Theor Chem Acc (1997) 96:61–70

Theoretical
Chemistry Accounts

Computationally useful bridge diagram series for the structure and thermodynamics of Lennard-Jones fluids

John Perkyns, B. Montgomery Pettitt

Chemistry Department, University of Houston, Houston, TX 77204-5641, USA

Received: 9 January 1997 / Accepted 28 January 1997

Abstract. The first two orders of bridge diagrams, those with two and three field points, have been calculated exactly for the Lennard-Jones fluid for several isotherms. The method of calculation was one of expansion in Legendre polynomials, and the dependence of the method on the number of polynomials needed for accurate results was investigated. Thermodynamic and structural properties of the Lennard-Jones fluid calculated from integral equation methods with the inclusion of bridge diagrams were found to be systematically improved. Two attempts at predicting the missing bridge diagrams of even higher order were discussed. The first, which uses the functional form of those diagrams that were calculated exactly, showed no significant improvement. The second, a series sum based on the first two orders of calculated diagrams and motivated by the success of a similar heuristic sum in the case of hard spheres, was extremely successful. When the series sum was employed, thermodynamic and structural quantities were improved to the point where the difference between simulation results and integral equation results was of the same order as the error in the simulations themselves.

Key words: Fluids – Integral equation methods – Radial distribution function

1 Introduction

The study of statistical mechanics of liquids and liquid mixtures contributes increasingly to the understanding of many chemical systems of current interest, particularly those of a biological nature [1–6]. While simulation methods have commanded the attention of most investigators in recent years, there has remained a niche for approximate integral equation methods because of the relatively low computational resources required.

Correspondence to: J. Perkyns

The reasons for the dominance of simulation over more theoretical approaches are several. Historically, simulation results have been more trusted since a careful simulation of a simple fluid model involving no long-range forces gives results essentially exact for the model [7]. The approximations involved in integral equation methods lead, in a small but sufficient number of cases, to poor results [7]. In addition, there is no truly satisfactory theory for molecular models, since those which approximate simulation results most reliably have celebrated formal deficiencies [8, 9]. It is also assumed, because of the problems with theoretical methods, that the superiority of simulation results will hold for very complicated systems where the simulations are not exact, although this remains largely unproven. Despite the problems with theories, there are many areas where they have given useful contributions. In addition, because they are based on first principles, it is clear that they would have wider use and acceptance if the results were exact, or at least if the approximations were less severe.

The problem of calculating highly accurate radial distribution functions from integral equation methods is one of computational tractability; the exact formal mathematical expression is known [7, 10]. An exact radial distribution function for a fluid can be written as a virial expansion in function-valued integrals, or diagrams. Integral equation theories divide these diagrams into subsets and express relations between the subsets in terms of the pair functions for the fluid. There is, however, one subset of diagrams known as bridge diagrams for which a simple expression in terms of pair functions does not seem possible. Usually these bridge diagrams are either ignored completely or approximated by other expressions. The additional computational effort required to exactly calculate the single bridge diagram of lowest order in density is at least double that of the problem without such improvement. As with all virial expansions, there are an infinite number of these diagrams, and there appears to be no recursive or iterative method for obtaining them all. There has, however, been no shortage of attempts to use approaches which either directly or indirectly approximate them [11–18]. Because of the difficulty in finding a method of systematic im-

provement to integral equation theories, development efforts have focused on adding expressions which correct at least one known problem, such as thermodynamic inconsistency [12, 14, 16–18] or dielectric inconsistency [19, 20]. Unfortunately, such approaches are either ad hoc or are corrections which cannot remedy all the problems of the method to which they are applied. Also, they are not generally transferable to other systems, while they may work well in one specific group of cases.

In the search for a general theoretical method of calculating accurate radial distribution functions there appears to be no substitute for directly evaluating at least some bridge diagrams. When one considers that computers have increased several orders of magnitude in speed since currently used equation systems became established, arguments of computational tractability have lost, and continue to lose, impact. It is also becoming much more common to improve the accuracy of long-range interactions in already expensive simulations with even more expensive approaches such as Ewald sums [21–25]. Furthermore, the size, complexity and simulation time of systems under consideration continues to increase [24]. It should be noted that the computation time used by a simulation grows as up to the square of the number of molecules being simulated, whereas the calculation of integral equation methods grows as the square of the number of species present in the system. These factors should ensure that the difference between computational resources needed for simulations and theories should become wider for future scientific investigations, even if some bridge diagrams are calculated.

Yet if the calculation of bridge diagrams is justifiable on a cost basis, it is still only worthwhile if the results are truly improved. The purpose of this study is to show the improvement in structure and thermodynamics of a system of pure Lennard-Jones spheres obtained by including all the four- and five-point diagrams in the bridge function. The formalism for the calculation of these diagrams in terms of expansions in Legendre polynomials was derived by Attard and Patey [10] and tested for a convenient, but less chemically useful model [26]. We have made a few small changes to the formalism (for computational convenience only) and so reiterate it here in Sect. 2, as well as restate in summary the integral equation system we wish to solve. Since this paper represents the first application of the formalism to a soft sphere model that can accurately represent a chemical system, part of what follows (Sect. 3) will show that indeed such diagrams can be calculated by this method. In Sect. 4 we then consider the accuracy of the results by examining thermodynamic and structural quantities and comparing them to simulation. We also examine two methods of approximating the infinite series of diagrams which are still missing from an exact solution, based on the diagrams we have explicitly calculated. Section 5 contains a summary and the main conclusions.

2 Theory

2.1 The pair functions

The expression of pair functions in terms of diagrammatic expansions has appeared before in many contexts and will be summarized here using standard terminology [7]. All pertinent pair functions can be expressed as infinite sums of diagrams consisting of two root points, density field points and bonds. No diagram contains an articulation point. The bonds represent Mayer f functions, $e^{-\beta u(r)} - 1$, where $u(r)$ is the pair potential and $\beta = 1/k_B T$, with k_B, T and r being the Boltzmann constant, the absolute temperature and the interparticle distance, respectively. The sum of all such connected diagrams is an expression of the total correlation function $h(r) = g(r) - 1$, where $g(r)$ is the radial distribution function. A subset of the diagrams of $h(r)$, formed by removing all those diagrams with nodal points, is an expression of $c(r)$, the direct correlation function. A different subset of the diagrams of $h(r)$, which expresses $w(r)$, is formed by removing all diagrams with no field points or where the root points form an articulation pair. The function $-w(r)$ is the excess potential of mean force. These subsets of the total correlation diagrams all contain an infinite number of diagrams, and all converge to bounded functions for physically meaningful systems. The functions which represent them are related by

$$w(r) = \beta u(r) + \ln g(r), \tag{1a}$$

and

$$h(r) = c(r) + \rho c(r) * h(r), \tag{1b}$$

where the $*$ represents a convolution and ρ the number density. Equation (1b) is known as the Ornstein-Zernike equation [27]. Equations (1a) and (1b) are not sufficient to solve the system of equations since we have two equations in three unknown functions. Another expression, known as a closure, is also necessary. An exact closure can be written

$$w(r) = t(r) + \mathrm{d}(r), \tag{2}$$

which represents the separation of the diagrams of $w(r)$ into those with nodes, $t(r) = h(r) - c(r)$, and those without nodes, $\mathrm{d}(r)$, the bridge diagrams. The bridge diagrams are highly connected and apparently cannot be evaluated simply. Fortunately there are relatively few of them, so setting $\mathrm{d}(r) = 0$ should not be expected to be a drastic approximation. The result is the hypernetted-chain (HNC) equation which is easily obtained from Eqs. (2) and (1a) and can be written [7]

$$c(r) = h(r) - \beta u(r) - \ln g(r). \tag{3}$$

A well-known closure can be derived by taking the bridge function

$$\mathrm{d}(r) = \ln(1 + t(r)) - t(r), \tag{4}$$

which leads to the closure known as the Percus-Yevick (PY) equation [28],

$$c(r) = (1 + t(r))\mathrm{e}^{-\beta u(r)} - t(r) - 1. \tag{5}$$

Once the $g(r)$ function for a given closure has been calculated, the excess internal energy per particle can easily be evaluated via

$$\frac{\beta U^{\mathrm{ex}}}{N\epsilon} = 2\pi\rho \int_0^\infty r^2 g(r)\beta u(r)\mathrm{d}r, \tag{6}$$

and the isothermal compressibility factor by

$$\frac{\beta P}{\rho} = 1 - \frac{2}{3}\pi\beta\rho \int_0^\infty r^3 g(r)\frac{\mathrm{d}u(r)}{\mathrm{d}r}\mathrm{d}r, \tag{7}$$

where P indicates system pressure.

The pair potential used for this work has the standard Lennard-Jones form,

$$u(r) = 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] \tag{8}$$

where σ and ϵ are the distance and well-depth parameters. Any phase point in a system of pure Lennard-Jones (LJ) spheres can be completely designated by the reduced temperature and density variables $T^* = 1/\beta\epsilon$ and $\rho^* = \rho\sigma^3$ [7].

2.2 The evaluation of bridge diagrams

What follows is a summary of the derivation of several expressions given fully in reference [10]. Expressions for the first two orders of bridge diagrams are rewritten in computationally tractable form. The simplification employs the method of expansion of angular dependent integrals in Legendre polynomials used quite widely in different contexts [29–31].

The function $\mathrm{d}(r)$, like all the pair functions, can be expressed as a virial series:

$$\mathrm{d}(r) = \rho^2\mathrm{d}^{(2)}(r) + \rho^3\mathrm{d}^{(3)}(r) + \cdots. \tag{9}$$

Each term in the series represents the sum of diagrams with a fixed number of field points, indicated by the superscripts. No bridge diagrams exist for lower orders of density, $\mathrm{d}^{(2)}(r)$ represents a single diagram, and $\mathrm{d}^{(3)}(r)$ represents the sum of 13 diagrams (Fig. 1).

The only diagram of $\mathrm{d}^{(2)}(r)$ is

$$\begin{aligned}\mathrm{d}^{(2)}(r) &= \frac{1}{2}\int \mathrm{d}r_3\mathrm{d}r_4 f(r_3)f(r_4)f(r_{23})f(r_{24})f(r_{34}) \\ &= \pi\int_0^\infty \mathrm{d}r_3 r_3^2 \int_0^\infty \mathrm{d}r_4 r_4^2 \int_0^\pi \mathrm{d}\theta_3 \sin\theta_3 \int_0^\pi \mathrm{d}\theta_4 \sin\theta_4 \\ &\quad\times \int_0^{2\pi} \mathrm{d}\phi_{34} f(r_3)f(r_4)f(r_{23})f(r_{24})f(r_{34}),\end{aligned} \tag{10}$$

where r_1 is chosen as the origin and $\boldsymbol{r}_2$ is aligned with the positive z axis, $\theta_3 = \theta_{23}$, $\theta_{24} = \theta_4$, $r_{ij} = \sqrt{r_i^2 + r_j^2 - 2r_i r_j\cos\theta_{ij}}$ and $\cos\theta_{34} = \cos\theta_3\cos\theta_4 + \sin\theta_3\sin\theta_4\cos\phi_{34}$. The details of the simplification of Eq. (10) are given elsewhere [10], but the general procedure is to expand each angular dependent function in the integrand in a series of Legendre polynomials and apply well-known mathematical relationships to simplify to an expression which can be written in terms of the coefficients of the expansions only. The coefficients can be expressed as

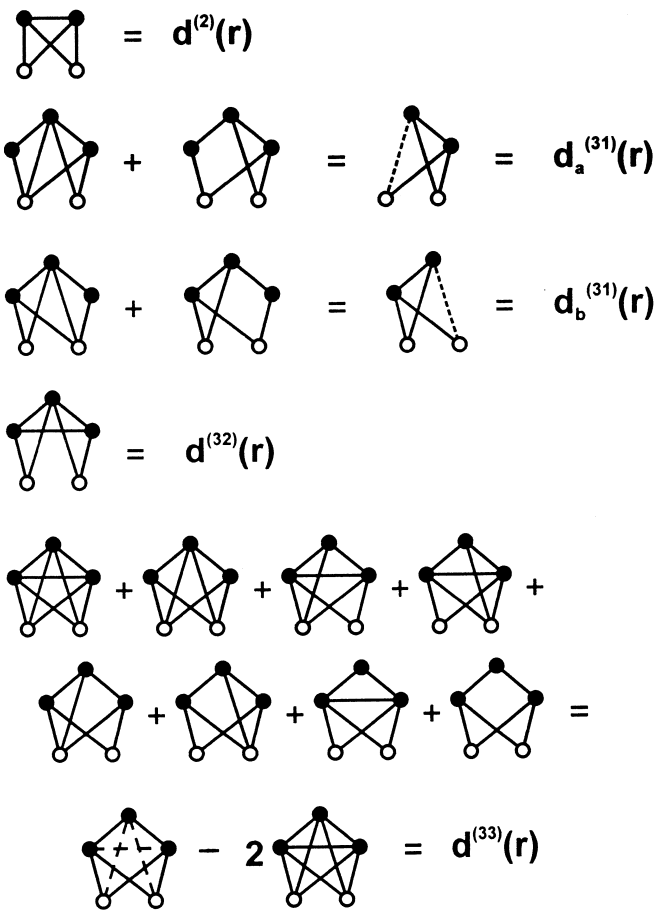

Fig. 1. The bridge diagrams of orders ρ^2 and ρ^3 and the functions to which they add. The *open circles* represent root points and the *solid circles* field points to be integrated. The *solid lines* represent Mayer f bonds, the *dashed lines* represent $h^{(1)}$ bonds and the *long-dashed lines* represent e bonds

$$\hat{f}_n(r_i, r_j) = \frac{2n+1}{2}\int_{-1}^{1} \mathrm{d}x P_n(x) f\left(\sqrt{r_i^2 + r_j^2 - 2r_i r_j x}\right), \tag{11}$$

where $x = \cos\theta$, and the final expression desired is

$$\begin{aligned}\mathrm{d}^{(2)}(r) &= 2\pi^2 \sum_{n=0}^{\infty}\left(\frac{2}{2n+1}\right)^2 \int_0^\infty \mathrm{d}r_3 r_3^2 f(r_3) \\ &\quad\times \int_0^\infty \mathrm{d}r_4 r_4^2 f(r_4)\hat{f}_n(r_2, r_3)\hat{f}_n(r_2, r_4) \\ &\quad\times \hat{f}_n(r_3, r_4).\end{aligned} \tag{12}$$

The 13 diagrams of $\mathrm{d}^{(3)}(r)$ are most conveniently rearranged into three groups, each of which is evaluated separately. The first group, consisting of four diagrams and designated $\mathrm{d}^{(31)}(r)$, is reduced to the sum of two contributions. Each contribution is formally the same as $\mathrm{d}^{(2)}(r)$ if one employs $h^{(1)}(r) = e(r)(f(r) * f(r))$ bonds to simplify the topology of the diagrams. We will write the two contributions separately here with notation which keeps track of which field points in the original five-point diagrams correspond to the field points in the

completely connected four-point diagram. For mixtures with more than one component at finite concentration the number of bridge diagrams with the same topology but different valued f bonds increases dramatically. Therefore, to avoid unnecessary complication in the counting of such diagrams it will be necessary to have all component functions of $\mathrm{d}^{(3)}(r)$ agree on which species is referred to by each of the three field points in a five-point diagram. This is unnecessary if the field points can only represent one species, such as in this work or in Ref. [10], but we forgo the simplification here for completeness and with a view to subsequent work. We write the two contributions to $\mathrm{d}^{(31)}(r)$ as

$$\mathrm{d}_a^{(31)}(r) = 4\pi^2 \sum_{n=0}^{\infty} \left(\frac{2}{2n+1}\right)^2 \int_0^{\infty} \mathrm{d}r_4 r_4^2 f(r_4) \times \int_0^{\infty} \mathrm{d}r_5 r_5^2 h^{(1)}(r_5) \hat{f}_n(r_2, r_4) \hat{f}_n(r_2, r_5) \hat{f}_n(r_4, r_5) \tag{13}$$

and

$$\mathrm{d}_b^{(31)}(r) = 4\pi^2 \sum_{n=0}^{\infty} \left(\frac{2}{2n+1}\right)^2 \int_0^{\infty} \mathrm{d}r_3 r_3^2 f(r_3) \times \int_0^{\infty} \mathrm{d}r_5 r_5^2 f(r_5) \hat{f}_n(r_2, r_3) \hat{f}_n(r_3, r_5) \hat{h}_n^{(1)}(r_2, r_5), \tag{14}$$

where

$$h^{(1)}(r_5) = e(r_5)(f(r_{13}) * f(r_{35})) \tag{15a}$$

and $\hat{h}_n^{(1)}(r_2, r_5)$ represents the polynomial coefficients (Eq. 11) in the expansion of

$$h^{(1)}(r_{25}) = e(r_{25})(f(r_{24}) * f(r_{45})). \tag{15b}$$

The second contribution to $\mathrm{d}^{(3)}(r)$ is the single diagram indicated in Fig. 1,

$$\mathrm{d}^{(32)}(r) = \int \mathrm{d}r_3 \mathrm{d}r_4 \mathrm{d}r_5 f(r_3) f(r_5) f(r_2, r_4, \theta_4) f(r_2, r_5, \theta_5) \times f(r_3, r_4, \theta_{34}) f(r_3, r_5, \theta_{35}) f(r_4, r_5, \theta_{45}), \tag{16}$$

which, when simplified in a similar manner in terms of the coefficients of Legendre polynomials, is

$$\mathrm{d}^{(32)}(r) = (2\pi)^3 \sum_{l,m=0}^{\infty} \frac{2}{2l+1} \frac{2}{2m+1} \int_0^{\infty} \mathrm{d}r_3 \mathrm{d}r_4 \mathrm{d}r_5 r_3^2 r_4^2 r_5^2 \times f(r_3) f(r_5) \hat{f}_m(r_2, r_4) \hat{f}_m(r_2, r_5) \hat{f}_l(r_3, r_4) \times \hat{f}_l(r_3, r_5) \int_{-1}^{1} \mathrm{d}x P_l(x) P_m(x) f(r_4, r_5, x). \tag{17}$$

The final contribution to $\mathrm{d}^{(3)}(r)$ consists of the sum of the eight remaining bridge diagrams with three field points and can be written

$$\mathrm{d}^{(33)}(r) = \int \mathrm{d}r_3 \mathrm{d}r_4 \mathrm{d}r_5 f(r_3) f(r_4) f(r_2, r_3, \theta_3) f(r_2, r_4, \theta_4) \times f(r_3, r_5, \theta_{35}) f(r_4, r_5, \theta_{45}) \left[\frac{1}{2} e(r_5) e(r_2, r_5, \theta_5) \times e(r_3, r_4, \theta_{34}) - \frac{1}{3} f(r_5) f(r_2, r_5, \theta_5) f(r_3, r_4, \theta_{34})\right]. \tag{18}$$

As shown in Fig. 1, the eight diagrams can be written as a sum of two completely connected diagrams in terms of e and f bonds. After significant manipulation, Eq. (18) can be shown to be equal to the expansion

$$\mathrm{d}^{(33)}(r) = 16\pi^3 \sum_{lmn=0}^{\infty} \sum_{l'=0}^{\min(lmn)'} \int_0^{\infty} \mathrm{d}r_3 \mathrm{d}r_4 \mathrm{d}r_5 r_3^2 r_4^2 r_5^2 f(r_3) \times f(r_4) \hat{f}_l(r_3, r_5) \hat{f}_m(r_4, r_5) F_{ln}^{l'}(r_2, r_3) F_{mn}^{l'}(r_2, r_4) \times \left[\frac{1}{2} e(r_5) \hat{e}_n(r_3, r_4) E_{lm}^{l'}(r_2, r_5) - \frac{1}{3} f(r_5) \hat{f}_n(r_3, r_4) F_{lm}^{l'}(r_2, r_5)\right], \tag{19}$$

where

$$F_{ln}^{l'}(r_2, r_3) = W_{ll'} W_{nl'} \int_{-1}^{1} \mathrm{d}x_3 f(r_2, r_3, x_3) P_l^{l'}(x_3) P_n^{l'}(x_3), \tag{20}$$

$$W_{ln} = \sqrt{\frac{(l-n)!}{(l+n)!}}, \tag{21}$$

$$\hat{e}_n(r_3, r_4) = \hat{f}_n(r_3, r_4) + \delta_{n,0}, \tag{22}$$

$$E_{lm}^{l'}(r_2, r_5) = F_{lm}^{l'}(r_2, r_5) + \frac{2}{2l+1} \delta_{l,m}, \tag{23}$$

the $P_n^l(x)$ are associated Legendre functions, and the prime on the second summation in Eq. (19) means that the $l' = 0$ term should be multiplied by 1/2.

The final function we require for the exact sum of the bridge diagrams of third order in density is

$$\mathrm{d}^{(3)}(r) = \mathrm{d}_a^{(31)}(r) + \mathrm{d}_b^{(31)}(r) + \mathrm{d}^{(32)}(r) + \mathrm{d}^{(33)}(r). \tag{24}$$

3 The convergence of Legendre expansions

In the calculation of Eqs. (13), (14), (17) and (19) both the evaluation of Legendre expansion coefficients and the integration over radial variables were evaluated using gaussian quadrature [32, 29]. If the trapezoidal or a higher-order rule needing a similar number of evaluation points had been used for integrating the radial variables, sufficient accuracy would have required significantly more computation time. Since the two types of integration necessary are quite distinct, the polynomial order ($\mathcal{N}_1$) used to evaluate the expansion coefficients $\hat{f}_n(r_i, r_j)$ and $F_{ln}^{l'}(r_2, r_3)$ is varied independently of the polynomial order ($\mathcal{N}_2$) used to evaluate the integrals over the radial variables r_3, r_4 and r_5. The weights and zeros needed for the quadratures were calculated and tested using standard methods [33].

The expressions given above and the programming have been checked against the analytical results for the gaussian model given in Ref. [26], and agreement of eight figures was easily obtained when the analytical form of $h^{(1)}(r)$ was used in the calculations. For a Lennard-Jones potential, however, a numerical convolution

is required in the calculation of $h^{(1)}(r)$ and sufficient accuracy is more difficult to achieve. In the numerical evaluation of the $h^{(1)}(r)$ functions, the forward transform of the f bonds was done using a fast Fourier transform (FFT) [33] with 2^{17} points spaced evenly over a range of 8.5 σ. Since only relatively few evaluated function values are required for $h^{(1)}(r)$, and since it is not possible to have all the required gaussian quadrature points exactly aligned with any FFT grid spacing chosen, the back transform was calculated using a standard Fourier integral for each function value required. This method provided four to five digits of agreement with the gaussian model. The $h^{(1)}(r)$ values can be precalculated before the main integration loops.

It remains to be shown that $\mathrm{d}^{(2)}(r)$ and $\mathrm{d}^{(3)}(r)$ can be calculated with values of $\mathcal{N}_1$ and $\mathcal{N}_2$ sufficiently small that numerical calculation is practical. For each of the values of the reduced temperature $T^* = 2.74$ and $T^* = 1.35$, the functions $\mathrm{d}^{(2)}(r)$ and $\mathrm{d}^{(3)}(r)$ were evaluated for all possible combinations of $\mathcal{N}_1 = 8, 16, 32, 64$ and $\mathcal{N}_2 = 8, 16, 32, 64, 128$, except for the case of $\mathrm{d}^{(3)}(r)$ with $\mathcal{N}_1 = 64$ and $\mathcal{N}_2 = 128$, for which it was estimated that calculation would require >36 h per point on a 175-MHz Digital Alpha workstation. Since the convergence pattern for $\mathrm{d}^{(2)}(r)$ and $\mathrm{d}^{(3)}(r)$ was so similar for all other orders, only the calculation of a few sample points was attempted. As expected, the convergence of the results for both $\mathrm{d}^{(2)}(r)$ and $\mathrm{d}^{(3)}(r)$ with respect to the polynomial order $\mathcal{N}_1$ is quite independent of the convergence with respect to the polynomial order $\mathcal{N}_2$. An example of each type of convergence is given in Fig. 2a, b. Figure 2a shows the $\mathrm{d}^{(2)}(r)$ results for $\mathcal{N}_1 = 32$ and all considered values of $\mathcal{N}_2$. While it is clear that $\mathcal{N}_2 = 8$ and 16 are unacceptable, the results for $\mathcal{N}_2 = 64$ and 128 are virtually indistinguishable. They agree to at least two significant figures everywhere. Figure 2b shows the result for $\mathcal{N}_1 = 8, 16, 32$ and 64 while $\mathcal{N}_2$ is held fixed at 128. Again, at least two-digit agreement everywhere is found for the two highest orders of $\mathcal{N}_1$. Figure 2c gives the results for $\mathrm{d}^{(3)}(r)$ corresponding to those of $\mathrm{d}^{(2)}(r)$ in Fig. 2a, and the convergence pattern is almost identical. Furthermore, the convergence pattern for each type of integration was unaffected by doubling the value of the Lennard-Jones well depth (not shown). The increase in computation time goes as $\mathcal{N}_1^5$ for $\mathrm{d}^{(3)}(r)$, a function which is an order of density less significant than $\mathrm{d}^{(2)}(r)$. Since the difference in results for $\mathcal{N}_1 = 16$ and $\mathcal{N}_1 = 64$ (in combination with any value of $\mathcal{N}_2$) is $<3\%$ everywhere that the corresponding $g(r)$ is not identically zero, the benefit of spending the extra time on a slightly more accurate $\mathrm{d}^{(3)}(r)$ function will probably not be worth the small change in any system functions or thermodynamic values obtained. Converged system solutions can therefore be found using $\mathcal{N}_1 = 32$ for $\mathrm{d}^{(2)}(r)$, $\mathcal{N}_1 = 16$ for $\mathrm{d}^{(3)}(r)$, and $\mathcal{N}_2 = 64$ for all calculations using potential functions similar to those examined here.

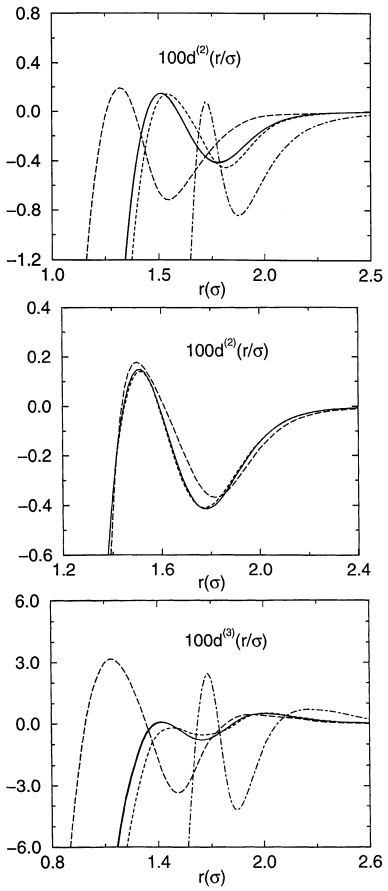

Fig. 2 **a** The function $\mathrm{d}^{(2)}(r)$ for $T^* = 2.74$ using $\mathcal{N}_1 = 32$. The *solid line* represents $\mathcal{N}_2 = 128$, the *dotted line* $\mathcal{N}_2 = 64$, the *dashed line* $\mathcal{N}_2 = 32$, the *long-dashed line* $\mathcal{N}_2 = 16$ and the *dot-dashed line* $\mathcal{N}_2 = 8$. **b** The function $\mathrm{d}^{(2)}(r)$ for $T^* = 2.74$ using $\mathcal{N}_2 = 128$. The *solid line* represents $\mathcal{N}_1 = 64$, *the dotted line* (which is virtually superimposed on the solid line) $\mathcal{N}_1 = 32$, the *dashed line* $\mathcal{N}_1 = 16$ and the *long-dashed line* $\mathcal{N}_2 = 8$. **c** The function $\mathrm{d}^{(3)}(r)$ for $T^* = 2.74$ using $\mathcal{N}_1 = 32$. The line types represent the same values of $\mathcal{N}_2$ as in **a**

4 Results and discussion

The results presented below are all calculated for a system of pure LJ spheres. Integral equation results are calculated by standard iterative methods [7]. All functions were calculated with 2048 evenly spaced points at a grid spacing of $\sigma/60$. Sample calculations using a halved grid spacing and doubled range were found to leave the thermodynamic quantities and radial distribution functions unchanged to four and six digits, respectively. The functions $\mathrm{d}^{(2)}(r)$ and $\mathrm{d}^{(3)}(r)$ for $T^* = 2.74$ and $T^* = 1.35$ were evaluated at each grid point using the converged values of $\mathcal{N}_1$ and $\mathcal{N}_2$ to a range of 4.5 σ. Extension of these functions to longer range left the thermodynamic quantities that depend on them unchanged to six digits. Following Ref. [10] we refer to the calculations performed using the bridge function

$$\mathrm{d}(r) = \rho^2 \mathrm{d}^{(2)}(r) \tag{25}$$

with Eqs.(1a, b) and (2) as the HNCD2 result, and that calculated using

$$\mathrm{d}(r) = \rho^2 \mathrm{d}^{(2)}(r) + \rho^3 \mathrm{d}^{(3)}(r), \tag{26}$$

together with the same equations, as the HNCD3 result. In order to compare radial distribution functions to

simulation, two molecular dynamics (MD) trajectories for the phase points ($T^* = 2.74, \rho^* = 1.0$) and ($T^* = 1.35, \rho^* = 0.7$) were calculated in the NVT ensemble using standard methods [34]. For both simulations a square box of length 10 σ was used. The calculations spanned 25,000 time steps for a total of 500 ps with a radial cutoff, r_c, at slightly less than half the box length, 4.2 σ. The first-phase point calculation used 1000 LJ spheres and resulted in values of excess internal energy, $U^{ex}/N\epsilon$, and isothermal compressibility factor, $\beta T/\rho$, of -4.219 and 7.307, respectively. The second-phase point calculation used 700 LJ spheres and resulted in the values $U^{ex}/N\epsilon = -4.665$ and $\beta T/\rho = 1.205$.

4.1 Thermodynamic quantities

Tables 1 and 2 contain the thermodynamic results of integral equation calculations for the HNC, HNCD2, HNCD3 and PY theories, as well as the Monte Carlo (MC) simulation results reproduced in Ref. [7] for comparison. The results presented for the HNC and PY closures differ slightly from those reported in Ref. [7]. The latter can be duplicated almost exactly by using a coarser grid and shorter range of 512 points with a grid spacing of $\sigma/50$. The results from MD and MC calculations also differ well outside the range of statistical error reported for both calculations. Besides the methods themselves, there are two differences between the simulations. The first is the method of correcting the tail beyond the point where the potential interaction is cut off. The MC calculations assume $g(r) = 1$ and apply the integrals in Eqs. (6) and (7) from the cutoff radius outwards. The MD result calculated here applied the same equations but uses the $g(r)$ function from an HNC calculation beyond the cutoff radius. This accounts for about 30% of the difference between MC and MD results. The remainder must be due to the difference in cutoff radius for the calculations. The MC result used $r_c = 2.4$ σ, whereas the MD uses $r_c = 4.2$ σ. While these differences in thermodynamic quantities are not large, many of the integral equation results fall within this wider bound of error. Using the statistical error alone to evaluate the precision of simulation results can thus be misleading when judging the veracity of theoretical results.

Table 1. Excess internal energies per particle, $U^{ex}/N\epsilon$ from MC simulation [7] and various integral equation theories

$T^* = 1.35$						
ρ^*	MC	HNC	HNCD2	HNCD3	HNCS	PY
0.70	−4.684	−4.524	−4.548	−4.568	−4.656	−4.625
0.65	−4.343	−4.259	−4.276	−4.290	−4.341	−4.322
0.55	−3.704	−3.678	−3.683	−3.686	−3.691	−3.695
0.50	−3.372	−3.378	−3.377	−3.377	−3.371	−3.382
0.45	−3.030	−3.081	−3.075	−3.072	−3.061	−3.073
0.40	−2.747	−2.792	−2.781	−2.776	−2.763	−2.773
0.35	−2.405	–	−2.468	−2.488	−2.477	−2.480
0.30	−2.090	–	–	–	−2.167	−2.188
$T^* = 2.74$						
ρ^*	MC	HNC	HNCD2	HNCD3	HNCS	PY
1.00	−4.180	−3.259	−3.338	−3.475	−4.137	−4.582
0.80	−4.281	−3.847	−3.907	−3.980	−4.196	−4.306
0.70	−3.902	−3.691	−3.734	−3.779	−3.884	−3.931
0.55	−3.207	−3.127	−3.147	−3.163	−3.189	−3.200
0.40	−2.371	−2.353	−2.358	−2.361	−2.365	−2.368
0.30	−1.783	−1.787	−1.788	−1.789	−1.790	−1.791

Table 2. Isothermal compressibility factors, $\beta P/\rho$ from MC simulation [7] and various integral equation theories

$T^* = 1.35$						
ρ^*	MC	HNC	HNCD2	HNCD3	HNCS	PY
0.70	1.166	2.097	1.926	1.783	1.145	1.697
0.65	0.850	1.500	1.351	1.235	0.769	1.262
0.55	0.415	0.758	0.653	0.581	0.357	0.695
0.50	0.303	0.560	0.475	0.421	0.275	0.534
0.45	0.280	0.444	0.377	0.339	0.249	0.436
0.40	0.272	0.389	0.339	0.313	0.261	0.388
0.35	0.298	–	0.340	0.327	0.299	0.377
0.30	0.352	–	–	–	0.353	0.396
$T^* = 2.74$						
ρ^*	MC	HNC	HNCD2	HNCD3	HNCS	PY
1.00	7.388	9.147	8.976	8.699	7.419	6.698
0.80	3.604	4.540	4.392	4.227	3.746	3.611
0.70	2.641	3.167	3.046	2.935	2.675	2.650
0.55	1.653	1.903	1.830	1.780	1.698	1.724
0.40	1.199	1.280	1.247	1.232	1.215	1.235
0.30	1.040	1.084	1.069	1.064	1.060	1.070

The HNC values for excess internal energy (Table 1) are overestimates for higher densities and underestimates for lower densities, with the crossover point in the range $0.50 < \rho^* < 0.55$ for the $T^* = 1.35$ isotherm and in the range $0.30 < \rho^* < 0.40$ for the $T^* = 2.74$ isotherm. The HNCD2 and HNCD3 each represent successive improvements in the energy values, no matter whether the direction of improvement is to higher or lower values. Furthermore, while these improvements always move toward the simulation result, they do not pass it. The PY energy results are clearly superior to the HNC results at all phase points calculated except one. The PY results are also closer to the MC than all but three of the HNCD3 results, but it should be noted that the HNCD3, PY and MC results are most often within a range similar to that of the difference between the two types of simulation compared above.

Similar systematic improvements to the isothermal compressibility factor from the HNC theory can be achieved by employing the bridge diagram functions $d^{(2)}(r)$ and $d^{(3)}(r)$. For all points on the phase diagram investigated, the results for the HNC theory were all too large (Table 2) and all were successively reduced towards the MC values. For these quantities, however, in the case with the strongest interaction between LJ spheres, $T^* = 1.35$, the HNCD3 results surpass those of the PY theory for all but the highest density.

Another interesting consequence of the addition of bridge diagrams for systems with $T^* = 1.35$ and $\rho^* < 0.40$ is that solution to the corrected HNC equation is now possible, at a lower density, where previously the HNC equation results did not give a solution. The lack of solution is usually interpreted as the crossing of a phase boundary, in this case the well-known separation of the liquid and gas phases of the Lennard-Jones fluid [35]. The HNC theory appears to misplace the phase boundary to a position of higher density on the phase diagram. The HNCD2 and HNCD3 theories appear not only to improve the thermodynamic quantities of the HNC, but also to move the position predicted for the liquid-gas phase boundary toward that which is correct for the model.

The HNCS results, which employ a heuristic series based on $d^{(2)}(r)$ and $d^{(3)}(r)$ to approximate the missing infinite series of bridge diagrams of density order 4 and higher, are also shown in Tables 1 and 2 and will be discussed below.

4.2 Radial distributions

When the radial distribution functions (RDFs) resulting from the theories under consideration are plotted together with the simulation result, the individual curves are difficult to distinguish, so instead we plot the difference functions, $\Delta g(r)$, defined by

$$\Delta g(r) = g_{\text{theory}}(r) - g_{\text{simulation}}(r) \tag{27}$$

for each theory. We use the RDF from the MD simulation performed for this work. The closer a result for a given theory is to zero, the closer to exact for the model it can be regarded. Figure 3a gives the $\Delta g(r)$ functions for the phase point $(T^* = 2.74, \rho^* = 1.0)$. As with the thermodynamic quantities, the HNCD2 and HNCD3 functions represent successive improvements to the HNC function. The RDF from the PY has the opposite deviation from the simulation to the HNC everywhere, which demonstrates a case where attempts to improve theories by taking a weighted average of HNC and PY would be reasonably successful. By contrast, such an approach would not be expected to work well for the phase point $(T^* = 1.35, \rho^* = 0.7)$ (Fig. 3b), where all deviations are in the same direction. Even if a negative coefficient for one of the PY and HNC functions were used in such an attempt, the result is unlikely to be satisfactory because the two functions are significantly out of phase with each other.

The statistical 'noise' from the simulation has not been subjected to any smoothing procedure and is evident in these results, particularly those of Fig. 3b, where the deviations are roughly half those of Fig. 3a. The statistical noise was not removed in order to demonstrate that the deviations from exactness from theories are less than an order of magnitude greater than the errors in even well-converged simulations, except at the highest densities.

The HNCS radial distributions are discussed in the next section.

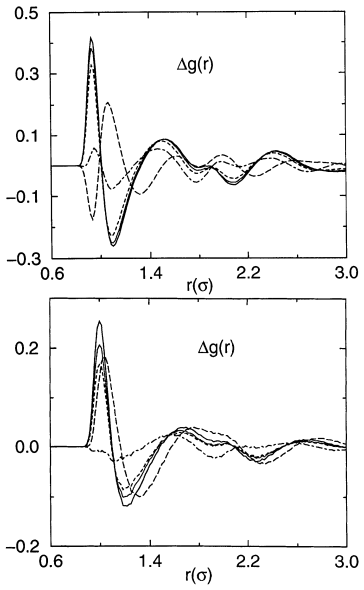

Fig. 3 a The function $\Delta g(r)$, the difference between the result from simulation and each theory. All calculations are for $T^* = 2.74$ and $\rho^* = 1.0$. The *solid line* represents the HNC result, the *dotted line* represents the HNCD2, the *dashed line* is the HNCD3 result and the *long-dashed* and *dot-dashed lines* come from the PY and HNCS calculations, respectively. **b** The function $\Delta g(r)$ for $T^* = 2.74$ and $\rho^* = 1.0$. The line types represent the same types of calculations as in **a**

4.3 Estimating the missing diagrams

While the HNCD2 and HNCD3 approximations to the missing bridge diagrams improve the HNC results everywhere, it has been found for hard spheres [10] that even greater improvement can be achieved by using the bridge functions $d^{(2)}(r)$ and $d^{(3)}(r)$ to estimate missing terms in the $d(r)$ series (Eq. 9) of higher order in density.

The simplest estimate can be created by assuming that those bridge functions calculated provide the correct functional form, and that we can improve the result by simply adjusting the magnitude using a coefficient, C_{ff}. The coefficient must depend on ρ since at the limit of small ρ it must approach 1, and it cannot have that value everywhere. If the coefficient is to be useful it must have little or no dependence on T^*, since such a dependence would not be simple if the estimate were applied to mixtures. We therefore assume that C_{ff} depends on ρ only and write

$$d(r) = C_{ff}(\rho)\left[\rho^2 d^{(2)}(r) + \rho^3 d^{(3)}(r)\right]. \tag{28}$$

Using this functional form, it should be possible to find values of C_{ff} which, when used to solve the corresponding integral equation system, duplicate a given thermodynamic quantity from a simulation. For each excess internal energy and isothermal compressibility factor value from MC simulations (Tables 1, 2) as well as those from the MD simulations of this work, a value of C_{ff} was calculated which duplicated a single simulation thermodynamic property to three digits, if such a procedure was possible. There were a few cases where it was not possible to duplicate a given simulation value,

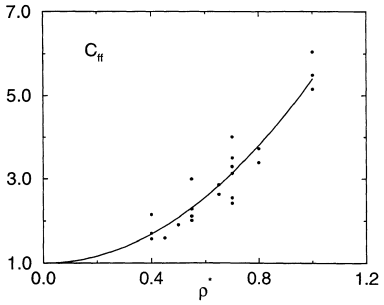

Fig. 4. The values of the functional form coefficient necessary to duplicate the thermodynamic quantities of simulation. The *line* is a regression fit to the data discussed in the text

although it was always possible to get within the bounds of error suspected for the simulation results (see above). Each such value of C_{ff} was given equal weight and plotted against density in Fig. 4. A linear regression of $(C_{ff}-1)/\rho^*$ was then performed to evaluate the constants in the equation

$$C_{ff}(\rho^*) = 1 - 0.086\rho^* + 4.48(\rho^*)^2, \tag{29}$$

which is also plotted in Fig. 4. While there appears to be no strong dependence on T^* and the fitted function for C_{ff} seems quite promising, integral equation results using the fitted C_{ff} values were no improvement over the HNCD3 results. If a given energy was improved, then the compressibility factor for that system was made worse, or vice versa. The most troubling feature of this attempt was, however, that even if the thermodynamics of a system were acceptable or improved, the corresponding RDFs inevitably deviated more strongly from the simulation results than did the HNCD3 result. This functional form approach also has no claim to exactness in all pair functions up to order ρ^3, unlike the HNCD3 closure.

While this functional form method was not a success, it emphasized an important condition in the estimation of approximate bridge functions. It appears to be extremely difficult to optimize a function for a given property of a system without losing accuracy in another property. Improvement in several system properties at once proves to be an extremely demanding criterion. If such an estimate improves the thermodynamics and structure of a system simultaneously, then it should probably be regarded as quite close to the true total bridge function, even if no direct evidence for such a claim is available.

In Ref. [10], for the case of hard sphere systems, the (1,2)-Padé approximant

$$\mathrm{d}(r) = \frac{\rho^2 \mathrm{d}^{(2)}(r)}{1 - \rho \mathrm{d}^{(3)}(r)/\mathrm{d}^{(2)}(r)} \tag{30}$$

proved to be just such an estimate, improving thermodynamics and structure alike. It is unfortunate that this approximation is impossible for Lennard-Jones systems. There are two reasons why this is true. The first is that there are several values of r for which $\mathrm{d}^{(2)}(r) = \rho d^{(3)}(r)$. At such points there is a zero in the denominator of the Padé approximant. The second is that for most of the range of r, $|\rho \mathrm{d}^{(3)}(r)/\mathrm{d}^{(2)}(r)| > 1$ and the Padé approximant, which is a type of functional geometric series, does not converge.

The success of the Padé approximant is quite compelling, however, so we will attempt to extract the essential features of it here in a way which can be used for LJ spheres, and perhaps other types of potential function as well. The problem of zeros in the denominator can be addressed by noticing that both $\mathrm{d}^{(2)}(r)$ and $\mathrm{d}^{(3)}(r)$, in the approximate range $0 < r < 1.5\ \sigma$, are remarkably gaussian in form, and also that this gaussian structure has an amplitude several orders of magnitude larger than the rest of the function. Furthermore, the relationship $\mathrm{d}^{(3)}(0)/\mathrm{d}^{(2)}(0) \approx \sqrt{\mathrm{d}^{(2)}(0)}$, general to all isotherms investigated, indicates that the geometric series property of the hard sphere bridge functions is also present at short range in the LJ sphere functions. We therefore assume that the missing orders of bridge functions, $\mathrm{d}^{(4)}(r), \mathrm{d}^{(5)}(r), \ldots$, will also be of gaussian form at short range, and that this short-range part is the most important feature of the bridge functions. Unfortunately, since the gaussian part of $\mathrm{d}^{(3)}(r)$ is not only of higher amplitude than that of $\mathrm{d}^{(2)}(r)$, but also of longer range, a geometric series of such functions will not sum to a finite result, and we must force convergence on the series. We therefore evaluate the gaussian $G^{(2)}(r) = A_2 e^{-a_2 r^2} \approx \mathrm{d}^{(2)}(r)$ using the calculated function values at the two points $r = 0$ and $r = \sigma$, and do the same for $G^{(3)}(r) = A_3 e^{-a_3 r^2} \approx \mathrm{d}^{(3)}(r)$. We then define the constants $R = A_3/A_2$ and $\delta = a_3 - a_2$ and define the function series sum

$$\mathrm{d}_s(r) = \rho^2 \mathrm{d}^{(2)}(r) + \rho^3 \mathrm{d}^{(3)}(r) + \rho^2 A_2 e^{-a_2 r^2} \sum_2^\infty \frac{(2R\rho)^n}{n!} e^{-n\delta r^2}. \tag{31}$$

Each gaussian in the series is narrower than the previous one by the constant δ. This ensures that $G^{(3)}(r)$ is the widest gaussian and that the amplitudes decrease rapidly as n becomes large. This series sum maintains the exactness to order ρ^3 of the HNCD3 closure by only adding terms of order ρ^4 or higher. The results of using this estimated bridge function, which we call gaussian series approximation (HNCS), are given in Tables 1 and 2. The HNCS thermodynamic values are superior to all others, including the PY values, everywhere except a few isolated cases where HNCS, PY and MC are all within the tail correction error in the simulation results. The HNCS $\Delta g(r)$ functions given in Fig. 3a and 3b are clearly closer to the zero than all others. In Fig. 3b the error in the HNCS result is only 2 or 3 times the error due to statistical noise from the simulation. In Fig. 5a the HNCD3 and HNCS bridge functions are compared for the two phase points $(T^* = 2.74, \rho^* = 1.0)$ and $(T^* = 1.35, \rho^* = 0.7)$. In both cases the addition of the series of gaussians has greatly decreased the first peak in $\mathrm{d}(r)$ while leaving the second almost unchanged. The RDFs which result from the HNCS closure are remarkably close to the simulation RDFs with all peaks at the correct height and almost exactly in phase (Fig. 5b). The only place where corresponding curves do not run

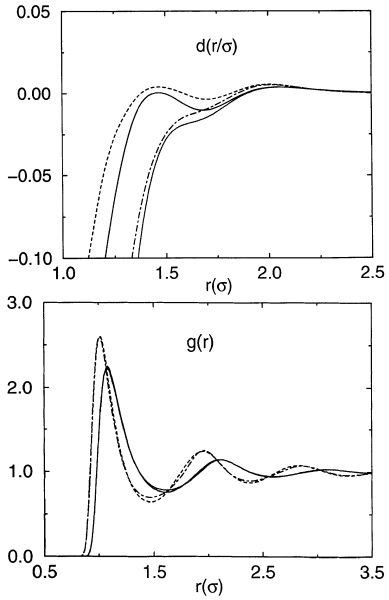

Fig. 5. a A comparison of the HNCD3 (*solid line*) and HNCS (*dotted line*) bridge functions for the phase point ($T^* = 2.74,\ \rho^* = 1.0$). The *dashed* and *dot-dashed lines* compare the HNCD3 and HNCS bridge functions, respectively, for the phase point ($T^* = 1.35,\ \rho^* = 0.7$). **b** Radial distributions from MD simulation (*solid line*) and HNCS theory (*dotted line*) at the phase point ($T^* = 1.35,\ \rho^* = 0.7$). The *dashed* and *dot-dashed lines* represent MD and HNCS calculations, respectively, for the phase point ($T^* = 2.74,\ \rho^* = 1.0$)

together is in the first well near 1.5 σ, but the differences are small. Finally, the HNCS expression has made solution of the integral equations at the point $(T^* = 1.35, \rho^* = 0.3)$ possible, further correcting the position of the liquid-gas phase boundary.

5 Summary and conclusions

Bridge diagram integrals of the first two orders of density, $d^{(2)}(r)$ and $d^{(3)}(r)$, have been evaluated exactly for pure fluids of LJ spheres for several points on the phase diagram. The bridge diagrams were evaluated using the expansion in Legendre polynomials proposed by Attard and Patey [10] for hard sphere fluids. The convergence properties of the expansions in Legendre polynomials, as well as the dependence on the number of polynomials used for the gaussian quadrature of the radial variables, were investigated. It was discovered that convergences with respect to these two types of integration were independent of each other and of the well depth of the LJ sphere potential function employed. Furthermore, the converged functions can be calculated within reasonable CPU times on a workstation.

The HNCD2 and HNCD3 integral equation theories were found to systematically correct the thermodynamic quantities and radial distributions of the HNC theory, with the HNCD3 results being superior to the PY theory results at some phase points. The addition of bridge diagrams was also found to move the position of the well-known liquid-gas phase separation, which is predicted at too high a density by the HNC theory, towards its correct position.

An attempt was made to predict the missing bridge diagrams of higher order by assuming the functional form from the HNCD3 theory. It was found that while some thermodynamic properties could be improved, it was not possible to improve all thermodynamic and structural properties at the same time. It was concluded that any approximate bridge function that improves all measurable criteria concurrently toward the corresponding correct results is probably a good approximation to the correct total bridge function.

Following a successful method of predicting high-order hard sphere bridge diagrams using a (1,2)-Padé approximant, we propose a series sum which uses the short-range gaussian structure of the known bridge diagrams to predict the diagrams which are not known. The result, which we call the HNCS result, gives improved thermodynamics and structure everywhere, surpassing the PY equation. The HNCS RDFs are virtually indistinguishable from simulation RDFs. In most cases the difference between simulation and theoretical result is of the same order as the error in the simulations themselves.

While this work is interesting simply as a contribution to the theory and statistical mechanics of fluids, it should properly be regarded as a preliminary step in the attempt to achieve accurate results for models of real systems. While the only chemical systems which are modeled well by pure LJ spheres are the noble gases [7], the LJ interaction is used widely to model the short-range repulsion and dispersive forces between the constituent atoms in models of biological systems [36, 37]. It will be most interesting to see whether similar improvements to theoretical results are possible in such models. Future study in this area will need to build on this work by examining mixtures, namely long-range ionic and dipolar species as well as multi-atom molecular species.

Acknowledgements. J. S. P. wishes to thank Dr. G. Lynch for sharing her expertise with the MD code ESP as well as generous gifts of time, and Dr. P. E. Smith for making his MD code available. The authors thank the Robert A. Welch Foundation, the N.I.H., the N.S.F. and the Texas Coordinating Board for partial support of this work.

References

1. Hirata F, Levy RM (1987) Chem Phys Lett 136:267
2. Beglov D, Roux B (1996) J Chem Phys 104:8678
3. Perkyns JS, Pettitt BM (1994) Biophys Chem 51:129
4. Perkyns JS, Pettitt BM (1995) J Phys Chem 99:1
5. Perkyns JS, Pettitt BM (1996) J Phys Chem 100:1323
6. Perkyns JS, Pettitt BM (1996) J Am Chem Soc 118:1164
7. Hansen JP, McDonald IR (1986) Theory of simple liquids, 2nd edn. Academic, London
8. Høye JS, Stell G (1976) J Chem Phys 65:18
9. Chandler D (1977) J Chem Phys 67:1113
10. Attard P, Patey GN (1990) J Chem Phys 92:4790
11. Verlet L (1964) Physica 30:95
12. Rowlinson JS (1965) Mol Phys 9:217
13. Lado F (1973) Phys Rev A 8:2548

14. Rosenfeld Y, Ashcroft NW (1979) Phys Rev A 20:1208
15. Rossky PJ, Dudowicz JB, Tembe BL, Friedman HL (1980) J Chem Phys 73:3372
16. Rogers FJ, Young DA (1984) Phys Rev A 30:999
17. Zerah G, Hansen JP (1986) J Chem Phys 84:2336
18. Vlachy V, Pohar C, Haymet ADJ (1988) J Chem Phys 88:2066
19. Perkyns JS, Pettitt BM (1997) Chem Phys Lett 190:626
20. Perkyns JS, Pettitt BM (1997) J Chem Phys 97:7656
21. Forester TR, McDonald IR (1991) Mol Phys 72:643
22. Smith PE, Dang LX, Pettitt BM (1991) J Am Chem Soc 113: 67
23. York DM, Wlodawer A, Pedersen LG, Darden TA (1994) Proc Natl Acad Sci USA 91:8715
24. Smith PE, Pettitt BM (1994) J Phys Chem 98:9700
25. Toukmaji AY, JAB Jr (1996) Comput Phys Commun 95:73
26. Uhlenbeck GE, Ford GW (1962) The theory of linear graphs with applications to the theory of the virial development of the properties of gases. In: DeBoer J, Uhlenbeck GE (eds) Studies in statistical mechanics, volume 1. North-Holland, Amsterdam, p 182
27. Ornstein LS, Zernike F (1914) Proc K Ned Akad Wet 17:793
28. Percus JK, Yevick GJ (1958) Phys Rev 110:1
29. Attard P (1989) J Chem Phys 91:3072
30. Barker JA, Monaghan JJ (1962) J Chem Phys 36:2564
31. Haymet ADJ, Rice SA, Madden WG (1981) J Chem Phys 74:3033
32. Abramowitz, Stegun IA (eds) (1970) Handbook of mathematical functions. Dover, New York
33. Press WH, Flannery BP, Teukolsky SA, Vetterling WT (1988) Numerical ecipes in C: the art of scientific computing. Cambridge University Press, New York
34. Allen MP, Tildesley DJ (1987) Computer simulation of liquids. Oxford University Press, Oxford
35. Hansen JP, Verlet L (1969) Phys Rev 184:151
36. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M (1983) J Comput Chem 4:187
37. Weiner SJ, Kollman PA, Nguyen DT, Case DA (1986) J Comput Chem 7:230