(r_{12})] d^3 \mathbf{q}_1 d^3 \mathbf{q}_2. \quad (33)$$

It is useful to introduce a change of coordinates  $\xi^\ell = q_1^\ell - q_2^\ell$  and  $\eta^\ell = q_1^\ell + q_2^\ell$  for  $\ell = x, y, z$ . We note that  $r_{12}$  is only de-

pendent on the  $\xi^\ell$  variables, therefore one can trivially integrate (33) through the  $\eta^\ell$  variables as the integrand has no dependence on these to obtain

$$f_k = \frac{L^3}{8} \int_{-L}^L \int_{-L}^L \int_{-L}^L F_1^k(r_{12}) \exp[-\beta U(r_{12})] d\xi^x d\xi^y d\xi^z,$$

where  $r_{12}$  is the minimum image inter-atomic distance (32). This integral can be written in terms of standard Euclidean distance  $r^2 = (\xi^x)^2 + (\xi^y)^2 + (\xi^z)^2$  as

$$f_k = 8L^3 \int_0^{L/2} \int_0^{L/2} \int_0^{L/2} F_1^k(r) \exp[-\beta U(r)] d\xi, \quad (34)$$

where  $d\xi = d\xi^x d\xi^y d\xi^z$ . In order to analyse  $f_k$  further by implementing a far field approximation, we need to make sure we are in a regime where the integrand is small - we do this by introducing a cutoff  $\gamma$ , which will divide the cube  $[0, L/2]^3$  into 8 cuboid subdomains, including

$$\begin{aligned} \Omega_1 &= [0, \gamma]^3, & \Omega_2 &= [0, \gamma]^2 \times [\gamma, L/2], \\ \Omega_3 &= [0, \gamma] \times [\gamma, L/2]^2, & \Omega_4 &= [\gamma, L/2]^3. \end{aligned}$$

Utilising the symmetry of the problem, we can rewrite integral (34) as

$$f_k = 8L^3 \left( \int_{\Omega_1} + 3 \int_{\Omega_2} + 3 \int_{\Omega_3} + \int_{\Omega_4} \right) F_1^k(r) \exp[-\beta U(r)] d\xi. \quad (35)$$

Considering (35) for  $k = 0$ , the integral over  $\Omega_1$  is independent of  $L$  and provides a bulk contribution to  $f_0$  that will depend on  $\gamma$ . The remaining three terms have integration domains that allow the integrand to be accurately described by a Taylor expansion giving the leading order contribution in the asymptotic limit  $L \rightarrow \infty$  as  $f_0 \propto L^6$ , which can be rewritten in terms of the density,  $n$ , in the form

$$f_0 \propto n^{-2} \quad \text{as } n \rightarrow 0. \quad (36)$$

Considering  $f_k$  for  $k \neq 0$ , the integral over  $\Omega_1$  in equation (35) is again independent of  $L$ . However in the far field expansion the integrals over  $\Omega_2$ ,  $\Omega_3$  and  $\Omega_4$  all decay with  $L$  due to the force factor. As the integration domain has essentially been transformed into that of inter-atomic distances about the three coordinates, when we increase the domain length, the inter-atomic force necessarily decays to 0. Therefore in the limit  $L \rightarrow \infty$  the dominant term arises from integrating over  $\Omega_1$ , and we see that, for  $k = 2, 4, 6, 8, \dots$ ,

$$f_k \propto n^{-1} \quad \text{as } n \rightarrow 0. \quad (37)$$

This leaves us with the final result that in the low density limit  $n \rightarrow 0$ , combining equation (8) with asymptotic expressions (36) and (37),

$$\alpha_k \propto n^{1-k/2} \quad \text{as } n \rightarrow 0. \quad (38)$$

While this result has been calculated for  $N = 2$  interacting atoms, it is also confirmed for larger values of  $N$  by estimating the  $k$ -th standardised moments using MD simulations, as it is shown in the next section.

$N$	$t_{\text{sim}}$	$L_0$	$n_0$
2	$10^9$	5	0.016
8	$10^7$	3	1/64
64	$10^6$	5	1/64
512	$10^4$	10	1/64

TABLE I. The length of MD simulation,  $t_{\text{sim}}$ , the (smallest) box width,  $L_0$ , used for simulations with  $N$  atoms and density  $n_0$  for MD simulations with varying temperatures.

## B. MD simulations with $N$ interacting atoms

In this section we present the results from MD simulations of many-body systems in three spatial dimensions using different values of  $N$ , including the case  $N = 2$  (analyzed in Section IV A). Atoms are subject to pairwise interactions governed by a Lennard-Jones potential, given in equation (3). For each system we use a velocity-Verlet<sup>23</sup> integrator and maintain the system in the canonical ensemble by incorporating a Nosé-Hoover thermostat<sup>37</sup>, see Appendix B. We perform two types of MD simulation studies: those that are used for studying how the number density,  $n$ , of a system affects standardised moments, and those that aim to probe temperature dependency. In all cases we utilise a time step  $\Delta t = 0.01$ . In the case of the simulation with  $N = 2$  atoms, we initialise the positions of atoms by setting  $\mathbf{q}_1 = \mathbf{0}$  and  $\mathbf{q}_2 = (L/2, L/2, L/2)$ , whereas for the  $N = 8, 64, 512$  atom systems, we choose to initialise these on a uniform cubic lattice.

The MD simulation parameters are summarised in Table I, where  $t_{\text{sim}}$  is the total simulation time used for calculating the required statistics, which is preceded by the initial simulation of length  $t_{\text{sim}}/10$  used for equilibrating the system. When investigating the number density dependence, we perform 20 simulations each with a box width of  $L = L_0 \times (6/5)^{i-1}$ , where  $i = 1, 2, \dots, 20$  labels the simulation number and  $L_0$  is the smallest cubic box width. We simulate the  $N = 8, 64, 512$ -atom systems with  $L_0 = 3, 5, 10$ , respectively. This enables direct comparison because we can identify triplets of simulated systems corresponding to systems of the same number densities. The two-atom system however is simulated in a sparser regime with  $L_0 = 5$ . We calculate statistics on the fly for every time step, for every atom and for each coordinate - therefore we average the computed results over the number of time steps ( $t_{\text{sim}}/\Delta t$ ) and atom coordinates ( $3N$ ). In particular, the statistics are calculated over  $3N t_{\text{sim}}/\Delta t$  data points. This is equal to  $6 \times 10^{11}$  (resp.  $1.536 \times 10^9$ ) data points in the simulation with  $N = 2$  (resp.  $N = 512$ ) atoms.

Calculating the number density in three spatial dimensions by  $n = N/L^3$ , we can study the behaviour of kurtosis  $\alpha_4$  as  $n$  varies. The results are presented in Figure 4. We see general agreement between behaviour of each of the four systems. We see when  $n$  is equal, the values of kurtosis are larger for  $N = 2$  than for the many-body systems with  $N = 8, 64, 512$ , which agree well amongst themselves.

The results in Figure 4 enable us to test the asymptotic expression (38) for  $k = 4$  derived in the limit  $n \rightarrow 0$ . Util-

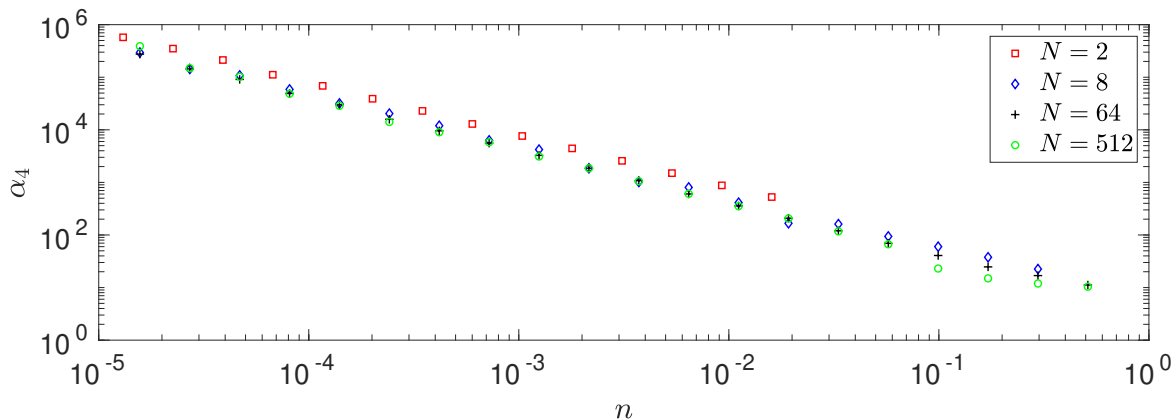


FIG. 4. Dependence of kurtosis  $\alpha_4$  on density  $n$ . Each of the larger atomic systems ( $N = 8, 64, 512$ ) is simulated over the same domain of number densities, while the  $N = 2$  system is simulated in a sparser domain, though all are simulated in three spatial dimensions. We truncate the results of the  $N = 2$  simulation in the plot, however the additional data points are used to calculate the results displayed in Figure 5.

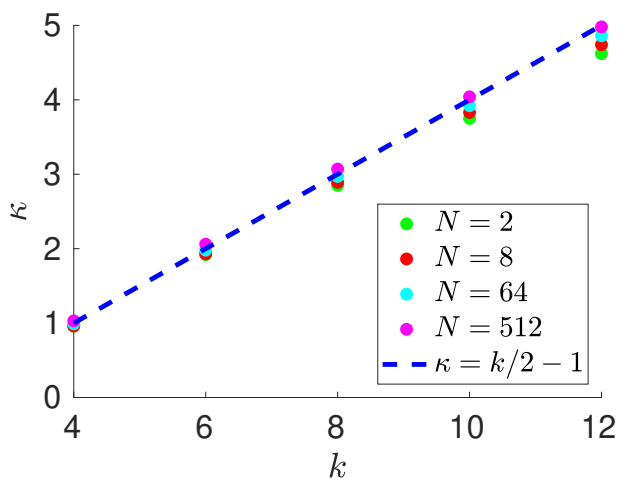


FIG. 5. Comparison of the results of MD simulations for a range of values of the number of atoms,  $N$ . After long time simulation, we compute the asymptotic behaviour  $\alpha_k \propto n^{-\kappa}$  and compare the leading order power scalings for each system. We compare this with the theoretical result (38) (denoted as a blue dashed line) that in the limit  $n \rightarrow 0$  we expect the universal behaviour  $\kappa = k/2 - 1$ , where  $k = 2, 4, 6, \dots$  denotes which standardised moment of force we are looking at.

Using similar log-log plots for MD data, we estimate the power law behaviour of each standardised moment,  $\alpha_k$ , for  $k = 4, 6, 8, 10, 12$ . Figure 5 illustrates the results. All systems agree well with the predicted asymptotic behaviour (38), in particular the  $N = 512$  atom system. There is a slight deviation between the results due to the fact that the smaller atom systems require a larger  $t_{\text{sim}}$  in order to converge fully to the predicted value. This discrepancy is amplified when looking at higher standardised moments due to the fact that we are calculating statistics resulting from  $F_1^{12}$  (i.e. for  $\alpha_{12}$ ) compared to  $F_1^4$  (i.e. for  $\alpha_4$ ), for example.

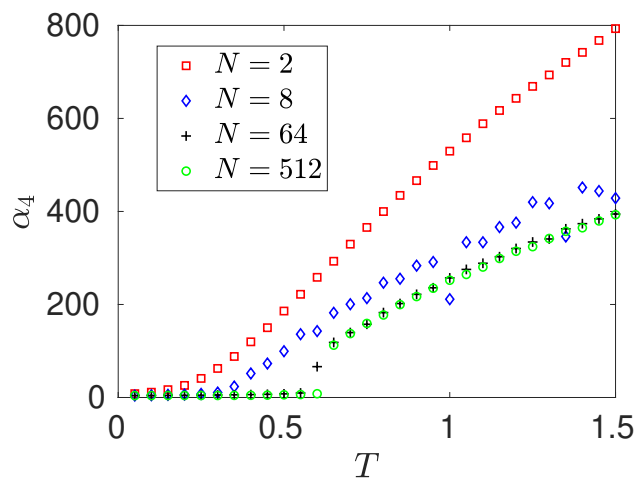


FIG. 6. Dependence of kurtosis  $\alpha_4$  on temperature  $T$ . Each atomic system is simulated at approximately the same density  $n = n_0$  given in Table I.

The dependence of kurtosis  $\alpha_4$  on temperature  $T$  is presented in Figure 6, where we keep the density fixed at  $n = n_0$  given in Table I. We observe that as temperature increases so does the kurtosis of the force distribution associated with each system. This can be explained in terms of the dynamics of the interacting atom system. If we maintain each system in the canonical ensemble, we expect on average that each atom will have a kinetic energy equivalent to  $3T/2$  (when in reduced units). As we increase this target temperature, the atoms become more energetic and thus are able to probe closer interatomic distances before a large repulsive force overcomes this inertial attraction. The range of forces on the tagged particle widens as temperature increases and therefore contributes to more outlier results in the distribution - leading to heavier tails and therefore distributions which become increasingly leptokurtic.



In Figure 6, we observe that there is a qualitative difference between the results for  $N = 2$  and larger atom systems. We see a bifurcation for the  $N = 64$  and  $N = 512$  systems at some temperature  $T_* \in (0.6, 0.65)$ , where a steady increase in kurtosis changes to a rapid increase. This bifurcation point in the phase plane lies on the coexistence boundary with  $(n, T) = (1/64, T_*)$  and is due to a clustering mechanism which has been seen in MD simulations of Lennard-Jones fluids<sup>29</sup>. From our results we see that the  $N = 2$  system has missed this behaviour completely. Snapshots of the  $N = 512$ -atom system at some  $T = 0.6 < T_*$ , and  $T = 0.66 > T_*$  are displayed in Figure 7. For  $T = 0.6$ , we see a large cluster has formed in the many-atom system. There would be far fewer outlier force results in this case due to the fact that the large majority of atoms are moving as a collective and effectively have fixed inter-atomic forces. Compared to the  $T = 0.66$  snapshot, where we see that the atoms are too kinetically unstable to form these larger stable cluster structures, this results in more outlier forces felt between atoms due to the fact that the system is intrinsically more disordered. It is useful to note that this bifurcation point is located on the vapour-liquid coexistence boundary, the mechanisms of which have been studied on dilute Lennard-Jones fluids<sup>30</sup>; here we see that this results in a bifurcation on standardised moments of the force distribution.

To understand the underlying variations of kurtosis,  $\alpha_4$ , with respect to changes in temperature and density, we use  $12 \times 16$  MD simulations with  $N = 512$  atoms and  $t_{\text{sim}} = 3 \times 10^6$ , varying simulation parameters  $(n, T)$ , where  $n = 10^{-2} + (i-1)/10$ , for  $i = 1, 2, \dots, 12$ , and  $T = 10^{-1} + j/10$ , for  $j = 1, 2, \dots, 16$ . The sampled values of excess kurtosis ( $\alpha_4 - 3$ ) are displayed in Figure 8. Here a bifurcation can be seen when using the smallest density  $n = 0.01$ , as the change in colour is prominent in this vertical strip, indicating a large change of kurtosis. This occurs around  $T = 0.6$ , which is consistent with the result in Figure 6, where we saw the bifurcation similarly located, though the slight shift in temperature is accounted for by the shift in density parameters used in each simulation (namely  $n = 0.01$  in Figure 8 and  $n = 1/64$  in Figure 6).

In general, this low density strip contains the largest values of kurtosis, and covers much of the purely gas phase of the Lennard-Jones fluid. This paper has so far probed the low density limit in an attempt to understand why the standardised moments of force are so large, though Figure 8 gives a good overview that in general, regardless of phase, a decrease in temperature, or an increase in density, systematically lead to a lower value of standardised moments. In this case as  $n \rightarrow \infty$  or  $T \rightarrow 0$ , we expect the  $\alpha_4 \rightarrow 3$  (excess kurtosis tends to zero). This limiting regime corresponds to the solid phase of a Lennard-Jones system, where the force variations are minimal and the distribution is Gaussian. There is not enough space, nor energy, that lead to (many) outlier forces experienced by any atom, so the force distribution becomes less and less skewed from Gaussian, the deeper we probe in these regions. This intuition was demonstrated analytically in Section III E when we showed this limiting behaviour on a 1D cartoon model with equation (31). It is interesting to note that

these changes in values of  $\alpha_4$  appear smooth about changes in temperature and density (in absence of the bifurcation point for larger values of  $n$ ), regardless of phase transitions.

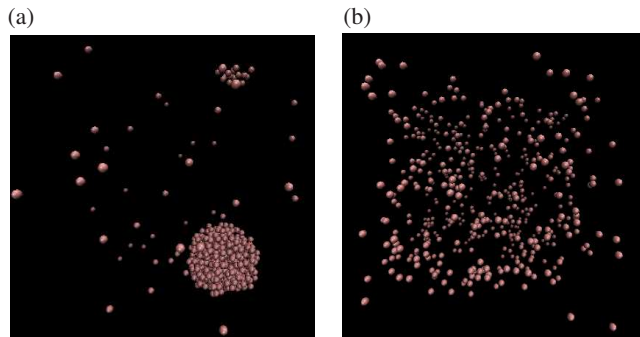


FIG. 7. Snapshots<sup>31</sup> of the MD simulation are taken for the system with  $N = 512$  atoms at time  $t = 7.5 \times 10^5$  for: (a)  $T = 0.6 < T_*$ ; and (b)  $T = 0.66 > T_*$ . Density is  $n = 1/64$ .

## V. DISCUSSION AND CONCLUSIONS

In Section III we have demonstrated use of a variety of methods to study the standardised moments of the force distribution in order to probe both their temperature and number density dependence. This gave way to a rich structure where we show that the partition function for a 1D system can be calculated entirely from these standardised moments. Extending the far field method introduced in Section III B to a system with  $N$  atoms in three-dimensional physical space, Section IV studies the dependence of  $\alpha_k$  on number density  $n$ ,

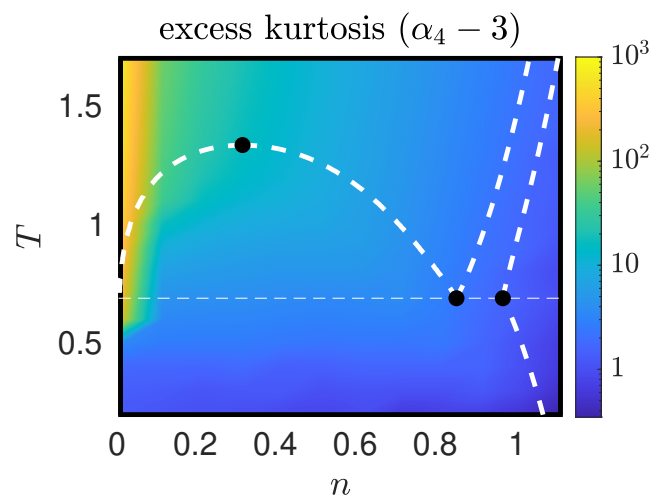


FIG. 8. The excess kurtosis,  $\alpha_4 - 3$ , calculated as a function of density  $n$  and temperature  $T$  for  $n \leq 1.11$  and  $T \leq 1.7$ . The white dotted lines describe coexistence lines of different phases of a Lennard-Jones fluid taken from the literature<sup>32-35</sup>. The solid black dots indicate (from left to right), the critical point and vapour-liquid-solid triple points.

deriving the asymptotic expression (38). Our analytic results are contrasted with MD simulations of four systems of  $N = 2, 8, 64, 512$  interacting Lennard-Jones atoms and these are compared. The results agree well with theoretical predictions though the results for systems with larger values of  $N$  are seen to converge more readily to the theoretically predicted results. In particular, rich dynamics such as clustering of Lennard-Jones fluids is completely missed by the systems with smaller values of  $N$ , but captured for systems with  $N$  as small as  $N = 64$  atoms. In general, as temperature increases  $\alpha_k$  increases due to energetic nature of atoms allowing them to push closer together and experience larger forces. Clustering exhibited at the vapour-liquid coexistence phase incurs a bifurcation point whereby a large increase is seen in the standardised moments of force in Figure 6, though a general increase in temperature, or decrease in number density, results in an increase in  $a_4$  regardless of the temperature/number density domain studied, as shown in Figure 8.

## ACKNOWLEDGMENTS

This work was supported by the Royal Society [grant number RGF\EA\180058] and by the Engineering and Physical Sciences Research Council [grant number EP/V047469/1].

## Appendix A: Constants $C_{k,1}$ and $C_{k,2}$ in equation (29)

The constants appearing in equation (29), namely  $C_{k,1}$  and  $C_{k,2}$ , are given by formulas

$$C_{k,1} = \frac{k(U''(r_*))^{k-1}U^{(3)}(r_*)}{2}, \quad (\text{A1})$$

$$C_{k,2} = \frac{k}{24}(U''(r_*))^{k-2} \times \left( 3(k-1)(U^{(3)}(r_*))^2 + 4U''(r_*)U^{(4)}(r_*) \right), \quad (\text{A2})$$

which can be derived in the following manner. Using  $F^k(r) = (U'(r))^k$  for even values of  $k$  and  $U'(r_*) = 0$ , we first note that

$$F^{k,k}(r_*) = k!(U''(r_*))^k, \\ F^{k,m}(r_*) = 0, \quad \text{for } m \leq k-1,$$

where  $F^{k,m}$  denotes the  $m$ -th derivative of  $F^k$ , i.e. the  $m$ -th derivative of the  $k$ -th power of  $F$ . Therefore, the first three non-zero terms of the Taylor expansion of  $F^k(r)$  around  $r = r_*$  are

$$F^k(r) \approx (r-r_*)^k (U''(r_*))^k + (r-r_*)^{k+1} \frac{F^{k,(k+1)}(r_*)}{(k+1)!} \\ + (r-r_*)^{k+2} \frac{F^{k,(k+2)}(r_*)}{(k+2)!}. \quad (\text{A3})$$

Therefore, we have  $C_{k,1} = F^{k,(k+1)}(r_*)/(k+1)!$  and  $C_{k,2} = F^{k,(k+2)}(r_*)/(k+2)!$  and, to derive equations (A1) and (A2), we need to express derivatives  $F^{k,m}(r_*)$  for  $m = k+1$  and  $m = k+2$  in terms of derivatives of  $U(r)$  at  $r = r_*$ . Using the

product rule, the  $m$ -th derivative of  $F^k(r)$  can be, in general, written as a finite sum of the form

$$F^{k,m}(r) = \sum_{\alpha_0, \alpha_1, \dots, \alpha_m=0}^k C(\alpha_0, \alpha_1, \dots, \alpha_m) \prod_{i=0}^m \left( F^{(i)}(r) \right)^{\alpha_i}, \quad (\text{A4})$$

where  $F^{(i)}(r)$  is the  $i$ -th derivative of function  $F(r)$  and  $C(\alpha_0, \alpha_1, \dots, \alpha_m)$  are constants, many of them equal to zero. In fact, all terms in the expansion (A4) have multiplicities that sum to  $k$ , that is we can only sum over sequences satisfying

$$\sum_{i=0}^m \alpha_i = k, \quad (\text{A5})$$

and all terms in the expansion (A4) have  $m$  derivatives, that is, we have

$$\sum_{i=0}^m i \alpha_i = m, \quad (\text{A6})$$

where  $\alpha_i \in \{0, 1, \dots, k\}$  for  $i = 0, 1, 2, \dots, m$ . Equation (A6) is of the form of a finite Diophantine equation, which has no closed form for the number of solutions. In particular, simplifying equation (A4) by solving equations (A5)–(A6) is, in general, not possible. However, noting the specific property that  $F(r_*) = 0 = U'(r_*)$ , we see that all terms that have  $\alpha_0 \neq 0$  will vanish when evaluated at this unique minimum  $r = r_*$ . In particular, we will obtain relatively simple forms of the sum (A4) for  $m = k+1$  and  $m = k+2$  by considering equations (A5)–(A6) with  $\alpha_0 = 0$ .

First, let us consider that  $m = k+1$ . Using  $\alpha_0 = 0$ , there is only one solution of equations (A5)–(A6) in non-negative integers, namely  $\alpha_1 = k-1$ ,  $\alpha_2 = 1$  and  $\alpha_3 = \alpha_4 = \dots = 0$ . Therefore, equation (A4) implies

$$F^{k,(k+1)}(r_*) = C(0, k-1, 1, 0, \dots, 0) \left( F^{(1)}(r_*) \right)^{k-1} F^{(2)}(r_*).$$

Using the general Leibniz rule<sup>36</sup>, we evaluate the combinatorial prefactor as  $C(0, k-1, 1, 0, \dots, 0) = k(k+1)!/2$ . Substituting into  $C_{k,1} = F^{k,(k+1)}(r_*)/(k+1)!$  and using  $F(r_*) = -U'(r_*)$  and that  $k$  is an even integer, we obtain formula (A1).

Second, we consider the case  $m = k+2$ . Using  $\alpha_0 = 0$ , there are two solutions of equations (A5)–(A6) in non-negative integers. The first solution is  $\alpha_1 = k-1$ ,  $\alpha_2 = 0$ ,  $\alpha_3 = 1$  and  $\alpha_4 = \alpha_5 = \dots = 0$ . The second solution is  $\alpha_1 = k-2$ ,  $\alpha_2 = 2$  and  $\alpha_3 = \alpha_4 = \dots = 0$ . Therefore, equation (A4) implies

$$F^{k,(k+2)}(r_*) \\ = C(0, k-1, 0, 1, 0, \dots, 0) \left( F^{(1)}(r_*) \right)^{k-1} F^{(3)}(r_*) \\ + C(0, k-2, 2, 0, 0, \dots, 0) \left( F^{(1)}(r_*) \right)^{k-2} \left( F^{(2)}(r_*) \right)^2.$$

Using the general Leibniz rule<sup>36</sup>, we evaluate these combinatorial prefactors as

$$C(0, k-1, 0, 1, 0, \dots, 0) = \frac{k}{6} (k+2)!,$$

$$C(0, k-2, 2, 0, 0, \dots, 0) = \frac{k(k-1)}{8} (k+2)!.$$

Substituting into formula  $C_{k,2} = F^{k,(k+2)}(r_*)/(k+2)!$  and using  $F(r_*) = -U'(r_*)$  and that  $k$  is an even integer, we obtain equation (A2). Thus, we have arrived at the expressions for  $C_{k,1}$  and  $C_{k,2}$  that are used in equation (29).

## Appendix B: Thermostats used in MD simulations

Considering 3D simulations in Section IV B, we use a Nosé-Hoover thermostat. Its parameter, originally<sup>37</sup> denoted  $Q$ , is the relaxation time of the thermostat. It is a measure of how strongly the thermostat is attached to the dynamics of the system. We choose a cautious value of  $Q = 10T$  for each simulation; this linear scaling with  $T$  is necessary as we need to more tightly couple the thermostat at lower temperatures in order to accurately maintain the system in the canonical ensemble<sup>38</sup>.

For 1D simulations in Section III, we maintain the canonical ensemble at a target (reduced) temperature  $T$  by implementing a Langevin thermostat. This is due to problems with ergodicity utilising the Nosé-Hoover thermostat for small systems<sup>39,40</sup>. Here the evolution of the free particle is modelled (in reduced units) as<sup>41,42</sup>

$$\ddot{x} = -\frac{dU}{dx} - \gamma\dot{x} + \sqrt{2\gamma T}R(t), \quad (\text{B1})$$

where  $R(t)$  is standard white noise, and  $\gamma$  acts as a friction parameter. We choose  $\gamma = 0.1$  when calculating our illustrative results presented Figures 1 and 3.

## REFERENCES

- <sup>1</sup>S. Joshi and S. Deshmukh, *A review of advancements in coarse-grained molecular dynamics simulations*. Molecular Simulation, DOI: 10.1080/08927022.2020.1828583 (2020)
- <sup>2</sup>Y. Wang et al. *Effective force coarse-graining*. Physical Chemistry Chemical Physics 11, p2002 (2009)
- <sup>3</sup>R. Erban and S. J. Chapman. *Stochastic modelling of reaction-diffusion processes*, Cambridge Texts in Applied Mathematics, Cambridge University Press (2020)
- <sup>4</sup>H. Ingólfsson et al. *The power of coarse graining in biomolecular simulations*. Wiley Interdisciplinary Reviews: Computational Molecular Science 4(3), p225 (2014)
- <sup>5</sup>A. Davtyan et al. *Dynamic force matching: A method for constructing dynamical coarse-grained models with realistic time dependence*. Journal of Chemical Physics 142, 154104 (2015)
- <sup>6</sup>R. Erban. *Coupling all-atom molecular dynamics simulations of ions in water with Brownian dynamics*. Proceedings of the Royal Society A 472(2186):20150556 (2016)
- <sup>7</sup>D. Wales. *Exploring energy landscapes*. Annual Review of Physical Chemistry 69, p401 (2018)
- <sup>8</sup>R. Gunaratne et al. *On short-range and long-range interactions in multi-resolution dimer models*. Interface Focus 9 (3), rsfs.2018.0070 (2019)
- <sup>9</sup>E. Rolls, Y. Togashi and R. Erban. *Varying the resolution of the Rouse model on temporal and spatial scales: application to multiscale modelling of DNA dynamics*. Multiscale Modeling and Simulation 15(4), p1672 (2017)
- <sup>10</sup>R. Erban. *From molecular dynamics to Brownian dynamics*, Proceedings of the Royal Society A 470(2167): 20140036 (2014)
- <sup>11</sup>L. DeCarlo. *On the meaning and use of kurtosis*. Psychological Methods 2(3), p292 (2014)
- <sup>12</sup>A. Carof, R. Vuilleumier and B. Rotenberg. *Two algorithms to compute projected correlation functions in molecular dynamics simulations*. Journal of Chemical Physics 140, 124103 (2014)
- <sup>13</sup>H. Shin et al. *Brownian motion from molecular dynamics*. Chemical Physics 375, p316 (2010)
- <sup>14</sup>R. Erban. *Coarse-graining molecular dynamics: stochastic models with non-Gaussian force distributions*. Journal of Mathematical Biology 80, p457 (2020)
- <sup>15</sup>S. Chandrasekhar. *Stochastic problems in physics and astronomy*. Review of Modern Physics 15:1 (1943)
- <sup>16</sup>A. Gabrielli et al. *Force distribution in a randomly perturbed lattice of identical atoms with  $1/r^2$  pair interaction*. Physical Review E 74:021110 (2006)
- <sup>17</sup>G. Rickayzen et al. *Single atom force distributions in simple fluids*. Journal of Chemical Physics 137, 094505 (2012)
- <sup>18</sup>A. C. Branka, D. M. Heyes and G. Rickayzen. *Pair force distributions in simple fluids*. Journal of Chemical Physics 135, 164507 (2011)
- <sup>19</sup>J. Jones. *On the determination of molecular fields. — II. From the equation of state of a gas*. Proceedings of the Royal Society A, 106:738 (1924)
- <sup>20</sup>H. Watanabe, N. Ito and C. Hu. *Phase diagram and universality of the Lennard-Jones gas-liquid system*. Journal of Chemical Physics 136, 204102 (2012)
- <sup>21</sup>J-P. Hansen and L. Verlet. *Phase transitions of the Lennard-Jones system*. Physical Review 184:151 (1969)
- <sup>22</sup>A. Rahman. *Correlations in the motion of atoms in liquid argon*. Physical Review A 136(2) (1964)
- <sup>23</sup>L. Verlet. *Computer "experiments" on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules*. Physical Review 159(1) (1967)
- <sup>24</sup>D. Frenkel and B. Smit. *Understanding molecular simulation: from algorithms to applications. 2nd ed.* Academic Press (1996)
- <sup>25</sup>L. Rowley, D. Nicholson and N. G. Parsonage. *Monte Carlo grand canonical ensemble calculation in a gas-liquid transition region for 12-6 Argon*. Journal of Computational Physics 17, p401 (1975)
- <sup>26</sup>D. Lide. *Properties of the elements and inorganic compounds; melting, boiling, triple, and critical temperatures of the elements*. CRC Handbook of Chemistry and Physics (86th ed.), CRC Press, Chapter 4 (2005)
- <sup>27</sup>A. Giró, E. Guardia and J. A. Padró. *Langevin and molecular dynamics simulations of Lennard-Jones liquids*. Journal of Molecular Physics 55, Issue 5 (1985)
- <sup>28</sup>C. M. Bender and S. A. Orszag. *Advanced mathematical methods for scientists and engineers: asymptotic methods and perturbation theory: v. 1*. Springer (1999)
- <sup>29</sup>N. Yoshii and S. Okazaki. *Molecular dynamics study of structure of clusters in supercritical Lennard-Jones fluid*. Fluid Phase Equilibria 144(1-2), p225 (1998)
- <sup>30</sup>J. Jung, J. Lee and J. Kim. *Cluster growth mechanisms in Lennard-Jones fluids: a comparison between molecular dynamics and Brownian dynamics simulations*. Chemical Physics 449, p1 (2015)
- <sup>31</sup>W. Humphrey, A. Dalke and K. Schulten. *VMD - Visual Molecular Dynamics*. Journal of Molecular Graphics 14, p33 (1996)
- <sup>32</sup>A. Schultz and D. Kofke. *Erratum: "Comprehensive high-precision high-accuracy equation of state and coexistence properties for classical Lennard-Jones crystals and low-temperature fluid phases"*. Journal of Chemical Physics 153, 059901 (2020)
- <sup>33</sup>S. Stephan et al. *Thermophysical properties of the Lennard-Jones fluid: database and data assessment*. Journal of Chemical Information and Modeling 2019 59 (10), 4248 (2019)
- <sup>34</sup>S. Stephan, J. Staubach and H. Hasse. *Review and comparison of equations of state for the Lennard-Jones fluid*. Fluid Phase Equilibria 523, 112772 (2020)
- <sup>35</sup>E. Mastny and J. de Pablo. *Melting line of the Lennard-Jones system, infinite size, and full potential*. Journal of Chemical Physics 127, 104504 (2007)
- <sup>36</sup>A. Traheem and A. Laradji. *Classroom note: a generalization of Leibniz rule for higher derivatives*. International Journal of Mathematical Education in Science and Technology 34(6): 905 (2003)
- <sup>37</sup>S. Nosé. *A molecular dynamics method for simulations in the canonical ensemble*. Molecular Physics 52(2), p255 (1984)
- <sup>38</sup>P. H. Hünenberger. *Thermostat algorithms for molecular dynamics simulations*. In: Dr. Holm C., Prof. Dr. Kremer K. (eds) Springer Advanced Computer Simulation. Advances in Polymer Science 173, p105 (2005)
- <sup>39</sup>M. E. Tuckerman et al. *Non-Hamiltonian molecular dynamics: generalizing Hamiltonian phase space principles to non-Hamiltonian systems*. Journal of Chemical Physics 115, p1678 (2001)
- <sup>40</sup>P. F. Tupper. *Ergodicity and the numerical simulation of Hamiltonian systems*. SIAM Journal on Applied Dynamical Systems 4(3), p563 (2005)
- <sup>41</sup>T. Schlick. *Molecular modeling and simulation*. Springer (2002)
- <sup>42</sup>B. Leimkuhler and C. Matthews. *Molecular dynamics with deterministic and stochastic numerical methods*. Springer Interdisciplinary Applied Mathematics 39 (2015)