

Statistical fluctuations along the Lennard-Jones melting curve

R. Patra*

*Department of Economics, Royal Holloway, University of London,
Egham, Surrey TW20 0EX, United Kingdom*

D.M. Heyes[†]

*Department of Physics, Royal Holloway, University of London,
Egham, Surrey TW20 0EX, United Kingdom[‡]*

(Dated: December 2, 2015)

Abstract

Statistical fluctuations and correlations between thermodynamic properties along the fluid side of the melting line of the Lennard-Jones (LJ) are determined using Molecular Dynamics (MD) computer simulation. Linear regression, the Pearson coefficient and other statistical measures are calculated. The cross correlation between the configurational part of the pressure and potential energy, and the repulsive and attractive parts of the potential energy are focussed on. Regression plots show that at constant temperature and constant total energy the Weeks-Chandler-Andersen (WCA) decomposition of the Lennard-Jones repulsive and attractive potential energy components show a qualitative change along the melting line. At low temperature the two components are correlated, while they are anticorrelated in the high temperature limit. There is an intermediate temperature range in which the two potential energy components are effectively uncorrelated. The various fluctuation trends along the melting line were found to be weakly dependent on the force field used to generate the distribution of states, namely, the LJ potential, inverse power potential with exponent 12, and the repulsive term in the WCA decomposition of the LJ potential.

PACS numbers:

*Electronic address: r.patra@rhul.ac.uk

†Electronic address: david.heyesh@rhul.ac.uk

‡Keywords: Pearson Coefficient, fluctuations, Molecular Simulation

I. INTRODUCTION

Statistical fluctuations and their analysis are used in a wide range of disciplines such as Economics and Physics to confirm the validity of conclusions derived from experimental data and also as model building tools in their own right . Although the source and nature of the data sets could hardly be more different in these two fields, the statistical methods used to extract significance from the data (apart perhaps from the terminology used to describe them) are remarkably similar. In condensed matter Physics any thermodynamic system over time exhibits a distribution of thermodynamic state variable values which depend on the ensemble adopted. Such fluctuations in first order thermodynamic quantities can be used as a route to calculate second order thermodynamic quantities (*e.g.*, heat capacity and compressibility), and they have been used in molecular simulation studies over many decades to perform this task, [1]. Relatively recently a new use for system property fluctuations has been proposed, and that is to identify those states on the phase diagram that have (to a good approximation) an underlying scale invariance which has been called *isomorphism*, [2–7] because of scaleable or self-similar underlying assembly structures of these thermodynamic state points. The statistical quantity called the Pearson coefficient, R_p , has been used to locate these states on the phase diagram. The Pearson coefficient while widely used in Economics, [8, 9] had not been used to any great extent in condensed matter Physics until this recent isomorphism application, as far as we are aware.

Consider a point in the configurational phase space of N molecules which may represent the molecules in a periodic simulation periodic cell, where \underline{r}_i is the coordinate of molecule i , and the configurational phase state point is represented in concise form by, $\underline{r}^N \equiv \underline{r}_1 \underline{r}_2 \cdots \underline{r}_N$. If $\rho = N/V$ is the number density of the N molecules in volume V , and $\tilde{r}_i \equiv \rho^{1/3} r_i$ is a

non-dimensionalised coordinate, two state points on the phase diagram defined by density and temperature are said to be isomorphic if the probability distribution function of these states, $P(\tilde{\underline{r}}^N)$, for all $\tilde{\underline{r}}^N$ in the two thermodynamic state points is the same. By extension an isomorphic line on the phase diagram in the density and temperature plane is one along which all state points have the same $P(\tilde{\underline{r}}^N)$ function.

Assuming pair-wise additivity of the potential energy surface, the analytic form of the pair potential is, in addition to the density and temperature, the most important factor in controlling the extent of isomorphic behaviour. The Lennard-Jones (LJ) potential is one of the most used classical representations of model condensed phase systems, and is defined by,

$$\phi_{LJ}(r) = 4\epsilon[(\sigma/r)^{12} - (\sigma/r)^6], \quad (1)$$

where ϵ and σ are the characteristic energy of interaction and diameter of the molecule, respectively, and r is the separation between the centres of two of the molecules. The inverse power potential, $\phi_{IP}(r) = 4\epsilon(\sigma/r)^n$, is another relevant potential in the present context, where n is an exponent which governs the steepness of the potential. The inverse power (IP) fluid and solid are examples of perfectly isomorphic condensed phases for which the isomorphic line is defined through the relationship, $\rho^{n/3}/T = const$, where T is the temperature (using the usual molecule-based reduced units). The LJ potential is the sum of two such inverse power terms, and the $n = 12$ IP fluid or solid could be considered to be a possible reference system for the LJ system, with the $n = 6$ IP attractive part of the potential being taken as a first order perturbation.

The ‘melting line’ on the phase diagram is where a transition between a fluid and solid (crystalline) state takes place. In fact, it is only a line when plotted in the P, T plane, where P

is the pressure. On the ρ, T and ρ, P planes, there are coexisting region ‘gaps’ between distinct fluid and solid single phase zones. Knowledge of the melting line (ML) of a chemical system is important in various chemically relevant fields as the physical state of the molecules can have a strong influence on the physical behaviour (*e.g.*, flow characteristics) of the system. This is important in, for example, geology and high pressure (elastohydrodynamic) lubrication. The melting line is already known to be almost isomorphic, which in part explains the success of various phenomenological ‘rules’ of melting that have been found to apply very well for many different types of molecule, [10].

The Pearson coefficient derived from the correlation between the configurational part of the pressure, P_c , and the potential energy, u , has been used as a convenient measure of the extent to which two state points are isomorphic [3–7, 11]. If the pressure-energy correlation measure, R_p , is equal to unity, the two states would be completely isomorphic, which is the case for the IP system. In reality only IP fluids form isomorphic lines, so $0 \leq R_p \leq 1$ for all other model systems having repulsive and attractive components in their interaction potential. The closer R_p is to unity the more ‘isomorphic’ the two state points can be said to be.

The present study uses the Pearson coefficient to assess the degree of correlation between the configurational part of the pressure, P_c , and the total potential energy per particle, u , along the melting line of the LJ system, using Molecular Dynamics computer simulation to generate the states and their fluctuations. This is relevant to the proof of isomorphism, and in particular the extent to which the melting line can be said to be isomorphic. In addition, and we think for the first time, the corresponding correlation between the potential energy components derived from repulsive and attractive decompositions of the pair potential is explored. The treatments explored here could be used to choose those decompositions of

the LJ potential which lead to the least correlation between the properties derived from the repulsive and attractive parts of the potential according to a chosen decomposition, of which there is an infinite number of possibilities. The Weeks-Chandler-Andersen decomposition of the LJ potential (see below) is a popular choice but by no means the only one. Ultimately such a procedure could lead to improved perturbation theories of the liquid state. The perturbation theories rely on the structure of the liquid being determined to a large extent by the repulsive part of the potential, and the attractive part leading to mean field background correction, so to be most effective it is desirable that these two terms be as weakly correlated as possible. This information we will show is manifest in simple regression plots in a pictorially revealing way.

II. SIMULATION AND DEFINITION DETAILS

The Lennard-Jones and other pair potentials were employed to generate the molecular configurations reported here using the leapfrog version of the Verlet algorithm to generate the molecular trajectories. All quantities presented are given in LJ reduced units (*i.e.*, ϵ for energy, and σ for distance). The potential energy, pair force and other static properties were obtained by including explicitly the LJ potential interaction of a molecule pair separation up to a demarcation distance, $r_d = 2.5$, [12] and then tapered to zero between $2.5 \leq r \leq 2.7$ using the Mason-Song formula, to minimise the effects of force truncation. [13] The usual mean field long range correction formulas, [12] were added to the potential energy and other static properties based on r_d . The time step was $0.005/\sqrt{T}$, and the simulations were conducted for up to 10^6 time steps during a post-equilibration stage. The number of particles in the simulation cell, N , was 2048 or 4000, to quantify any finite size effects. Molecular dynamics (MD) simulations were carried out in the microcanonical (NVE) ensemble, and the constant temperature

ensemble using either velocity rescaling, [14, 15] or Nosé-Hoover thermostats. [16–18] State points on the fluid side of the melting line terminating at the triple point at *ca.* T, ρ values of 0.69, 0.85 [19] were simulated (ρ is defined here as the reduced number density in dimensionless units, $N\sigma^3/V$, where V here is the volume of the cubic simulation cell). The state points simulated were determined via a polynomial fit to several sources of molecular simulation fluid-solid coexistence data taken from the literature (*e.g.*, [19–21]).

Simulations were carried out using different pair potentials to generate the state points, with some being carried out with the LJ potential. The Weeks-Chandler-Andersen (WCA) decomposition of the LJ potential into a steeply repulsive (‘r’) and a smoothly varying (‘background’) attractive (‘a’) part respectively is as follows, [22–31]

$$\begin{aligned}
 \phi_{WCA,r}(r) &= \phi_{LJ}(r) + \epsilon, & r \leq r_c, \\
 \phi_{WCA,r}(r) &= 0 & r > r_c, \\
 \\
 \phi_{WCA,a}(r) &= -\epsilon, & r \leq r_c, \\
 \phi_{WCA,a}(r) &= \phi_{LJ}(r) & r > r_c,
 \end{aligned} \tag{2}$$

where $r_c = 2^{1/6}\sigma$ is the position of the minimum of the LJ potential, and $\phi_{LJ}(r) = \phi_{WCA,r}(r) + \phi_{WCA,a}(r)$ for all r . Some simulations were carried with $\phi_{WCA,r}(r)$, and others using the inverse power potential, [32] $\phi(r) = 4\epsilon(\sigma/r)^{12}$ to generate the configurations, for the same values of T and ρ . The values of the thermodynamic properties of the ‘virtual’ LJ and WCA potential systems were also computed even for state distributions generated by the other two force fields.

The virial expression for the pressure, P , was used in the simulations, [1, 12]

$$P = \frac{1}{3V} \left[\sum_{i=1}^N \frac{1}{m} \underline{p}_i \underline{p}_i + \frac{1}{2} \sum_{i=1}^N \sum_{j \neq i}^N r_{ij} f_{ij} \right] \quad (3)$$

where V is the volume of the system, \underline{p}_i is the translational momentum of molecule, i , of mass m , $\underline{r}_{ij} = \underline{r}_i - \underline{r}_j$, and \underline{r}_i and \underline{r}_j are the coordinates of molecules using the *nearest image*, i and j , respectively. The pair force between the molecules is $f_{ij} = -d\phi_{ij}/dr_{ij}$, again applying the nearest image convention between molecules i and j . The kinetic contribution to the total pressure is the first term in Eq. (3), which by equipartition can be replaced for equilibrium systems by $\underline{P}_k = \rho k_B T$, where k_B is Boltzmann's constant. The second term on the right hand side of Eq. (3), is the configurational part of the pressure, denoted by P_c . The potential energy per particle is $u = \langle \sum_{i < j} \phi(r_{ij}) \rangle / N$, where $\langle \dots \rangle$ represents a simulation average (the configurational part of the pressure is similarly averaged). For the LJ potential this can be decomposed into repulsive (' r ') and attractive (' a ') parts, *i.e.*, $u_r = 4 \langle \sum_{i < j} \epsilon(\sigma/r_{ij})^{12} \rangle / N$ and $u_a = -4 \langle \sum_{i < j} \epsilon(\sigma/r_{ij})^6 \rangle / N$, respectively. The LJ potential can also be decomposed into the two WCA contribution parts as defined in Eq. (2), $u_{WCA,r} = \langle \sum_{i < j} \phi_{WCA,r}(r_{ij}) \rangle / N$ for the WCA repulsive potential component, and $u_{WCA,a} = \langle \sum_{i < j} \phi_{WCA,a}(r_{ij}) \rangle / N$ for the attractive component. In the literature, the potential term, $\phi_{WCA,r}$ is often just referred to as the 'WCA' potential.

Three temperature and density fluid states along the LJ melting line were considered. The three temperatures were 0.7, 4.0 and 60, and the corresponding densities were 0.847, 1.229 and 2.289, respectively. The theory of statistical fluctuations relating to linear regression and the Pearson coefficient is covered in Sec. III. Application of this theory to the simulation data is made in Sec. IV. The correlation between P_c and u is computed, as these two quantities were first used to test for isomorphism in previous molecular simulation studies [6]. Correlations

between two decompositions of the total potential energy are also assessed for these three state points and various force fields used to generate the dynamics. The results from IP and WCA (repulsive part only) and full LJ dynamics are compared. Section V is mainly concerned with a time-dependent extension of the Pearson coefficient criterion. Conclusions are made in Sec. VI.

III. THEORY AND RESULTS FOR P_c AND u CORRELATIONS FOR LJ

In this section the directional relationships between the several variables are analysed. The analysis of the variance, Pearson's Rank Correlation Coefficient, [33] which is sometimes referred to as *Pearson's Product-Moment Correlation coefficient* or PCC for short, and Ordinary Least Squares (OLS) regression techniques, [34] are used for this purpose. Pairs of variables are treated and the standard t -test, [35] is carried out to establish the statistical significance of the derived relationship. The goodness of fit of the correlation between the variables, for example, $u_{WCA,r}$ and $u_{WCA,a}$ is used to verify to what extent the relationship between these variables is linear.

Several basic statistical concepts and the relationship between them are covered first, in order to interpret properly the PCC. One of the most commonly used measures of how the points in a data set are distributed is the second central moment around the mean. The 'variance' of a variable, A , or σ_A^2 , is the mean squared deviation from its mean for a given sample of data, $\sigma_A^2 \equiv Var(A) = \mathbb{E}[(A - \mathbb{E}[A])^2] = \mathbb{E}[A^2] - (\mathbb{E}[A])^2$, where \mathbb{E} is the expectation value of A . (*i.e.*, $\mathbb{E}[A] = \sum_{i=1}^N A_i/N$ for the i -th value of A in a data set). The variance measures how spread out about the mean the distribution of data points is. A variance of zero means all the values of A in a data set have the same value, and the variance is always ≥ 0 , of course. The 'standard deviation', denoted by σ_A , is the square root of the variance, which in standard notation is, $\sigma_A = \sqrt{\mathbb{E}[A^2] - (\mathbb{E}[A])^2}$, which should not be confused here with the particle diameter, σ , in the potential.

A related quantity, the 'covariance' is a measure of the 'strength' of the linear relationship

between two variables A and B ,

$$\text{Cov}(A, B) = \mathbb{E}[(A - \mathbb{E}[A])(B - \mathbb{E}[B])] = \mathbb{E}[AB] - \mathbb{E}[A]\mathbb{E}[B]. \quad (4)$$

If $\text{Cov}(A, B) > 0$, then on average A increases as B increases and *vice versa*. If $\text{Cov}(A, B) < 0$, then A tends to decrease as B increases and *vice versa*. These quantities are important when it comes to defining the PCC, which is denoted here by, $R_{p,A,B}$, to specify explicitly the involvement of quantities A and B . Correlation is a measure of the directional relationship between the paired elements in two data sets, A and B , and,

$$R_{p,A,B} = \frac{\text{Cov}(A, B)}{\sigma_A \sigma_B} = \frac{\mathbb{E}[(A - \bar{A})(B - \bar{B})]}{\sqrt{\mathbb{E}(A^2) - [\mathbb{E}(A)]^2} \sqrt{\mathbb{E}(B^2) - [\mathbb{E}(B)]^2}}, \quad (5)$$

where $\text{Cov}_{A,B}$ is the covariance between data sets A and B , and σ_A is the standard deviation of data set A (and the same notation for B). The average value of A is denoted by \bar{A} and the average value of B is denoted by \bar{B} . Note that the PCC is dimensionless while covariance has units obtained by multiplying the units of the two variables. The PCC is a measure of the ‘strength’ of the relationship between the two variable sets, but does not define any particular functional relationship (or ‘causality’) between the two variables taken at the same time or in a particular order. By ‘causality’ we do not necessarily mean that a value in A leads directly to the corresponding value in B or *vice versa*, but that both quantities may be determined by an underlying third parameter of the system (*e.g.*, the partition function of the system in statistical mechanics). This latter point limits our ability to draw a causal relationship between the two variables, and for this reason an additional procedure, known as ‘Regression Analysis’, (RA) or in the present context of assumed proportionality between two variables, ‘Linear Regression’ (LR), which involves minimising the sum-of-the-squares of the errors is widely used to draw inferences about any causal relationship between the variables. The RA involves the method of Ordinary Least Squares (OLS). Below the procedure of regression is

defined and used to establish a formal link between the PCC and the OLS regression coefficient.

Regression analysis is the process of constructing a mathematical model or function that can be used to predict or determine the value of one variable from that of another variable, or other variables. The most elementary regression model is called ‘simple regression’. In simple regression, the variable to be predicted is called the dependent variable, and is usually designated by Y . The independent variable, or ‘explanatory’ variable, usually designated by X is also called the ‘predictor’. The procedure of simple regression involves fitting a straight line through a set of N_p points in such a way that the sum of the squared residuals of the model is minimised. The equation of this line is,

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i, \quad (6)$$

where, \hat{Y}_i is the predicted value of Y_i using a finite number of sample sets, $\hat{\beta}_0$ is the y -intercept of the line of best fit, and $\hat{\beta}_1$ is the slope of the line of best fit. The difference between the actual and predicted value of the dependent variable, called the ‘residual’, is,

$$\hat{U}_i = Y_i - \hat{Y}_i = Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i. \quad (7)$$

The residuals are the vertical distances between the points of the data set and the fitted line. Intuitively it is readily appreciated that the smaller the residuals the closer the fit line is to the distribution of input pair values. To avoid the problem of positive residuals offsetting negative residuals, the principle of Ordinary Least Squares (OLS) is employed, which involves finding the values of $\hat{\beta}_0$ and $\hat{\beta}_1$ which minimise the sum of the squares of the residuals, S ,

$$S = \sum_{i=1}^{N_p} \hat{U}_i^2 = \sum_{i=1}^{N_p} (Y_i - \hat{Y}_i)^2 = \sum_{i=1}^{N_p} (Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_i)^2. \quad (8)$$

By minimising the above equation with respect to $\hat{\beta}_0$ (intercept) and $\hat{\beta}_1$ (slope) expressions for these two quantities are obtained,

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}, \quad (9)$$

where $\hat{\beta}_0$ is the predicted intercept, and the predicted slope, $\hat{\beta}_1$ is,

$$\hat{\beta}_1 = \frac{\mathbb{E}[(X - \bar{X})(Y - \bar{Y})]}{\mathbb{E}(X^2) - [\mathbb{E}(X)]^2} = \frac{Cov(X, Y)}{Var(X)}, \quad (10)$$

for the N_p data points, and where, \bar{X} is the mean value of the explanatory variable X , and \bar{Y} is the mean value of the dependent variable, Y .

The relationship between the OLS estimator and Pearson's Correlation coefficient is now derived. The formula for the estimator, $\hat{\beta}_1$, is given by,

$$\hat{\beta}_1 = \frac{Cov(A, B)}{\sigma_B^2} = \frac{\mathbb{E}[(A - \bar{A})(B - \bar{B})]}{\mathbb{E}(B^2) - [\mathbb{E}(B)]^2}. \quad (11)$$

which using Eq. (5) gives, $\hat{\beta}_1 = R_{p,A,B}\sigma_A/\sigma_B$. This relationship reveals that the regression coefficient is the PCC times the ratio of the standard deviations of the independent variable divided by that of the dependent variable. This signifies that regression analysis provides additional information when compared to the Pearson coefficient, namely, the relative distribution spreads of the two variables. The OLS coefficient, $\hat{\beta}_1$, will be used here to analyze the relationship between the two variables, in addition to the PCC. The sign of Pearson's correlation coefficient and that of the OLS coefficient are the same, as $\sigma_A \geq 0$ and $\sigma_B \geq 0$. To establish how well the predicted line fits the data, the ratio,

$$R^2 = \frac{Var(\hat{Y})}{Var(Y)}, \quad (12)$$

is used, where R^2 measures the fractional variation in the dependent variable given by the model. The predicted value of Y is \hat{Y} , and the input value of the treatment is Y .

The so-called t -test, [36] is used here to establish whether the OLS estimator is significantly different from zero, (*i.e.*, the slope is statistically significant based on the number and distribution of data points) through the parameter, t , where $t = \hat{\beta}_1/SE(\hat{\beta}_1)$ and $SE(\hat{\beta}_1)$ is the

standard error of $\hat{\beta}_1$.

The above analysis is now used to establish the extent of correlation between the following pairs of variables, (u, P_c) , (u_r, u_a) and $(u_{WCA,r}, u_{WCA,a})$ where the right entry is taken to be the dependent variable and the left entry to be the independent variable. In the latter two sets it is reasonable to take the repulsive energy term to be the independent variable as this is consistent with perturbation theories of liquids where the structure of the liquid is assumed to be dominated by the repulsive part of the potential. For (u, P_c) there appears to be no clear preference for which of the quantities should be taken to be the independent variable as they are formally different system measures, and both include the repulsive and attractive parts of the potential energy (although weighted differently). The adopted choice is therefore arbitrary.

IV. LINEAR REGRESSION AND THE PEARSON COEFFICIENT

The linear regression and PCC as defined in Fig. (5) are first explored for system states generated using the LJ potential for three state points in the low and high temperature limits, and one in the middle.

Figure 1 shows (a) P_c as a function of u , (b) u_a as a function of u_r and (c) $u_{WCA,a}$ as a function of $u_{WCA,r}$ respectively from left to right. The dynamics and state points are generated using the LJ potential. The state point used has a temperature of 60 and a density of 2.289 in LJ reduced units. The degree of correlation is measured by the extent to which the data points fall on or near the regression straight line. Figure 1 shows data points for $N = 2048$ using the Nosé-Hoover (NH) thermostat (with a time constant of 3 LJ units) and the velocity rescaling constant temperature thermostat (VS), and $N = 2048$ using the NVE or microcanonical ensemble. Also some symbols are shown for $N = 4000$ using the NH thermostat. There is no significant difference between the distributions generated through these different constraints, and they all show the same trends. There is very little difference between $N=2048$ and 4000 within the simulation statistics. It would be surprising if there were no difference at all, but any differences between using $N=2048$ and 4000 do not affect the conclusions.

Figures 2 and 3 give the corresponding plots for the temperature, density pairs of (4.00, 1.229) and (0.70, 0.847), respectively. These three state points are on the fluid boundary side of the LJ melting line. The dynamics and state points are again generated using the LJ potential. Table I presents further statistical analysis of these nine data sets, which quantifies the degrees and nature of the correlation between the three pair fluctuation quantities. The (u, P_c) correlations are strongest, and have a positive slope with very little dispersion or scatter of

points about the line. The extent of linearity in the correlation between this pair has been used to determine the extent to which lines on the density-temperature planes of the phase diagram are isomorphic (*i.e.*, have an underlying structural invariance),[3–7]. The u_r and u_a are also quite strongly linked but with a negative slope, which indicates that they are *anticorrelated*. Even at constant temperature, one expects the attractive part of the potential to change in the opposite direction to a change in the repulsive part. Both of these trends are evident along the whole of the melting line.

The behaviour of the $u_{WCA,a}, u_{WCA,r}$ pair shown in Figs. 1-3, in contrast, changes qualitatively along the melting line. The slope goes from being negative to positive as the temperature (density) decreases, and is approximately infinite in the region, $T \sim 10$. This change in behaviour has to be associated with the analytic form of the repulsive and attractive parts of the WCA decomposition of the LJ potential, and the dynamic distribution of near neighbour molecules. As the (temperature) density decreases the near neighbour particles move further apart, and the repulsive part of the potential is weaker. This suggests a weakening of the anticorrelation coupling between $\phi_{WCA,r}$ and $\phi_{WCA,a}$ energy terms. In fact it becomes slightly correlated close to the triple point, as shown in Fig. 3.

The radial distribution function for the three fluid state points generated using the LJ potential force field are shown in Fig. 4. The lower set of curves uses the pair separation, r in LJ σ on the abscissa. The top set of radial distribution functions expressed in isomorphic distance units, $\tilde{r} = \rho^{1/3}r$, show excellent isomorphic collapse. The peaks of $g(r)$ shift to smaller distances with increasing density. In fact, the first peak of all three are to varying extents within the truncation distance of $\phi_{WCA,r}$ *i.e.*, $r_c = 2^{1/6}$, which is shown as a vertical line on the figure. For $T = 60$ the whole of the first peak is within r_c while only about half of it is in this range for

$T = 0.7$. The correlation to anticorrelation transition with increasing T and ρ could possibly be attributable to these variations, or more many-body consequences of these trends. Note that the $g(r)$ are liquid-like for all three state points, showing no evidence of any crystalline order.

The analysis used to generate Figs. 1-3 was made for systems generated by the LJ potential. Two purely repulsive potentials, which are formed from the repulsive region of the LJ potential are now considered as generators of the system dynamics. One form, called here, ‘IP12’, is the IP potential with $n = 12$, *i.e.*, $\phi(r) = 4\epsilon(\sigma/r)^{12}$, which does not include any of the attractive part of the LJ potential. The other purely repulsive potential is the repulsive part of the LJ or WCA interaction, $\phi_{WCA,r}$, which does include the short range region of the attractive part of the LJ potential, up to r_c . The potential (or derived force) used to generate the system of states is referred to as the ‘force field’ here. The results of these simulations are summarised in Figs. 5-7, and in Table II.

Table III compares the various averages for different N , thermostats and from an NVE simulation. There is no statistically significant difference between the velocity rescaling and NH thermostats. The NVE fluctuation quantities show the same trends but can be up to *ca.* 10 % different in magnitude from the corresponding NVT cases. One would not expect fluctuation-based quantities to be exactly the same in these two ensembles, but for these cross-correlation quantities the trends are the same, as is also evident in Figs. 1-3.

Figure 5 compares the same three pairs of computed property as in Figs. 1-3, given along the rows. Each row gives the results from a different force field. The bottom row is derived from LJ potential dynamics. The middle row used the (repulsive) WCA force field, and the top row

from the IP12 potential. The temperature and density of the state point are 0.70 and 0.847 for each of the nine frames. First, the figure shows that the WCA and IP12 potentials generate very similar pair-property correlation behaviour to the LJ case. The u_r and u_a are strongly anticorrelated more or less equally for the three force fields. The figure also indicates that the $u_{WCA,r}$, $u_{WCA,a}$ pair are relatively weakly correlated, especially for IP12, indicated by the closer to circular ellipsoidal pattern of symbols on the figure (top right frame). This weak correlation trend is understandable as the WCA decomposition of the LJ potential was originally chosen to partition it into a strongly repulsive part and a slowly varying component that are weakly correlated, as an objective for their use in perturbation theories of the liquid state, [22, 37–39]. In the perturbation theory the attractive part of the WCA potential is treated essentially as a background term and the structure is governed by the repulsive part of the WCA decomposition.

Figure 6 presents the same set of correlations for a state point in which the temperature and density are 4.00 and 1.229, respectively. The u_r , u_a pair are again strongly anticorrelated for all force fields. The $u_{WCA,r}$, $u_{WCA,a}$ pair are even more weakly correlated than in Fig. 5, for all force fields, as evident by the nearly circular pattern of symbols for all frames in the rightmost column. Superficially at least it appears that the two potential terms are statistically independent, which could be made use of in developing perturbation theories of the liquid state, as these two components appear to be statistically independent over a certain temperature (density) range along the melting curve. Another noteworthy feature is that again the distribution of points for each type of correlation is largely independent of the dynamics generating force field, along the melting line at least.

Figure 7 presents the corresponding data for a temperature of 60 and density of 2.289. The three u_r and u_a are anticorrelated to more or less the same extent as found in the previous

two figures. The three $u_{WCA,r}$, $u_{WCA,a}$ pair reveal strong *anticorrelation* this time. This change in behavior could be explained because the repulsive part of the potential becomes relatively more important at higher temperatures (density), and larger fluctuations in this part of the potential energy will take place which will induce oppositely signed fluctuations in the attractive part of the potential (*i.e.*, the constant temperature system is not too far from the microcanonical ensemble in this limit). As discussed above, a key issue in this respect could be where the first peak in the radial distribution function lies in relation to the distance r_c in the WCA decomposition of the LJ potential at a given density.

Table II gives a summary of the simulation average property values for the three state points considered in Figs. 5-7, which are well separated along the melting line. The table shows that as temperature (density) increases the total LJ potential energy, u , shifts in the positive direction, especially for simulations carried out with the IP potential using $n = 12$ (or ‘IP12’) force field dynamics. The difference in the total energy from LJ and WCA dynamics is insignificant at $T = 60$ and not very great for $T = 0.7$, a result which is consistent with the applicability of the WCA potential in liquid state perturbation theory. The average potential energy, u , values from the *LJ* and *WCA* dynamics are not too different, and typically within a few percent of each other, while that of the IP12 force field is much more positive, which becomes more accentuated with increasing temperature along the melting curve. The PCC for the three pairs of quantities are shown in the last three columns of the table, which shows that the Pearson coefficient for the P_c and u pair of quantities is very close to unity for all of the state points and force fields considered. Its value increases towards unity with increasing temperature. Just why the Pearson Coefficient is so close to unity for this pair of system quantities is not immediately obvious. One might expect there to be a reasonably strong correlation between u and P_c as the latter has a component of u_r in its definition. In fact,

for the Lennard-Jones potential, $P_c/\rho T = 4u_r + 2u_a$, [40] Indeed, all static properties of the LJ system can be expressed as a linear combination of the average repulsive and attractive parts of the potential, apart from some known constants or numerical factors. The strong anticorrelation between u_r and u_a may also contribute to the proximity of the PCC to unity, as then the repulsive and attractive terms can be combined into one effective (less repulsive) quantity. The table shows that the quantity, $R_p[u_r.u_a]$, is close to -1 for the three state points. The behaviour in R_p for the pair, $u_{WCA,r}$ with $u_{WCA,a}$, is quite different, as noted above. The absolute value is much less than unity and for all types of force field dynamics. It is sensitive to state point and R_p goes from positive to negative in the temperature interval between 4 and 60. It is also quite sensitive to the force field, in going more positive through the sequence: IP12, LJ to WCA. At higher temperature there is evident more anticorrelation between the positive and attractive parts of the potential.

A limitation of linear regression and the Pearson coefficient is that it does not give any indication of the time or chronological persistence of the correlation between the two quantities along the data set. In fact, any randomly sorted array of a two column table would give the same regression plots and PCC values. However consecutive data values in a table can be correlated with each other, which generally decays to a statistically uncorrelated state between two data points far enough apart along the table. This is useful information which could give further insights into the underlying physics. An extension of the Pearson correlation concept which gives this additional information is proposed and tested in the next section.

V. TIME CORRELATION PEARSON MODIFICATION

The degree of correlation between the same or two different quantities at times separated by an interval, t , can be expressed as,

$$R_{p,A,B}(t) = \frac{\langle \delta A(0)\delta B(t) \rangle}{\langle \delta A(0)^2 \rangle^{1/2} \langle \delta B(0)^2 \rangle^{1/2}}. \quad (13)$$

where A and B are again the two system quantities of interest. In Eq. (13) the quantity, $\delta A(t) = A(t) - \bar{A}$ and $\delta B(t) = B(t) - \bar{B}$. The function in Eq. (13) is an extension of the Pearson coefficient definition (which is the $t = 0$ value) to account for temporal correlations between the two quantities. Such a formula is widely used in Economics and is known as regression with n -lagged explanatory variable [34]. In liquid state physics, if $A = B$ then this is called an ‘autocorrelation’ function, whereas if $A \neq B$ in Eq. (13) it is referred to as a ‘cross-correlation or perhaps ‘Pearson’ correlation function is a more appropriate term in the present context. The quantity defined in Eq. (13) is closely related to the time-correlation function used to explore the dynamics and calculate transport coefficients of fluids by MD with Green-Kubo formulae, [41] but in that case the denominator is set to $\langle \delta A(0)\delta B(0) \rangle$ or unity (respectively) instead. The only significant difference between the various forms is the normalisation factor used is therefore in the denominator. An informative step in the present context is to express time in isomorphic units defined by, $\tilde{t} = \rho^{1/3}T^{1/2}t$. Along an isomorphic line time dependent properties scale with time expressed as \tilde{t} . Time dependent properties along an isomorph should collapse onto the same curve if the ordinate quantity is suitably normalised (this is referred to as *isochronal* scaling, [42]). The time-dependent function, $R_{p,A,B}(\tilde{t})$ from Eq. (13) expressed in terms of isomorphic time quantifies the time persistence of any correlation between A and B over time. It gives some information on how long it takes the values between two variables to become essentially statistically uncorrelated.

The above treatment is another statistical analysis tool which quantifies the relationship between variables, where an explanatory variable can influence the dependent variable even with a time lag. These are known as distributed lag models in the time series literature, which are usually formulated as follows,

$$Y_t = \beta_0 + \beta_1 X_{t-1} + U_t \quad (14)$$

where Y_t is the functional value at time t for the input vector, X_{t-1} , and β_1 measures the effect of the explanatory variable one increment of time in the past to the dependent variable, *ceteris paribus*. The residual vector at time t is denoted by U_t . Many lagged variables $t - 1, t - 2, \dots$ as far back as is necessary can be included in the set of explanatory variables to account fully for memory effects. The extent of the time lag can be chosen by using t -tests for every subsequent addition of a lagged explanatory variable. The OLS estimation gives the best fit to the data, the statistical significance of which can be established using the t -test and other statistical measures to prove the data is stationary, that is when the mean, variance, autocorrelation of the data are constant within the data statistics.

Figure 8 shows $R_{p,A,B}(\tilde{t})$ for the same fluid state points and quantities as given in Table I, where the LJ force field has been used to generate the dynamics. The three Pearson cross-correlation functions shown on the figure are, $\langle u(0)P_c(\tilde{t}) \rangle$, $\langle u_r(0)u_a(\tilde{t}) \rangle$ and $\langle u_{r,WCA}(0)u_{a,WCA}(\tilde{t}) \rangle$. The first two functions decay monotonically to zero from a positive or negative initial value, and to a very good approximation exhibit isochronal collapse along the studied melting line. The corresponding WCA quantity has a quite different time dependence even when cast in isomorphic units, which is consistent with the data in Table I (*i.e.*, the time equal to zero value of this function). The $t = 0$ value goes from being positive to negative with increasing temperature, and at a certain temperature $R_{p,A,B}(0)$ must be zero for each type of dynamics.

Simulations carried out at that state point could therefore be useful in the development of perturbation theory descriptions of the liquid state. There is a long-time tail in these functions, having not achieved zero by 0.5 isomorphic time units. Figures 9 and 10 show the corresponding $R_p(\tilde{t})$ produced by WCA and IP12 forcefield dynamics. The features and trends are qualitatively and quantitatively basically the same as for LJ dynamics as shown in Fig. 8.

VI. CONCLUSIONS

The statistical fluctuation behaviour of pairs of thermodynamic properties are examined for the Lennard-Jones (LJ) fluid along its melting curve. Regression plots of the instantaneous values of the two variables are used to determine the slope and intercept using standard linear regression analysis. The Pearson coefficient was also calculated, which has been used recently to determine the extent to which a line on the phase diagram is isomorphic (*i.e.*, has an underlying structural invariance) taking the two variables to be the configurational part of the pressure and the potential energy of the system. The statistical analysis has been extended here to include the correlation between the repulsive and attractive parts of the LJ potential, and also that of its Weeks-Chandler-Andersen (WCA) decomposition. The regression plots between these two energy components are shown to be informative in regard to assessing the degree and nature of any correlations between the two quantities. At constant temperature and constant total energy the former are strongly anticorrelated along the melting line whereas the latter goes from being weakly correlated near the triple point to being moderately anticorrelated in the high temperature (density) limit.

The present analysis approach gives new insights into the nature of any correlation between the attractive and repulsive parts of the total LJ potential energy which will underpin its structure and thermodynamic properties. This treatment shows the extent to which the repulsive part of the potential determines the behaviour of the attractive part, that is, the extent to which they are correlated. As many perturbation theories of the liquid state are based on minimising the correlation between the two, the present approach could in future work be used as a useful complementary technique to analyse various decompositions of the LJ potential (other than the WCA one) which may have advantages in this respect.

An extension of the Pearson coefficient method to determine time dependent correlations is also proposed, and shown to give new insights into the temporal behaviour or persistence of system property correlations.

The statistical trends are shown to be relatively insensitive to the potential used to generate the dynamics provided it is purely repulsive and constructed from the LJ potential either as the r^{-12} inverse power part or the repulsive part of the WCA reconstruction of the LJ potential.

References

- [1] J.-P. Hansen and I.R. McDonald, *Theory of simple liquids*, 4th Ed. Academic Press: 2013.
- [2] D.M. Heyes, D. Dini and A.C. Brańka, *Scaling of Lennard-Jones liquid elastic moduli, viscoelasticity and other properties along fluid-solid coexistence*, Phys. Stat. Solidi B, **252**, 1-12 (2015).
- [3] T.S. Ingebrigtsen, T.B. Schröder and J. C. Dyre, *Isomorphs in model molecular liquids*, J. Phys. Chem. B **116**, 1018-1034 (2012).
- [4] J.C. Dyre, *Hidden Scale Invariance in condensed matter*, J. Phys. Chem. **118**, 10007-10024 (2014).
- [5] N.P. Bailey, T.B. Schröder and J.C. Dyre, *Variation of the dynamic susceptibility along an isochrone*, Phys. Rev. E, **90**, 042310 (2014).
- [6] N. Gnan, T.B. Schröder, U.R. Pedersen, N.P. Bailey, J.C. Dyre, *Pressure-energy correlations in Liquids. IV. "Isomorphs" in liquid phase diagrams*. J. Chem. Phys. **131**, 234504 (2009).
- [7] U.R. Pedersen, N.P. Bailey, T.B. Schröder and J.C. Dyre, *Strong pressure-energy correlations in van der Waals liquids*, Phys. Rev. Lett. **100**, 015701 (2008).
- [8] C. Cortinhas and K. Black, *Statistics for Business and Economics*, Wiley: 2012.
- [9] M.K. Fung, *Are knowledge spillovers driving the convergence of productivity among firms?*, Economica **72**, 287 - 305 (2005).
- [10] W.G. Hoover and M. Ross, *Statistical theories of melting*, Contemp. Phys. **12**, 339-356 (1971).
- [11] D. Coslovich and C.M. Roland, *Pressure-energy correlations and thermodynamics scaling in viscous Lennard-Jones liquids*, J. Chem. Phys. **130**, 014508 (2009).
- [12] D.M. Heyes, *The liquid state - applications of molecular simulations*, John Wiley & Sons: 1998.
- [13] J.R. Morris and X. Song, *The melting lines of model systems calculated from coexistence simulations*, J. Chem. Phys. **116**, 9352-9358 (2002).

- [14] L. V. Woodcock, *Isothermal molecular dynamics calculations for liquid salts*, Chem. Phys. Lett. **10**, 257-261 (1971).
- [15] L.V. Woodcock, *Comparison of thermostats*, CCP5 Information Quarterly No. 24, March 1987, p. 297.
- [16] S. Nosé, *A molecular dynamics method for simulations in the canonical ensemble*, Mol. Phys. **52**, 255-268 (1984).
- [17] W. G. Hoover, *Canonical dynamics: Equilibrium phase-space distributions*, Phys. Rev. A **31**, 1695-1697 (1985).
- [18] S. Nosé, *Constant temperature molecular dynamics methods*, Prog. Theor. Phys. Supplement **103**, 1-117 (1991).
- [19] M.A. Barroso and A.L. Ferreira, *Solid-fluid coexistence of the Lennard-Jones system from absolute free energy calculations*, J. Chem. Phys. **116**, 7145-7150 (2002).
- [20] A. Ahmed and R.J. Sadus, *Solid-liquid equilibria and triple points*, J. Chem. Phys. **131**, 174504 (2009).
- [21] G. C. McNeil-Watson and N.B. Wilding, *Freezing line of the Lennard-Jones fluid: A phase switch Monte Carlo study*, J. Chem. Phys. **124** 064504 (2006)
- [22] J.-P. Hansen and L. Verlet, *Phase transitions of the Lennard-Jones system*, Phys. Rev. **184**, 151-161 (1969).
- [23] J.D. Weeks, D. Chandler and H.C. Andersen, *Role of repulsive forces in determining the equilibrium structure of simple liquids*, J. Chem. Phys. **54**, 5237-5247 (1971).
- [24] J.D. Weeks, D. Chandler and H.C. Andersen, *Relationship between the hard-sphere fluid and fluids with realistic repulsive forces*, Phys. Rev. A **4**, 1597-1607 (1971).
- [25] H.C. Andersen, D. Chandler and J.D. Weeks, *Roles of repulsive and attractive forces in liquids: The optimized Random Phase approximation* J. Chem. Phys. **56**, 3812-3823 (2006).

- [26] D. Chandler, J.D. Weeks and H.C. Andersen, *van der Waals picture of liquids, solids, and phase transformations*, Science **220**, 787-794 (1983).
- [27] H.C. Andersen, D. Chandler and J.D. Weeks, *Roles of repulsive and attractive forces in liquids: The equilibrium theory of classical fluids*, Adv. Chem. Phys. **34** 105-156 (1976).
- [28] D.M. Heyes and H. Okumura, *Equation of state and structural properties of the Weeks-Chandler-Andersen fluid*, J. Chem. Phys. **124**, 164507 (2006).
- [29] A. de Kuiper, J.A. Schouten and J.P.J. Michels, *The melting line of the Weeks-Chandler-Andersen Lennard-Jones reference system*, J. Chem. Phys. **93**, 3515 (1990).
- [30] S. Pikatan Tan, H. Adidharma and M. Radosz, *Weeks-Chandler-Andersen model for solid-liquid equilibria in Lennard-Jones systems*, J. Phys. Chem. C **106**, 7878-7881 (2002).
- [31] J.-P. Hansen, *Phase transition of the Lennard-Jones system. II. High-temperature limit*, Phys. Rev. A, **2**, 221-230 (1970).
- [32] D.M. Heyes, S.M. Clarke and A.C. Brańka, *Soft-sphere soft glasses*, J. Chem. Phys. **131**, 204506 (2009).
- [33] J. P. Onyango and A.M. Plews, *A textbook of basic statistics*, East African Publishers : 1987.
- [34] J.M. Wooldridge, *Introductory econometrics: A modern approach*, South-Western Publishers, 4th Ed. : 2009.
- [35] J.K. Adams, *Basic statistical concepts*, McGraw-Hill Book Company, Inc.: 1955.
- [36] D.A. McQuarrie, *Mathematical methods for scientists and engineers*, University science books: 2003.
- [37] D. Ben-Amotz and G. Stell, *Hard sphere perturbation theory for fluids with soft-repulsive-core potentials*, J. Chem. Phys. **120**, 4844-4851 (2004).
- [38] S. Hess, M. Kröger and H. Voigt, *Thermomechanical properties of the WCA-Lennard-Jones model system in its fluid and solid states*, Physica A **250**, 58-82 (1998).

- [39] A. Ahmed and R.J. Sadus, *Phase diagram of the Weeks-Chandler-Andersen potential from very to high temperatures and pressures*, Phys. Rev. E **80**, 061101 (2009).
- [40] R. Zwanzig and R.D. Mountain, *High-frequency elastic moduli of simple fluids*, J. Chem. Phys. **43**, 4464 (1965).
- [41] D.M. Heyes, G. Rickayzen and A.C. Bránka, *Static properties and time correlation functions of fluids with steeply repulsive potentials*, Mol. Phys. **102**, 2057-2070 (2004).
- [42] T.B. Schröder, N. Gnan, U.R. Pedersen, N.P. Bailey, J.C. Dyre, *Pressure-energy correlations in Liquids. V. Isomorphs in generalised Lennard-Jones systems*. J. Chem. Phys. **134**, 164505 (2011).
- [43] © StataCorp. 2001. Statistical Software: Release STATA 12. College Station, TX: Stata Corporation.

The statistical analysis for this table was carried out using the software package, © STATA [43]

| State | Pair | RP | Coefficient | Std. Error | t | $p > t $ | 95% Conf. Interval | R^2 |
|-------|-----------------------|-----------------|------------------------|-----------------------|--------|-----------|--|--------|
| S1 | $P_c u$ | $\hat{\beta}_1$ | 0.932 | 0.037 | 25.3 | 0.000 | 0.860 - 1.005 | |
| S1 | $P_c u$ | $\hat{\beta}_0$ | 0.0106 | 0.0168 | 0.630 | 0.529 | -0.0225 - 0.0438 | 0.764 |
| S1 | $u_a u_r$ | $\hat{\beta}_1$ | -0.917 | 0.023 | -39.08 | 0.000 | -0.964 - 0.8714 | |
| S1 | $u_a u_r$ | $\hat{\beta}_0$ | 0.958 | 0.013 | 71.79 | 0.000 | 0.932 - 0.984 | 0.885 |
| S1 | $u_{WCA,a} u_{WCA,r}$ | $\hat{\beta}_1$ | 0.557 | 0.059 | 9.51 | 0.000 | 0.441 - 0.672 | |
| S1 | $u_{WCA,a} u_{WCA,r}$ | $\hat{\beta}_0$ | 0.119 | 0.030 | 3.93 | 0.000 | 0.059 - 0.178 | 0.314 |
| S2 | $P_c u$ | $\hat{\beta}_1$ | 1.001 | 0.005 | 207.3 | 0.000 | 0.9917 - 1.011 | |
| S2 | $P_c u$ | $\hat{\beta}_0$ | 4.96×10^{-4} | 2.46×10^{-3} | 0.20 | 0.840 | -4.35×10^{-3} - 5.35×10^{-3} | 0.995 |
| S2 | $u_a u_r$ | $\hat{\beta}_1$ | -0.897 | 0.022 | -40.3 | 0.000 | -0.940 - 0.853 | |
| S2 | $u_a u_r$ | $\hat{\beta}_0$ | 0.945 | 0.012 | 81.8 | 0.000 | 0.922 - 0.968 | 0.892 |
| S2 | $u_{WCA,a} u_{WCA,r}$ | $\hat{\beta}_1$ | 0.289 | 0.063 | 4.59 | 0.000 | 0.165 - 0.412 | |
| S2 | $u_{WCA,a} u_{WCA,r}$ | $\hat{\beta}_0$ | 0.340 | 0.034 | 9.99 | 0.000 | 0.273 - 0.407 | 0.0963 |
| S3 | $P_c u$ | $\hat{\beta}_1$ | 0.999 | 0.001 | 835 | 0.000 | 0.997 - 1.002 | |
| S3 | $P_c u$ | $\hat{\beta}_0$ | -9.72×10^{-6} | 5.78×10^{-4} | -0.02 | 0.987 | -1.15×10^{-3} - 1.13×10^{-3} | 0.9997 |
| S3 | $u_a u_r$ | $\hat{\beta}_1$ | -0.801 | 0.033 | -24.6 | 0.000 | -0.866 - 0.737 | |
| S3 | $u_a u_r$ | $\hat{\beta}_0$ | 0.886 | 0.0178 | 49.7 | 0.000 | 0.850 - 0.921 | 0.753 |
| S3 | $u_{WCA,a} u_{WCA,r}$ | $\hat{\beta}_1$ | -0.540 | 0.061 | -8.92 | 0.000 | -0.659 - 0.420 | |
| S3 | $u_{WCA,a} u_{WCA,r}$ | $\hat{\beta}_0$ | 0.704 | 0.0283 | 24.9 | 0.000 | 0.648 - 0.760 | 0.287 |

TABLE I: Linear single variable OLS regression for $N = 2048$ and NVT (velocity rescaling) data.

Key: State point S1 is T, ρ equal to 0.7, 0.8468, state point S2 is T, ρ equal to 4.0, 1.229 and state point S3 is 60.0, 2.289. There are 200 data points. The regression parameter is 'RP'. The t -test value is denoted by ' t ', which should not be confused with time here, which also is denoted by t . 'Conf. Interval' is the confidence interval. For the first two rows, for example, the p value on $\hat{\beta}_0$ is greater than 0.05, hence $\hat{\beta}_0$ is not a significant predictor³¹ of the real y -intercept. The p value on $\hat{\beta}_1$ does not exceed 0.05, and therefore P_c is a significant predictor of u . As P_c increases by one unit u increases by 0.93 units. The value of R^2 indicates that P_c explains 76.4% of the variation in u . The statistical

| Dynamics | T | ρ | u_r | u_a | u | $u_{WCA,r}$ | $u_{WCA,a}$ | $P_c u$ | $u_r u_a$ | $uWruWa$ |
|----------|-----|--------|-------|---------|--------|-------------|-------------|---------|-----------|----------|
| IP12 | 0.7 | 0.847 | 11.7 | -14.03 | -2.35 | 4.11 | -6.46 | 0.995 | -0.930 | 0.425 |
| LJ | 0.7 | 0.847 | 5.80 | -11.94 | -6.13 | 0.601 | -6.74 | 0.959 | -0.944 | 0.537 |
| WCA | 0.7 | 0.847 | 6.10 | -12.06 | -5.96 | 0.712 | -6.67 | 0.972 | -0.949 | 0.603 |
| IP12 | 4 | 1.230 | 58.43 | -30.48 | 27.94 | 37.27 | -9.33 | 0.999 | -0.926 | 0.151 |
| LJ | 4 | 1.230 | 24.27 | -24.62 | -0.349 | 9.12 | -9.46 | 0.998 | -0.951 | 0.335 |
| WCA | 4 | 1.230 | 24.39 | -24.66 | -0.265 | 9.19 | -9.46 | 0.998 | -0.952 | 0.344 |
| IP12 | 60 | 2.289 | 781.7 | -108.67 | 673.01 | 690.53 | -17.53 | 1.000 | -0.920 | -0.298 |
| LJ | 60 | 2.289 | 293.6 | -85.39 | 208.17 | 225.43 | -17.25 | 1.000 | -0.951 | -0.571 |
| WCA | 60 | 2.289 | 293.4 | -85.38 | 208.02 | 225.27 | -17.25 | 1.000 | -0.952 | -0.580 |

TABLE II: Thermodynamic averages and the Pearson coefficient, R_p , for three fluid state points the configurations of which are generated according to three force fields specified in the first column. The velocity rescaling thermostat was used and $N = 2048$. Note that $u = u_r + u_a = u_{WCA,r} + u_{WCA,a}$. The acronym, ‘IP12’ indicates dynamics generated using the IP potential with $n = 12$. The acronym ‘WCA’ indicates that the MD dynamics were produced using the repulsive part of the LJ potential, *i.e.*, $\phi_{WCA,r}$. Also $uWruWa$ stands for $u_{WCA,r}u_{WCA,a}$. The R_p values for the three quantity correlations are given in the last three columns. The standard errors are *ca.* 2 in the third significant figure for the pressure, energy averages, and $P_c u$ and $u_r u_a$, and *ca.* 0.02 for $uWruWa$.

| T | ρ | P_{cu} | P_{cu} | $u_r u_a$ | $u_r u_a$ | $uWra$ | $uWra$ | $uWruWa$ | $uWruWa$ |
|-----|---------|----------|----------|-----------|-----------|--------|--------|----------|----------|
| | | 2048VS | 2048NH | 2048VS | 2048NH | 2048VS | 2048NH | 4000NH | 2048NVE |
| 0.7 | 0.84685 | 0.995 | 0.995 | -0.930 | -0.930 | 0.425 | 0.434 | 0.432 | 0.309 |
| 0.7 | 0.84685 | 0.959 | 0.958 | -0.944 | -0.940 | 0.537 | 0.540 | 0.550 | 0.377 |
| 0.7 | 0.84685 | 0.972 | 0.974 | -0.949 | -0.951 | 0.603 | 0.611 | 0.609 | 0.502 |
| 4 | 1.2295 | 0.999 | 0.999 | -0.926 | -0.928 | 0.151 | 0.153 | 0.133 | 0.166 |
| 4 | 1.2295 | 0.998 | 0.998 | -0.951 | -0.952 | 0.335 | 0.320 | 0.327 | 0.220 |
| 4 | 1.2295 | 0.998 | 0.998 | -0.952 | -0.954 | 0.344 | 0.349 | 0.315 | 0.273 |
| 60 | 2.2893 | 1.000 | 1.000 | -0.920 | -0.926 | -0.298 | -0.316 | -0.332 | -0.234 |
| 60 | 2.2893 | 1.000 | 1.000 | -0.951 | -0.956 | -0.571 | -0.597 | -0.582 | -0.434 |
| 60 | 2.2893 | 1.000 | 1.000 | -0.952 | -0.955 | -0.580 | -0.594 | -0.579 | -0.470 |

TABLE III: Thermodynamic averages and the Pearson coefficient, R_p , for the three fluid state points using various system sizes, thermostats and an NVE simulation. Key to simulation details: $N = 2048$ and a velocity rescaling thermostat is labelled, ‘2048VS’. $N = 2048$ and the Nosé-Hoover thermostat is labelled, ‘2048NH’. $N = 4000$ and the Nosé-Hoover thermostat is labelled, ‘4000NH’. $N = 2048$ and the NVE ensemble is labelled, ‘2048NVE’. The standard errors are *ca.* 2 in the third significant figure for P_{cu} and $u_r u_a$, and *ca.* 0.01 for $uWruWa$.

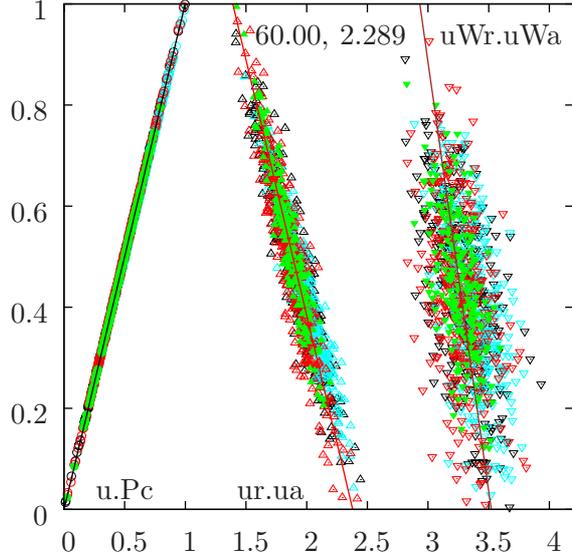


FIG. 1: The correlation from left to right (a) P_c as a function of u (b) u_a as a function of u_r and (c) $u_{WCA,a}$ ($'uWa'$ on the figure) as a function of $u_{WCA,r}$ ($'uWr'$ on the figure), where the quantities plotted are the differences from their means. The state point is $T = 60.00$ and $\rho = 2.289$, using the LJ potential to generate the dynamics, which is high up on the melting curve. The data in each pair quantity correlation is normalised to fall within $0 - 1$ for the abscissa and ordinate quantities. The black symbols are for $N = 2048$ using the Nosé-Hoover thermostat, the green symbols are for $N = 2048$ using the velocity rescaling constant temperature thermostat, and the red symbols are for $N = 2048$ in the NVE or microcanonical ensemble. The cyan coloured symbols are for $N = 4000$ using the Nosé-Hoover thermostat. The solid lines are least square fits to the $N = 2048$ and velocity rescaling data sets.

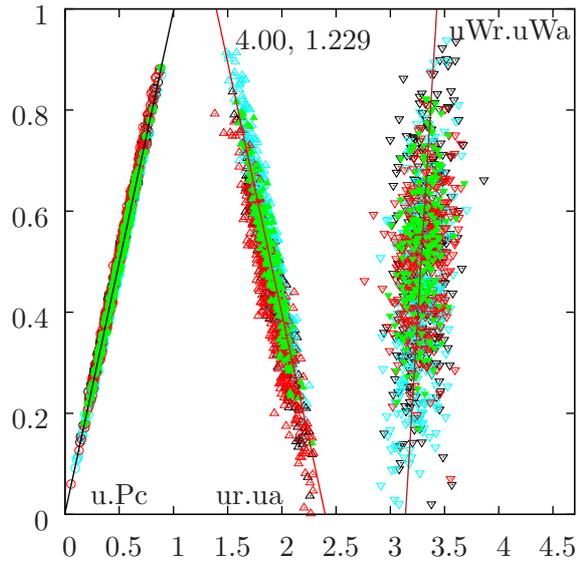


FIG. 2: As for Fig. 1 except that the state point is $T = 4.00$ and $\rho = 1.229$ is used.

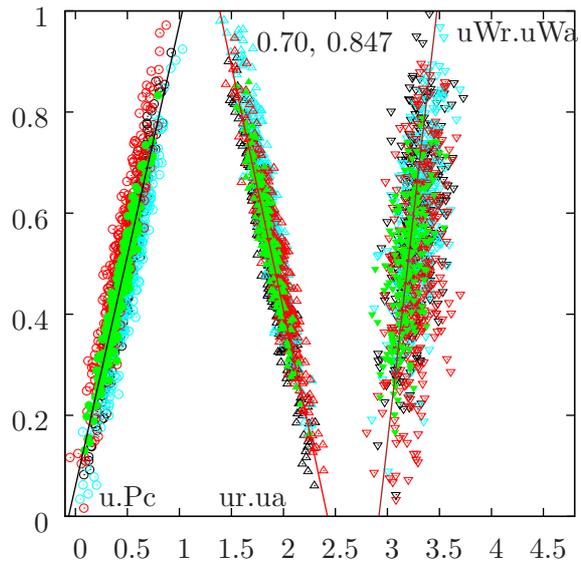


FIG. 3: As for Fig. 1 except that the state point is $T = 0.700$ and $\rho = 0.847$ is used.

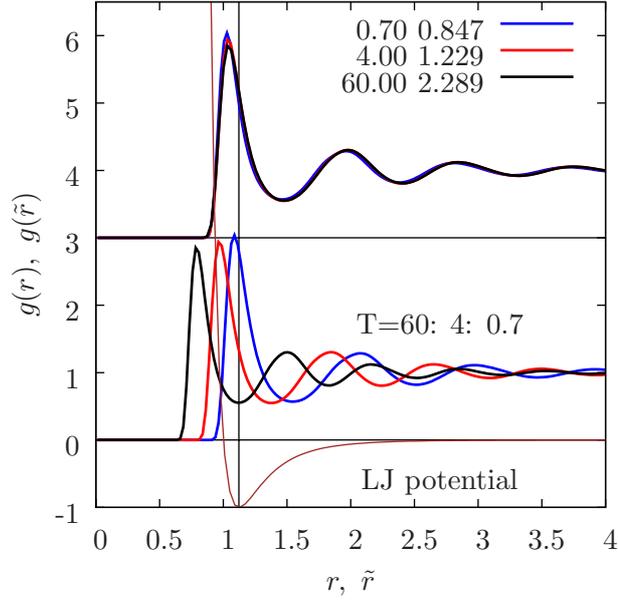


FIG. 4: The radial distribution function, $g(r)$ expressed in LJ distance units, and in isomorphic distance units, $\tilde{r} = \rho^{1/3}r$ for the three state points along the fluid side of the coexistence curve. The upper set of curves, shifted upwards by 3 is $g(\tilde{r})$ and the lower set are for the $g(r)$. The LJ potential is also shown as a thin red line curve. The vertical black line corresponds to the position of the minimum in $\phi_{LJ}(r)$, which is equal to $2^{1/6}$. The velocity rescaling thermostat and $N = 2048$ were used for these simulations.

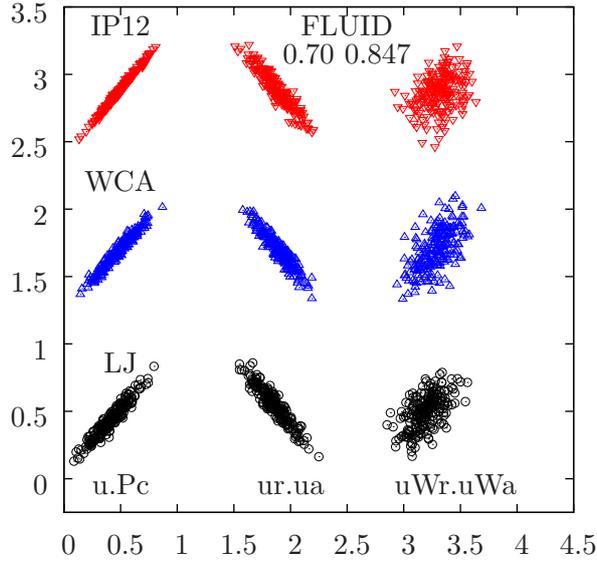


FIG. 5: The correlation between (from left to right columns, respectively), (a) LJ u and P_c , denoted by $u.Pc$ on the figure, (b) LJ u_r and u_a , denoted by $ur.ua$ on the figure, and (c) $u_{WCA,r}$ and $u_{WCA,a}$, denoted by $uWr.uWa$ on the figure, where the quantities are the differences from their means. The rows indicate from bottom to top, (a) LJ, (b) WCA and (c) IP12 force field dynamics. The near-triple point state point, $T = 0.70$ and $\rho = 0.847$, is considered. The data in each frame is normalised to fall within $0 - 1$ for the abscissa and ordinate quantities.

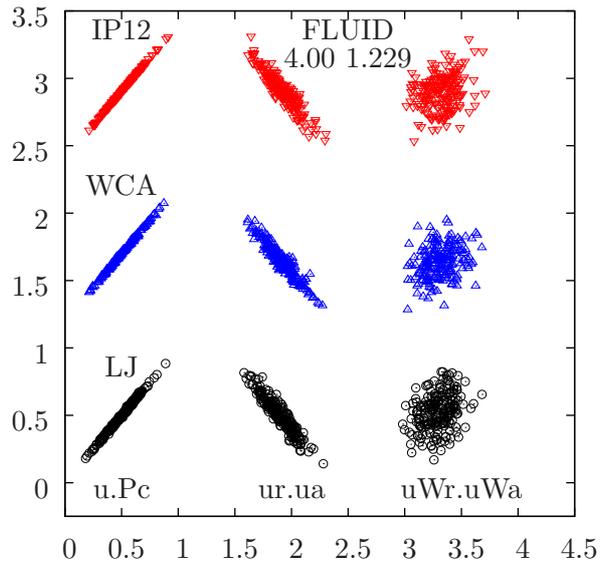


FIG. 6: As for Fig. 5, except the state point $T = 4.00$ and $\rho = 1.229$ is considered.

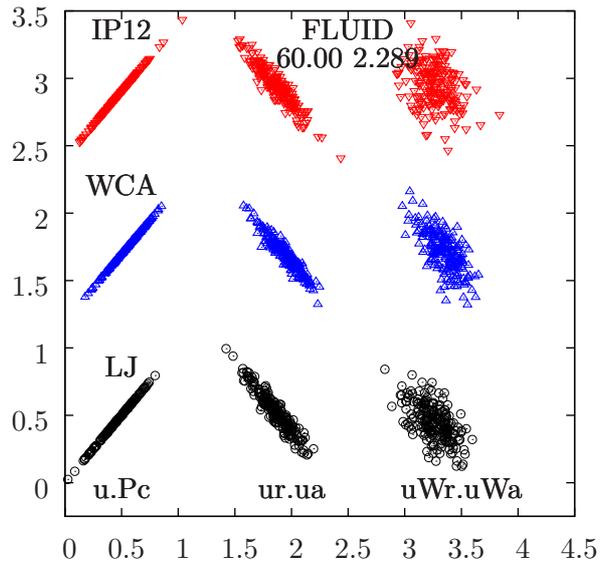


FIG. 7: As for Fig. 5, except the state point $T = 60.0$ and $\rho = 2.289$ is considered.

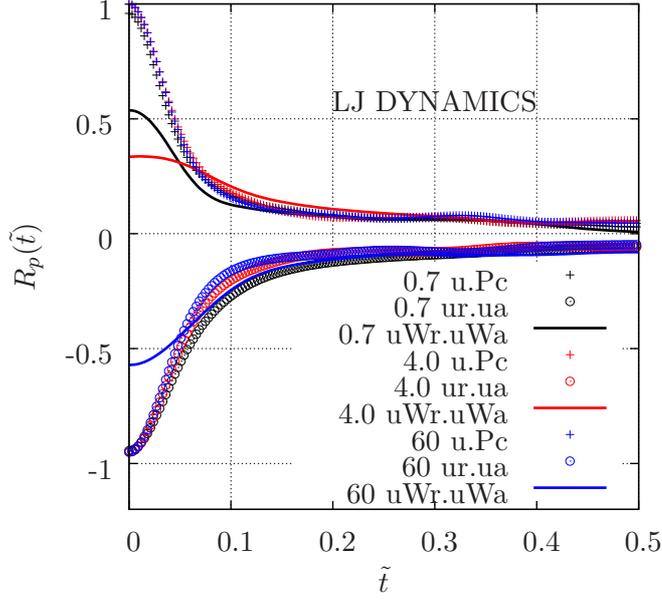


FIG. 8: The function, $R_p(\tilde{t})$ defined in Eq. (13) is plotted for three cross-correlations, $\langle u(0)P_c(\tilde{t}) \rangle$ and $\langle u_r(0)u_a(\tilde{t}) \rangle$ using the LJ potential terms, and $\langle u_{r,WCA}(0)u_{a,WCA}(\tilde{t}) \rangle$. The annotation for the quantities is the same as for Figs. 5-7. The LJ potential was used in each case to calculate the forces used in the equations of motion. For each quantity the difference from the mean is used. Data for the fluid phase state points, [60.00, 2.289] (blue symbols and curve online), [4.00, 1.229] (red online) and [0.70, 0.847] (black online), are considered, in the same order from bottom to top on the figure in the WCA case.

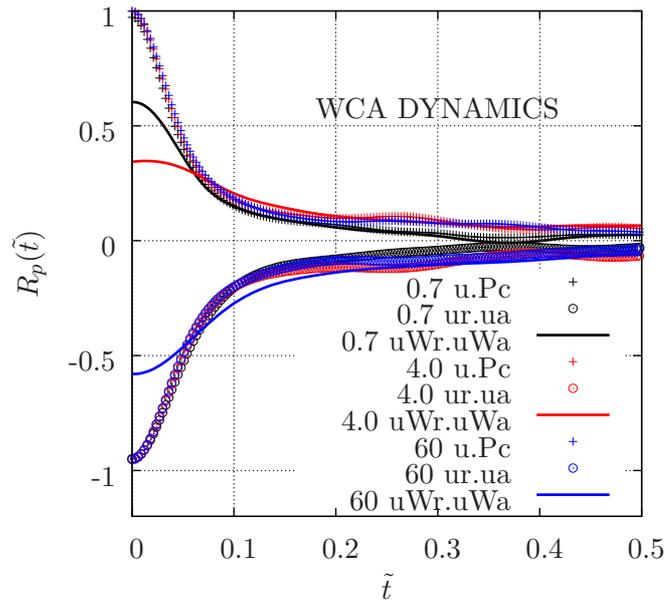


FIG. 9: As for Fig. 8 except that WCA potential was used in each case to generate the dynamics.

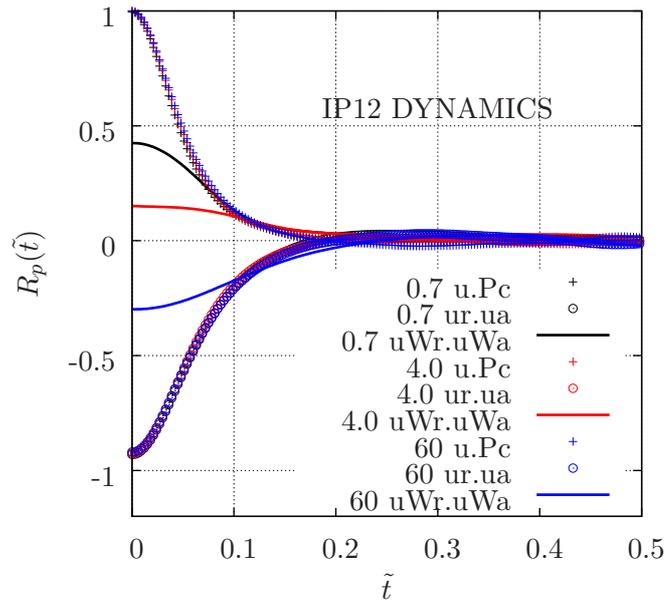


FIG. 10: As for Fig. 8 except that IP12 potential was used for the dynamics.