Finite-size effects in molecular dynamics simulations: Static structure factor and compressibility.
II. Application to a model krypton fluid

J. J. Salacuse

*Science and Mathematics Department, GMI Engineering and Management Institute, 1700 West Third Avenue,
Flint, Michigan 48502-2276*

A. R. Denton* and P. A. Egelstaff

Department of Physics, University of Guelph, Guelph, Ontario, Canada N1G 2W1

M. Tau

Dipartimento di Fisica dell'Università, Viale delle Scienze, I 43100, Parma, Italy

L. Reatto

Dipartimento di Fisica dell'Università, via Celoria 16, 20133, Milano, Italy

The method described in the preceding paper [J. J. Salacuse, A. R. Denton, and P. A. Egelstaff, preceding paper, Phys. Rev. E **53**, 2382 (1996)] for computing the static structure factor $S(Q)$ of a bulk fluid is used to analyze molecular dynamics computer simulation data for a model krypton fluid whose atoms interact via a truncated Aziz pair potential. Simulations have been carried out for two system sizes of $N=706$ and 2048 particles and two thermodynamic states, described by a common reduced temperature $T^*=1.51$ and reduced densities $\rho^*=0.25$ and 0.4. Results presented include the N-particle radial distribution function $g_N(r)$ and the bulk static structure factor $S(Q)$. In addition we calculate the direct correlation function $c(r)$ from the full $S(Q)$. In comparison with corresponding predictions of the modified hypernetted chain theory, the results are generally in excellent agreement at all r and Q, to within random statistical errors in the simulation data.

PACS number(s): 61.20.Gy, 61.20.Ja, 02.70.Ns, 05.70.Ce

I. INTRODUCTION

In the preceding paper [1] (paper I) we have presented a method for computing the static structure factor $S(Q)$ of a uniform bulk fluid from computer simulation data for a finite (N-particle) system. Starting from the Fourier transform relation between $S(Q)$ and the radial distribution function $g(r)$ [Eq. (1a) in paper I] and exploiting a theoretical $O(1/N)$ finite-size correction for the N-particle function $g_N(r)$ [Eq. (18) in paper I], the method prescribes a corresponding analytic correction for the N-particle distribution function $S_N(Q,R)$ that yields the bulk $S(Q)$ at radial distances R that lie within the asymptotic range of $g_N(r)$. Consequently, the thermodynamic state should not be near the critical point and the density should not be large, in order that $g(r)$ may decay to near its asymptotic limit within the range of the simulation box. The method is expected to be valid for *all* wavevector magnitudes Q (including $Q=0$) and for thermodynamic states for which $S(0)$ (proportional to the isothermal compressibility) remains of order unity. For practical applications it is important that the derivatives of $g(r)$ with respect to density [see Eq. (19) of paper I] are negligible, so that the simplified Eqs. (22) and (23) of pa-

per I become applicable.

In the present paper, in order to demonstrate and test the above method, we determine $S(Q)$ at all Q from molecular dynamics (MD) simulations of a dense krypton fluid. Also we compute the correlation functions for the infinite system on the basis of the modified hypernetted chain equation (MHNC), which gives accurate results at the states used here. In Sec. II we begin by describing the details of our simulations and then explain our analysis of the data to extract $S(Q)$. In Sec. III we present our (MD) results, which include standard plots of $g_N(r)$ vs r, illustrating the finite-size effects, and of the final $S(Q)$ vs Q. For comparison we show the corresponding predictions of MHNC integral equation theory. In Sec. IV we discuss the accuracy of our results and potential limitations of our method and present the data on the direct correlation function. In Sec. V we summarize our results and draw conclusions regarding the utility of our method. Finally, in the Appendix we briefly review our use of the MHNC integral equation.

II. SIMULATIONS AND ANALYSIS

We have carried out a series of MD simulations for a dense krypton fluid in which the particles were made to interact via the Aziz pair potential [2], characterized for krypton by the parameters $\sigma_0=3.579$ Å (atomic diameter) and $\epsilon/k_B=200$ K (well depth), with the potential truncated and shifted to zero at a cutoff distance of

*Present address: Institut für Theoretische Physik, Technische Universität, Wiedner Hauptstrasse 8-10, A-1040 Wien, Austria.

$r_c = 4\sigma_0$. The simulation program is based on a fifth-order Gear predictor-corrector algorithm [3] that solves the classical equations of motion for the particle trajectories in the microcanonical ensemble, with periodic boundary conditions. The simulations were performed for two system sizes $N = 706$ and 2048 and at two different thermodynamic states, defined by a common reduced temperature $T^* \equiv k_B T / \epsilon \cong 1.51$ and reduced densities $\rho^* \equiv \rho \sigma_0^3$ of 0.25 and 0.4. (For comparison, the critical point of krypton occurs at $T^* = 1.05$, $\rho^* = 0.3$.) These system sizes and states were chosen to best illustrate the application of our method since their compressibilities are relatively large and since implicit size effects are expected to be negligible [1]. During the initial equilibration stage, the velocities were periodically rescaled to maintain the desired constant temperature. For each choice of N and ρ^* several independent simulations were performed, ranging in duration from 25×10^4 to 10^6 time steps, corresponding to a physical time range of roughly 1250–5000 ps. The simulations were run on IBM RS/6000 computer workstations.

During the simulations, the configurations (particle coordinates) were periodically stored and subsequently analyzed to determine $S_N(Q, R)$ by the following procedure. From a given coordinate set at time t_k say, corresponding to the kth time step, a particle i is chosen arbitrarily and all particles j contained within a sphere of radius R centered on the chosen particle are identified. The factor $[\sin(Qr_{ij})/Qr_{ij}]$, r_{ij} being the distance between particles i and j, is then summed over all j (including $j = i$). Next, another particle i is chosen, the same sum is calculated, and the process is repeated until every particle has served as the central particle. Finally, the average over all $i = 1, \ldots, N$ of the calculated sums represents a single estimate, at time t_k, of the quantity $\langle N(Q, R) \rangle$, which appears in Eq. (4a) of paper I. This estimate, which we denote by $X_k(Q)$, may be expressed more concisely in the form

$$X_k(Q) = \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} \left[\frac{\sin[Qr_{ij}(t_k)]}{Qr_{ij}(t_k)} \right] \Delta_{ijk}(R)$$

$$= \frac{2}{N} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \left[\frac{\sin[Qr_{ij}(t_k)]}{Qr_{ij}(t_k)} \right] \Delta_{ijk}(R) + N ,$$

(1a)

where

$$\Delta_{ijk}(R) \equiv \begin{cases} 1 & \text{if } r_{ij}(t_k) \leq R \\ 0 & \text{if } r_{ij}(t_k) > R , \end{cases}$$

(1b)

and $r_{ij}(t_k) = |\mathbf{r}_i(t_k) - \mathbf{r}_j(t_k)|$ is defined as the minimum distance at time t_k between particle i at position $\mathbf{r}_i(t_k)$ and either particle j or its nearest periodic image (minimum image convention).

Now the average of the estimates $X_k(Q)$ over some number M of *independent* coordinate sets, defined by

$$\overline{X}(Q, M) = \frac{1}{M} \sum_{k=1}^{M} X_k(Q) ,$$

(2)

where M is limited in practice by the duration of the simulation, represents an approximation for $\langle N(Q, R) \rangle$ that clearly improves in accuracy with increasing M. Furthermore, each estimate is derived from a sample whose variance may be approximated by

$$\sigma_k^2 \approx \frac{1}{M} \sum_{k=1}^{M} [X_k(Q) - \overline{X}(Q, M)]^2 ,$$

(3)

and thus the variance in $\overline{X}(Q, M)$ is given by

$$\sigma_M^2 = \frac{1}{M} \sigma_k^2 = \frac{1}{M^2} \sum_{k=1}^{M} [X_k(Q) - \overline{X}(Q, M)]^2 .$$

(4)

As a measure of the uncertainty in $\overline{X}(Q, M)$, we take the standard deviation σ_M and write

$$\langle N(Q, R) \rangle = \overline{X}(Q, M) \pm \sigma_M ,$$

(5)

with the usual understanding that about 68% of the estimates fall within σ_M of $\overline{X}(Q, M)$, about 95% within $2\sigma_M$, etc. Now combining Eq. (4) of paper I with Eq. (5) above, we obtain

$$S_N(Q, R) = \overline{X}(Q, M) - \tfrac{4}{3}\pi\rho R^3 u(QR) \pm \sigma_M ,$$

(6)

with the $Q = 0$ limit

$$S_N(0, R) = \overline{X}(0, M) - \tfrac{4}{3}\pi\rho R^3 \pm \sigma_M .$$

(7)

Equation (7) provides a practical means of extracting the function $S_N(0, R)$ from simulation data, which when substituted into Eq. (23) of paper I at sufficiently large R leads to an approximate $S(0)$. Finally, substituting this $S(0)$, together with the large-R $S_N(Q, R)$ [from Eq. (6)], into Eq. (22) of paper I yields in turn a well-defined approximation for $S(Q)$ at arbitrary Q. This procedure assumes that the density derivative terms of $g(r)$ give a negligible contribution to the correction, and it will be shown later that this is so for the present cases.

Now Eq. (5) is strictly valid only if the estimates $X_k(Q)$, $k = 1, \ldots, M$, used to compute $\overline{X}(Q, M)$ are statistically independent. To ensure independence, successive configurations must be separated by a time interval sufficiently long to eliminate significant correlations. As a means of determining a sufficient length of interval, we have computed the N-particle intermediate scattering function $I_N(Q, t, R)$ [Eqs. (27b) and (28) of paper I] at $Q = 0$, which has the character of a time correlation function. The key quantity in the calculation $\langle N(Q, t, R) \rangle$ may be computed by a procedure similar to the one described above for determining $\langle N(Q, R) \rangle$, with the exception that at time $(t_k + t)$ particles are counted within a sphere of radius R that at an *earlier* time t_k was centered on a particle. The function $I_N(0, t, R)$ thus reflects, as a function of time, the loss of memory of the initial presence at $t = 0$ of a central particle. Equivalently, it measures the degree of correlation between configurations separated by time t. Note that in general, for fixed N, $I_N(Q, t, R) \rightarrow 0$ as $t \rightarrow \infty$.

An estimate of $\langle N(Q, t, R) \rangle$ at time t_k, which we denote by $X_k(Q, t)$, may be expressed in the form [cf. Eq. (1a)]

$$X_k(Q,t) = \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{N} \left[\frac{\sin[Qr_{ij}(t_k,t)]}{Qr_{ij}(t_k,t)} \right] \Delta_{ijk}(t,R) , \qquad (8)$$

where $r_{ij}(t_k,t) = |r_i(t_k) - r_j(t_k+t)|$ and $\Delta_{ijk}(t,R)$ is defined as in Eq. (1b) but with $r_{ij}(t_k)$ replaced by $r_{ij}(t_k,t)$. As before, an average of the estimates $X_k(Q,t)$ over configurations gives an approximation for $\langle N(Q,t,R) \rangle$, which may be substituted into Eq. (28a) of paper I to obtain an approximation for $I_N(Q,t,R)$. Since our objective is to roughly assess the correlation (or decay) time of $I_N(0,t,R)$, we do not concern ourselves with the standard deviation of the estimates.

III. RESULTS: MD AND MHNC DATA AND ANALYSIS

As mentioned above, we have carried out MD simulations at two reduced densities $\rho^* = 0.4$ and 0.25 and analyzed the data to compute $S(Q)$ by the method described in paper I. In this section we first present MD results for the $\rho^* = 0.4$ case, followed by results for the $\rho^* = 0.25$ case. For comparison, we present results from the corresponding MHNC calculations [4], obtained using the method described in the Appendix. Since the MHNC data are expressed in units of the reduced variables r/σ_m and $Q\sigma_m$, where $\sigma_m = 4.012$ Å is the distance at which the Aziz potential reaches its minimum, functions of r and Q will often be expressed in these same units.

A. Case 1: $\rho^* = 0.4$

We begin by presenting our results for the N-particle radial distribution function $g_N(r)$. Although it is not strictly necessary to compute $g_N(r)$ in order to determine $S(Q)$ by our method, it nevertheless serves as a simple and useful check on the accuracy of the simulations and directly illustrates the finite-size effects. Furthermore, knowing the behavior of $g_N(r)$ helps in understanding the behavior of the distribution function $S_N(Q,R)$. As an example, Fig. 1 shows $g_N(r)$ vs r for $N=706$, determined by analyzing the MD configurations using standard methods [3]. The error bars in all figures reflect statistical uncertainty of plus or minus one standard deviation. These data represent an average over 188 configurations, each configuration yielding one estimate of $g_N(r)$.

To ensure independence of the estimates, as discussed previously, we selected configurations separated by a fixed time interval of 15 ps plus a *random* interval between 0 and 15 ps, giving an average separation of about 23 ps. Inclusion of a random interval allows a different subset of configurations to be selected simply by varying the seed of the random number generator. Averaging results over different seeds then allows for optimal use of the MD data, a point we elaborate on below with regard to the calculation of $S_N(0,R)$. The average separation between configurations necessary to ensure independence of the estimates was determined by examining the time dependence of the time correlation function $I_N(0,t,R)$. As Fig. 2 illustrates, $I_N(0,t,R)$ generally decays with increasing time as successive configurations become un-

correlated with an arbitrarily chosen initial configuration. When the function has decayed sufficiently, after an interval Δt say, estimates obtained from configurations separated in time by Δt may be assumed independent. Of course, the more rigorously that independence is en-
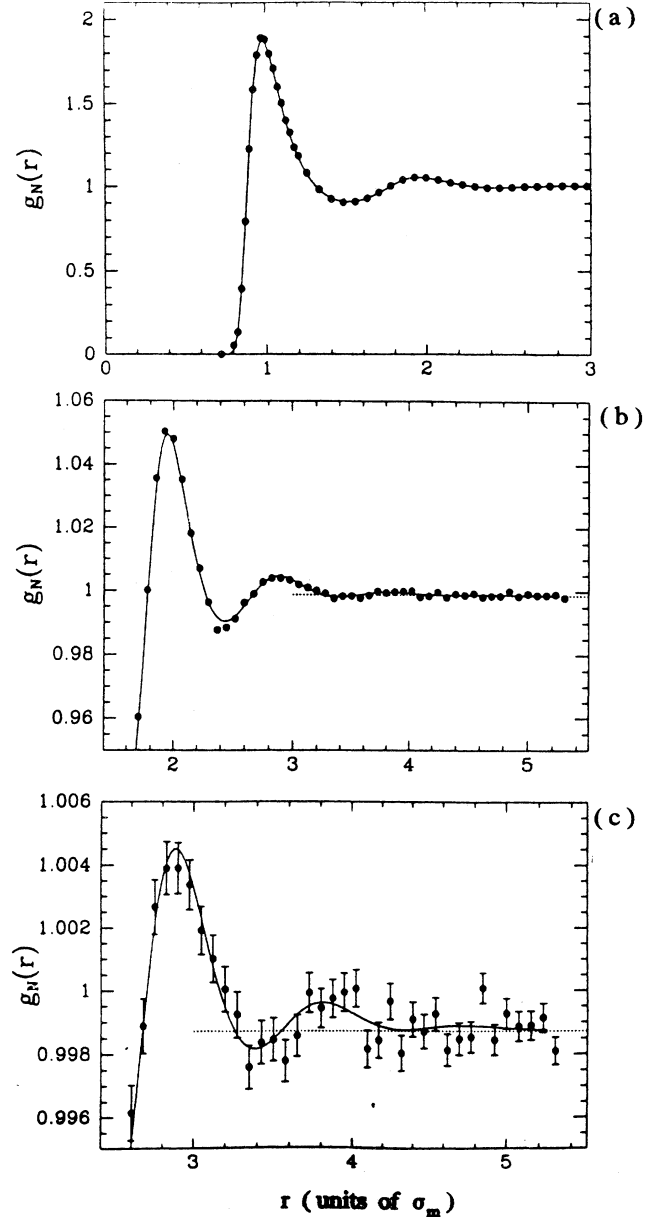

FIG. 1. Radial distribution function $g_N(r)$ vs radial distance r in units of σ_m (distance of Aziz pair potential minimum) for $N=706$, $\rho^* = 0.4$. The dots denote our MD simulation data, the solid curve the MHNC results and the dotted curve the asymptotic limit $[1-S(0)/N]$, according to Eq. (2) of paper I. Panels (a), (b), and (c) highlight the first, second, and third peaks, respectively. Panel (c) also includes the MD error bars.

forced, by lengthening Δt, the fewer the number of estimates available (for a given total simulation time) and hence the greater the statistical uncertainties. As a practical compromise, we have adopted the following simple criterion for independence: Two configurations separated by a time interval Δt are deemed to be independent if $I_N(0,\Delta t,R)/I_N(0,0,R) < 0.15$ for all R. The cutoff value

FIG. 2. Time correlation function $I_N(0,t,R)$ [related to intermediate scattering function $I_N(Q,t)$ at $Q=0$] vs time t for various system sizes, states, and radial distances R in units of σ_m: (a) $N=706$, $\rho^*=0.4$, $R=3$ (squares), $R=5$ (circles); (b) $N=706$; $\rho^*=0.25$, $R=3$ (squares), $R=6$ (circles); and (c) $N=2048$, $\rho^*=0.25$, $R=6$ (squares), $R=8$ (circles).

of 0.15 was chosen empirically. Smaller values (for longer Δt's) were found not to significantly change the results, but merely to increase the statistical errors. For $N=706$, Fig. 2(a) shows that after an interval $\Delta t \cong 23$ ps, $I_N(0,t,R)$ has decayed to less than 15% of its initial value (at $t=0$).

Also shown in Fig. 1, for comparison, are the corresponding MHNC predictions for $g_N(r)$. Since MHNC theory actually produces the infinite-system function $g(r)$, we convert (see the Appendix) to $g_N(r)$ by subtracting $S(0)/N$, according to Eq. (2) of paper I, using the MHNC value of $S(0)=0.9078$ for this purpose. For reference, the expected asymptotic limit $g_N(r) \sim 1 - S(0)/N$ is also indicated in Fig. 1.

We now turn to our results for the N-particle distribution function $S_N(Q,R)$, its infinite-system counterpart $S(Q,R)$, and ultimately the desired static structure factor $S(Q)$. The finite-size correction that converts $S_N(Q,R)$ to $S(Q,R)$ [Eq. (22) of paper I] requires a determination of $S(0)$. To this end, we first compute the function $S_N(0,R)$, i.e., $S_N(Q,R)$ at $Q=0$, by means of Eq. (7). The data for $N=706$ represent an average over 188 estimates, obtained from configurations separated by an average of 23 ps [composed of a fixed 15-ps interval plus a random interval between 0 and 15 ps, as noted above for $g_N(r)$]. For $N=2048$, we averaged over 88 estimates, separated by an average of 30 ps. A longer separation was used because $I_N(0,t,R)$ decays more slowly for the larger system. This behavior is consistent with the tendency of $I_N(0,t,R)$ towards a time-independent constant as N increases, the bulk limit ($N \to \infty$) being $I(0,t,R)=S(0)$, for R sufficiently large and t finite. We note that the use of a random interval here ensures that $S_N(0,R)$ at different values of R is obtained from different subsets of configurations. Figure 3 and Table I give our results for $S_N(0,R)$ vs R for $N=706$ and 2048, together with the corresponding MHNC predictions. To obtain the latter, we used Eq. (22) of paper I, with the MHNC values of $S(0)$ and $S(0,R)$.
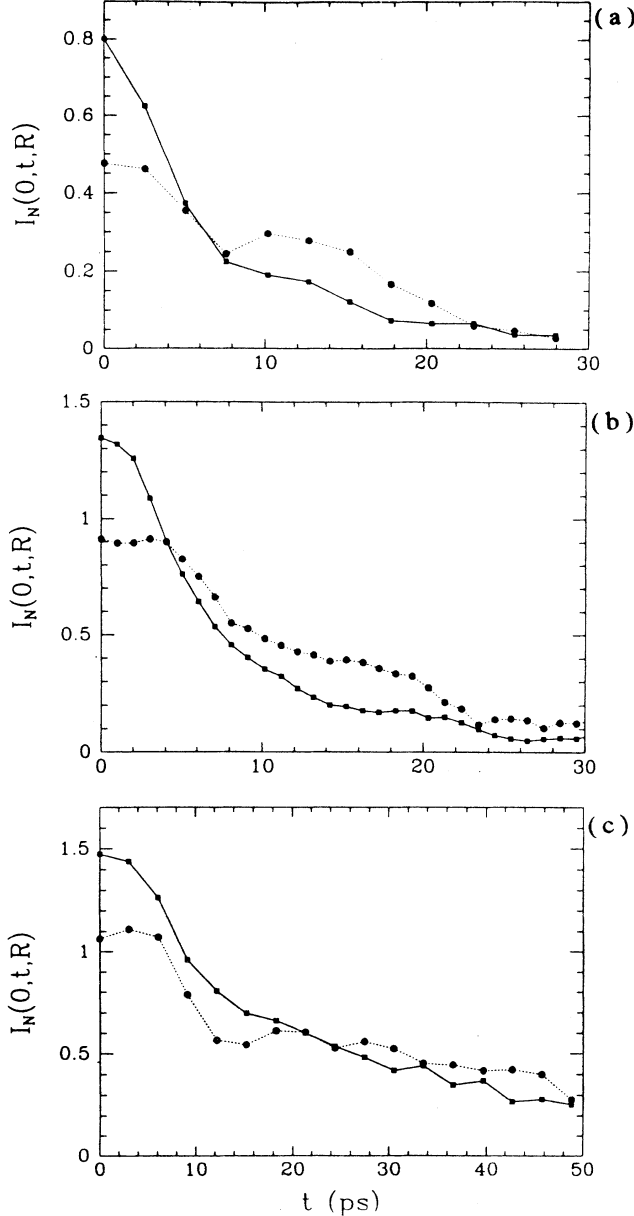
TABLE I. Simulation and MHNC data for $S_N(0,R)$ at $\rho^*=0.4$.

R/σ_m	$N=706$		$N=2048$	
	Simulation	MHNC	Simulation	MHNC
1.5	0.724±0.006	0.712	0.724±0.005	0.717
2.0	0.648±0.011	0.609	0.650±0.008	0.624
2.5	0.736±0.015	0.719	0.762±0.012	0.749
3.0	0.763±0.021	0.735	0.823±0.017	0.788
3.5	0.741±0.026	0.716	0.824±0.023	0.800
4.0	0.708±0.031	0.681	0.846±0.028	0.807
4.5	0.631±0.035	0.609	0.825±0.038	0.789
5.0	0.506±0.038	0.517	0.802±0.040	0.765
5.5			0.751±0.047	0.726
6.0			0.668±0.047	0.678
6.5			0.607±0.052	0.618
7.0			0.546±0.051	0.548
7.5			0.434±0.054	0.466

From the behavior of $g_N(r)$, the general appearance of $S_N(0,R)$ may be understood as follows. Differentiating Eq. (3b) of paper I with respect to R at $Q=0$ yields the relation

$$\frac{dS_N(0,R)}{dR} = 4\pi\rho R^2 [g_N(R) - 1] . \tag{9}$$

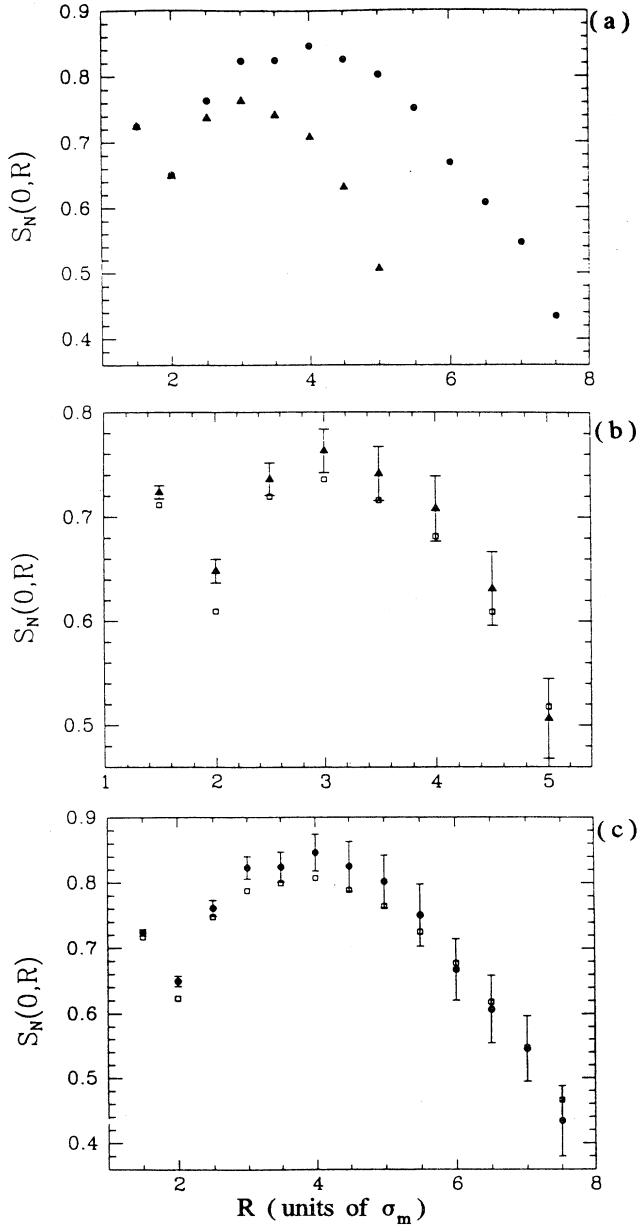Hence $S_N(0,R)$ is an increasing function of R wherever

$g_N(R) > 1$ and a decreasing function wherever $g_N(R) < 1$. As seen from Fig. 1, when $N=706$, $g_N(R) < 1$ for $R \geq 3.2$, which accounts for the monotonic decrease of $S_N(0,R)$ in the same region seen in Fig. 3. In general, the concave shape of $S_N(0,R)$ can be attributed to the structure of $g_N(R)$. As N increases, the asymptotic limit of $g_N(R) \sim 1 - S(0)/N$ also increases, causing the maximum of $S_N(0,R)$ to grow in amplitude and to shift to larger values of R.

The technique of separating estimates by a time interval consisting of both a fixed and a random element is used consistently throughout this work. As an illustration of this technique, we elaborate on the calculation of $S_N(0,R)$. For a given value of R, say R_1, starting with the initial configuration, estimates of $S_N(0,R_1)$ are obtained, each configuration giving one sample. The average separation in time of the configurations is made large enough to ensure that the estimates are essentially independent. Use of a random time interval eliminates the possibility of sampling the data set at an interval corresponding to a local maximum in the correlation function, which would increase the correlation between estimates. Assuming independence, estimates of $S_N(0,R_1)$ and its standard deviation are obtained. Another value of R, say R_2, is selected and the entire process, starting with the initial configuration, is repeated to obtain an estimate of $S_N(0,R_2)$. The random time interval ensures that, in general, estimates of $S_N(0,R_1)$ and $S_N(0,R_2)$ are not obtained from the same set of configurations. Samples associated with *distinct* value of R, however, may not be separated by a large enough time interval to guarantee the independence of the estimates of $S_N(0,R_1)$ and $S_N(0,R_2)$. Thus, for example, the entries in Table I of the simulation $S_N(0,R)$ may not be statistically independent of one another.

Having computed $S_N(0,R)$ by MD simulation, we next approximate $S(0)$ by means of Eq. (23) of paper I, which, for sufficiently large R, we expect to be accurate to $O(1/N)$. Table II lists the resulting estimates of $S(0)$

FIG. 3. Distribution function $S_N(0,R)$ vs R (in units of σ_m) for $\rho^* = 0.4$: (a) MD data for $N=706$ (triangles) and $N=2048$ (circles); (b) MD data (triangles) compared with MHNC results (squares) for $N=706$; and (c) MD data (circles) and MHNC results (squares) for $N=2048$.

TABLE II. Estimate of $S(0)$ via Eq. (23) of paper I at $\rho^* = 0.4$.

	$S(0)$	
R/σ_m	$N=706$	$N=2048$
1.5	0.73 ± 0.01	0.73 ± 0.01
2.0	0.67 ± 0.01	0.66 ± 0.01
2.5	0.78 ± 0.02	0.78 ± 0.01
3.0	0.84 ± 0.02	0.85 ± 0.02
3.5	0.86 ± 0.03	0.87 ± 0.02
4.0	0.90 ± 0.04	0.91 ± 0.03
4.5	0.91 ± 0.05	0.92 ± 0.04
5.0	0.87 ± 0.07	0.94 ± 0.05
5.5		0.93 ± 0.06
6.0		0.89 ± 0.06
6.5		0.89 ± 0.08
7.0		0.90 ± 0.08
7.5		0.84 ± 0.11

over a range of R. The *minimum* value of R for which the approximation is reliable may be identified as the value beyond which $g_N(R)$ is essentially constant to within statistical fluctuations. For $N=706$, Fig. 1 indicates that $g_N(R) \approx$ const for $R/\sigma_m \geq 4.0$. Furthermore, from Table II, $S(0)$ is also seen to be essentially constant in this region. Thus, averaging the $N=706$ entries in Table II at $R/\sigma_m = 4.0$, 4.5, and 5.0 yields the prediction $S(0)=0.89\pm0.05$, which is in good agreement with the MHNC value of 0.9078.

For $N=2048$, the same technique may be used to define the appropriate large-R limit of Eq. (23) of paper I from inspection of the long-range tail of $g_N(R)$. Alternatively, *in the absence of* $g_N(R)$, the limit may also be identified as the distance at which $S_N(0,R)$ reaches its maximum. The decrease of $S_N(0,R)$ beyond this distance implies that $g_N(R)-1<0$. As seen from Table I, from $N=2048$, $S_N(0,R)$ reaches its maximum at approximately $R/\sigma_m = 4.0$. Assuming that the upper and lower bounds of the function $[g_N(R)-1]$ are roughly equidistant from its asymptotic limit of $-S(0)/N$ leads to the

bounds $-2S(0)/N < g_N(R)-1 < 0$ for $R/\sigma_m \geq 4.5$. This in turn suggests that oscillations in $g_N(R)$ are relatively weak for $R/\sigma_m \geq 4.5$ and we assume $g_N(R)$ to be essentially constant in this region to within statistical uncertainties inherent in the simulation data. Hence the entries in Table II for $R/\sigma_m \geq 4.5$ may be taken to represent the asymptotic limit of Eq. (23) of paper I. Averaging the entries at $R/\sigma_m = 4.5-7.5$ then yields the independent prediction $S(0)=0.90\pm0.07$, the statistical error being larger than for $N=706$ because of shorter runs for $N=2048$. The averaging technique is particularly useful when $S(0)$ exhibits significant fluctuations at large R. Although the large- and small-system predictions of $S(0)$ are consistent with each other, the latter were naturally considerably easier to compute. In Sec. IV we discuss the relative advantages of large- and small-system simulations for the determination of $S(Q)$.

Another form of comparison between our MD simulation method and MHNC theory is illustrated by Fig. 4, where the MD $S_N(0,R)$ has been converted to $S(0,R)$ via Eq. (22) of paper I, using the appropriate MD values of
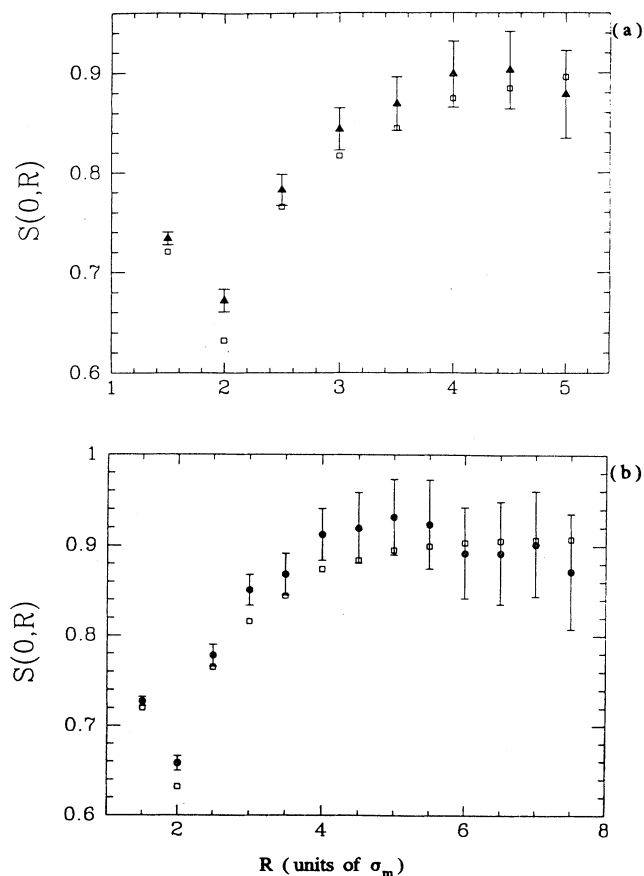

FIG. 4. Distribution function $S(0,R)$ vs R (in units of σ_m) for $\rho^*=0.4$: (a) MD data (triangles) compared with MHNC results (squares) for $N=706$ and (b) MD data (circles) and MHNC results (squares) for $N=2048$.
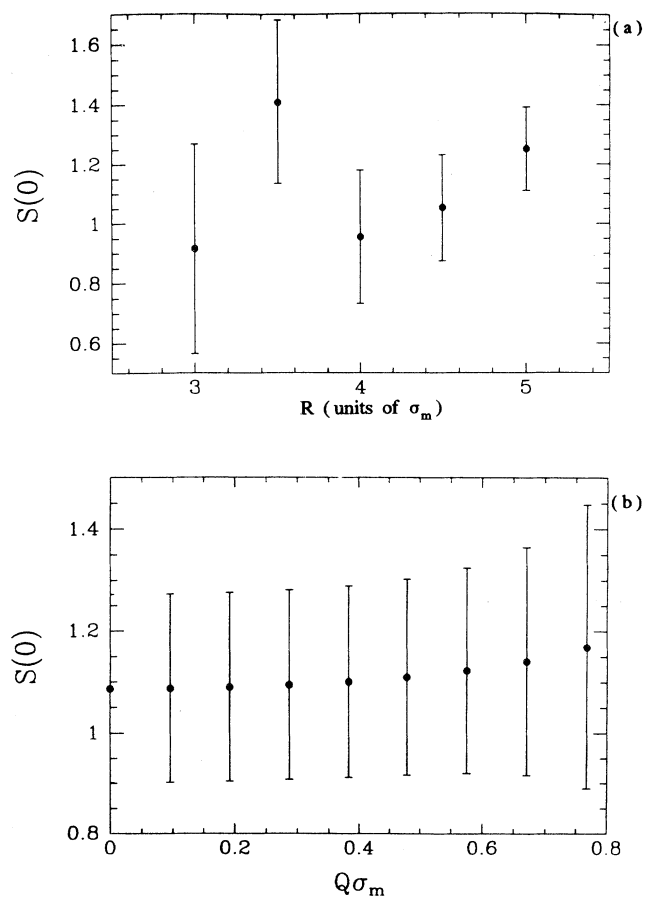
FIG. 5. Compressibility ratio $S(0)$, obtained from MD data for $\rho^*=0.4$ using the "difference formula" approach: (a) $S(0)$ vs R for $Q=0$ and (b) $S(0)$ vs Q for $R=5$ (in units of σ_m). The error bars are determined by the uncertainties in $S_N(Q,R)$ in Eq. (24) of paper I [1].

$S(0)$. The fact that $S(0,R)$ reaches its asymptotic (large-R) limit for $R < L/2$ confirms that the $N=706$ system is sufficiently large for our method to be applicable. In general, however, we expect the method not to be reliable for systems so small that $S(0,R)$ fails to reach a plateau by $R = L/2$, since in this case $S(0)$ cannot be accurately predicted. Note that the plateau level of $S(0,R)$ agrees with the value of $S(0)$ predicted by Eq. (23) of paper I, as it should for consistency.

We have also attempted to determine $S(0)$ by the approach based on the "difference formula" [Eq. (24) in paper I]. Figure 5 shows the resulting $S(0)$, both as a function of R at fixed Q and as a function of Q at fixed R. The results are in fair agreement with the previous prediction and are seen to be essentially independent of Q, as expected, for Q smaller than the first zero of $u(QR)$, which for $R = 5\sigma_m$ occurs at $Q \cong 1\sigma_m^{-1}$. Although the rather large inherent uncertainties limit the utility of this approach, the consistency between predictions obtained

by two different means does serve as a useful check on the calculations.

The distribution function $S_N(Q,R)$ was calculated via Eq. (6) and is shown in Fig. 6. Finally, having determined $S(0)$ as described above, we obtain $S(Q,R)$ after making the finite-size correction prescribed by Eq. (22) of paper I. For R sufficiently large, i.e., within the asymptotic range of $g_N(r)$, $S(Q,R)$ is approximately $S(Q)$. Averaging over the asymptotic range of $g_N(r)$ (specifically $R/\sigma_m = 4.0$, 4.5, and 5.0 for $N=706$) then yields a more precise result for the bulk $S(Q)$, which is plotted vs Q in Fig. 6, together with the corresponding
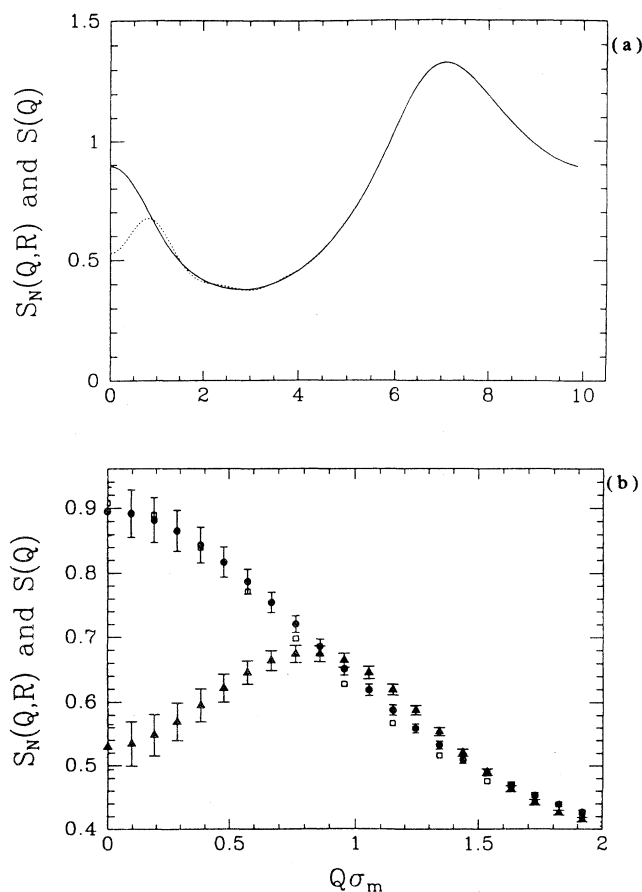

FIG. 6. N-particle distribution function $S_N(Q,R=5)$ (R in units of σ_m) [dotted curve in (a), triangles in (b)] and bulk static structure factor $S(Q)$ [solid curve in (a), circles in (b)] vs wavevector magnitude $Q\sigma_m$ for $N=706$, $\rho^*=0.4$. Also shown in (b) are the MHNC results (squares) for $S(Q)$.
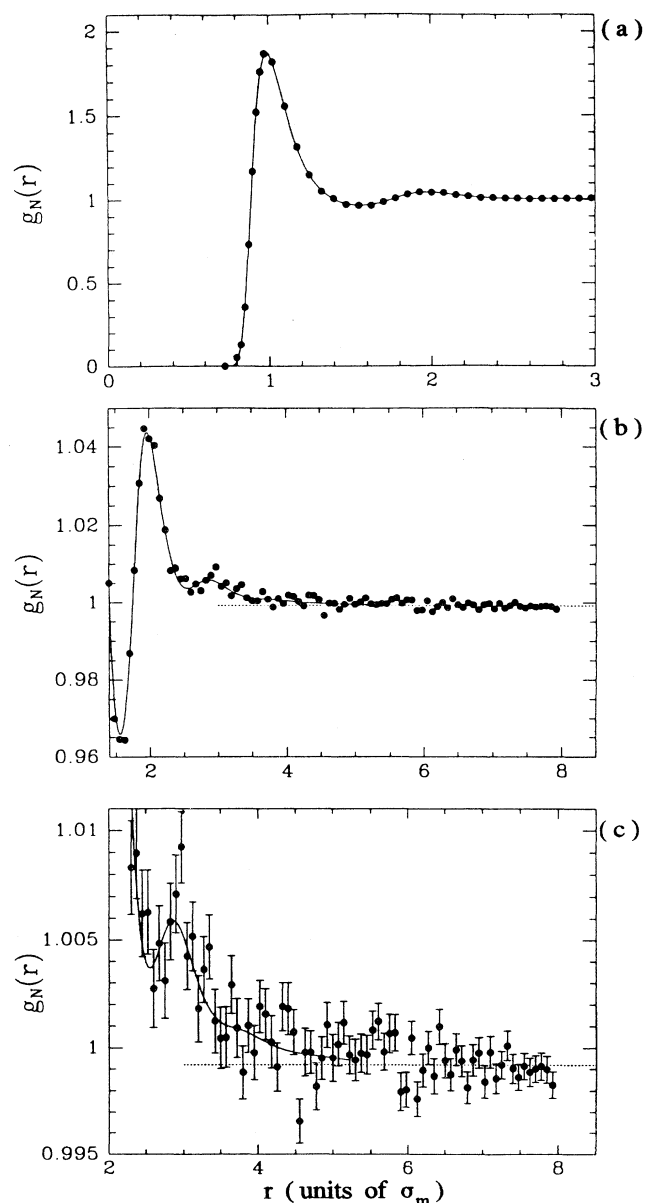

FIG. 7. Same as Fig. 1, but for $N=2048$, $\rho^*=0.25$.

MHNC predictions for $S(Q)$. As Fig. 6 clearly illustrates, the finite-size effects become increasingly significant as $Q \to 0$.

B. Case 2: $\rho^* = 0.25$

The second state that we have considered presents more of a challenge to our method, as it is somewhat closer to the critical point. This is reflected in the slightly longer range of $g_N(r)$ in comparison with the first state, characteristic of near-critical behavior, as seen in Fig. 7 for $N = 2048$. As in Fig. 1, the MHNC theoretical predictions, adjusted for finite-size effects using the MHNC value of $S(0) = 1.6388$, are also shown, together with the expected asymptotic limit. The MD data were obtained by averaging over 66 independent configurations, independence being ensured by separating configurations with a fixed interval of 50 ps plus a random interval between 0 and 50 ps, for an average separation of $\Delta t = 75$ ps.

Our results for $S_N(0,R)$ are presented in Table III, alongside the size-adjusted MHNC predictions. For $N = 706$, an average separation interval of $\Delta t = 22.5$ ps (comprising a fixed 15-ps interval plus a random interval between 0 and 15 ps) was found to be sufficient, allowing an average over 333 independent configurations. Note that this is a significantly shorter interval than the 75 ps found necessary for the larger system. These intervals were again determined empirically by the requirement that the time correlation function $I_N(0,t,R)$ be smaller than 15% of its initial value. As Figs. 2(b) and 2(c) illustrate, the decrease in the decay rate of $I_N(0,t,R)$ with increasing N is more pronounced for this state, which may reflect the closer proximity to the critical point.

In Table IV we list estimates of $S(0)$, over a range of R, as obtained via Eq. (23) of paper I. Averaging the entries for $N = 706$ over $R/\sigma_m = 5.0$–6.0 yields the predic-

TABLE IV. Estimate of $S(0)$ via Eq. (23) of paper I at $\rho^* = 0.25$.

R/σ_m	$S(0)$	
	$N = 706$	$N = 2048$
1.5	1.05 ± 0.01	1.05 ± 0.01
2.0	1.12 ± 0.01	1.11 ± 0.01
2.5	1.30 ± 0.01	1.32 ± 0.02
3.0	1.42 ± 0.02	1.39 ± 0.03
3.5	1.47 ± 0.03	1.55 ± 0.03
4.0	1.50 ± 0.03	1.58 ± 0.06
4.5	1.58 ± 0.04	1.63 ± 0.06
5.0	1.63 ± 0.05	1.69 ± 0.07
5.5	1.68 ± 0.06	1.59 ± 0.08
6.0	1.63 ± 0.08	1.61 ± 0.09
6.5		1.63 ± 0.12
7.0		1.66 ± 0.13
7.5		1.82 ± 0.16
8.0		1.76 ± 0.19
8.5		1.76 ± 0.22

tion $S(0) = 1.65 \pm 0.06$. Similarly, averaging the $N = 2048$ entries over $R/\sigma_m = 6.0$–8.5 yields the independent prediction $S(0) = 1.71 \pm 0.15$, which is in agreement with the MHNC value of 1.6388. Using the MD values of $S(0)$ to make the finite-size correction to $S_N(0,R)$, we obtain the results for $S(0,R)$ plotted vs R in Fig. 8, where the pla-

TABLE III. Simulation and MHNC data for $S_N(0,R)$ at $\rho^* = 0.25$.

| R/σ_m	$N = 706$		$N = 2048$	
	Simulation	MHNC	Simulation	MHNC
1.0	0.350 ± 0.002	0.348	0.351 ± 0.003	0.349
1.5	1.045 ± 0.005	1.038	1.049 ± 0.006	1.044
2.0	1.097 ± 0.008	1.082	1.105 ± 0.011	1.099
2.5	1.260 ± 0.012	1.261	1.303 ± 0.019	1.295
3.0	1.336 ± 0.018	1.320	1.367 ± 0.028	1.379
3.5	1.337 ± 0.023	1.346	1.501 ± 0.033	1.441
4.0	1.303 ± 0.028	1.324	1.511 ± 0.054	1.466
4.5	1.279 ± 0.034	1.266	1.523 ± 0.057	1.468
5.0	1.206 ± 0.036	1.172	1.537 ± 0.067	1.450
5.5	1.097 ± 0.039	1.045	1.402 ± 0.071	1.416
6.0	0.892 ± 0.043	0.884	1.361 ± 0.077	1.366
6.5			1.311 ± 0.094	1.303
7.0			1.250 ± 0.098	1.226
7.5			1.27 ± 0.11	1.135
8.0			1.11 ± 0.12	1.030
8.5			0.98 ± 0.12	0.911
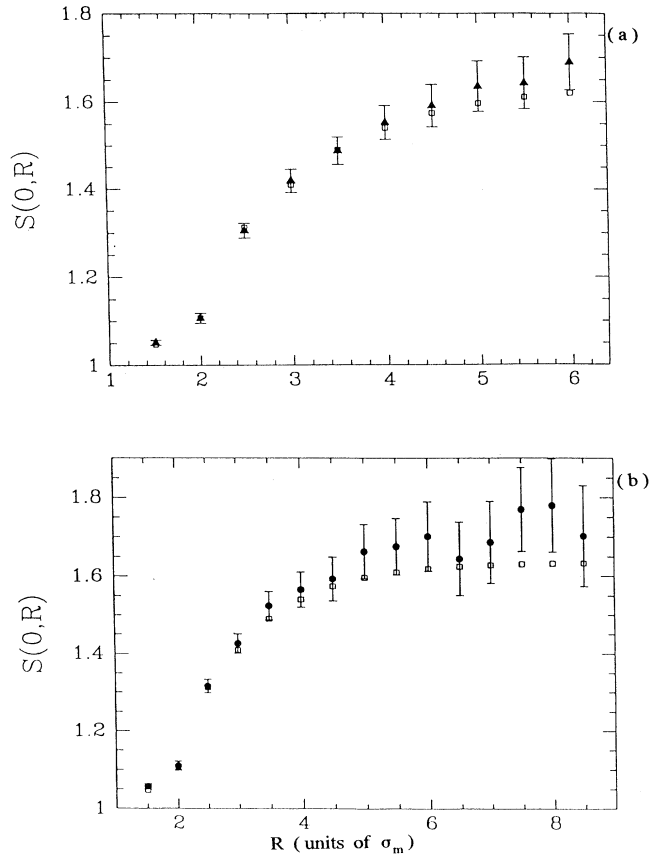
FIG. 8. Same as Fig. 4, but for $\rho^* = 0.25$.

teau level is seen to be consistent with the previous pre-dictions of $S(0)$.

The distribution function $S_N(Q,R)$, calculated via Eq. (6), is shown in Figs. 9 and 10. Finally, using $S(0)$ to correct $S_N(Q,R)$, as in the previous case, leads to $S(Q,R)$. Averaging $S(Q,R)$ over the highest few values of R (for $N=706$, $R/\sigma_m=5.0$, 5.5, and 6.0, and for $N=2048$, $R/\sigma_m=6.5-8.0$) then yields the bulk static structure factor $S(Q)$, which is plotted in Figs. 9 and 10 for $N=706$ and 2048, respectively. Comparing these two figures, it is interesting to observe that for the larger sys-tem the finite-size effects are not only weaker, but also set in at a smaller value of Q.

We conclude this section by justifying our neglect, in the calculation of $S(Q,R)$, of the density derivative term in Eq. (21) of paper I. First, we note that Eqs. (21) and (3a) of I may be rewritten in the form

$$S(Q,R)=S_N(Q,R)+\frac{1}{N}F(Q,R)+O(1/N^2) , \quad (10)$$

where

$$F(Q,R)\equiv\rho S(0)\frac{4}{3}\pi R^3 u(QR)$$

$$+\frac{S(0)}{2}4\pi\rho\int_0^R dr\, r^2\left[\frac{\sin(Qr)}{Qr}\right]$$

$$\times\frac{\partial^2}{\partial\rho^2}\{\rho^2[g(r)-1]\} . \quad (11)$$

We discuss the role of the two contributions to $F(Q,R)$. $F(Q,R)$ is largest at $Q=0$ and at the densities $\rho^*=0.25$ and 0.4 the second contribution in Eq. (11), which de-pends on the density derivative, is only of order of 1% of the first. Therefore this term is completely negligible at $Q=0$ and at small Q, so that the simplified expressions (22) and (23) of paper I can be used in the analysis of the simulation results. At larger Q the second term in Eq. (11) becomes comparable to the first, but in this range of Q the function $F(Q,R)/N$ becomes so small that it can be neglected.

We can also perform another check on the assumptions behind this method. Under the conditions of the simula-
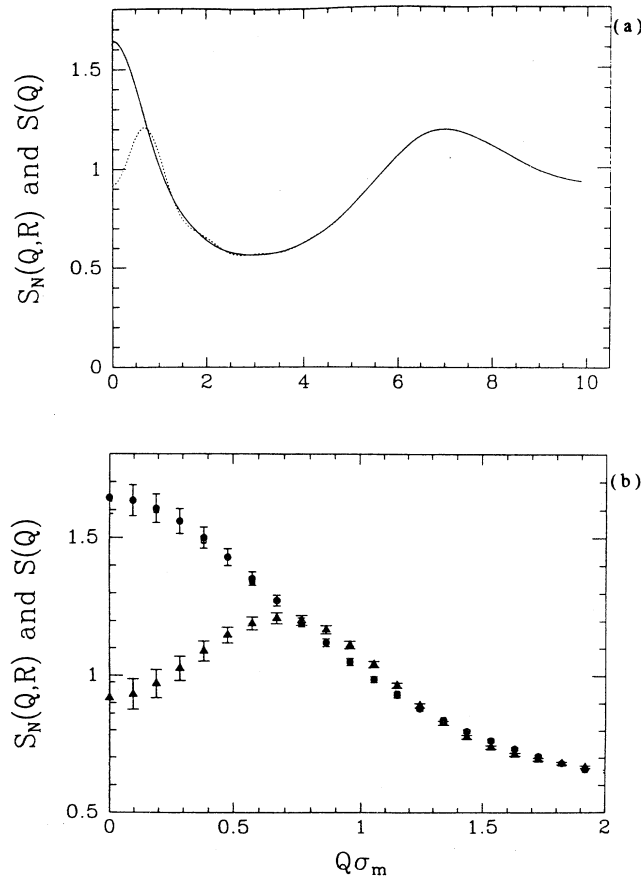

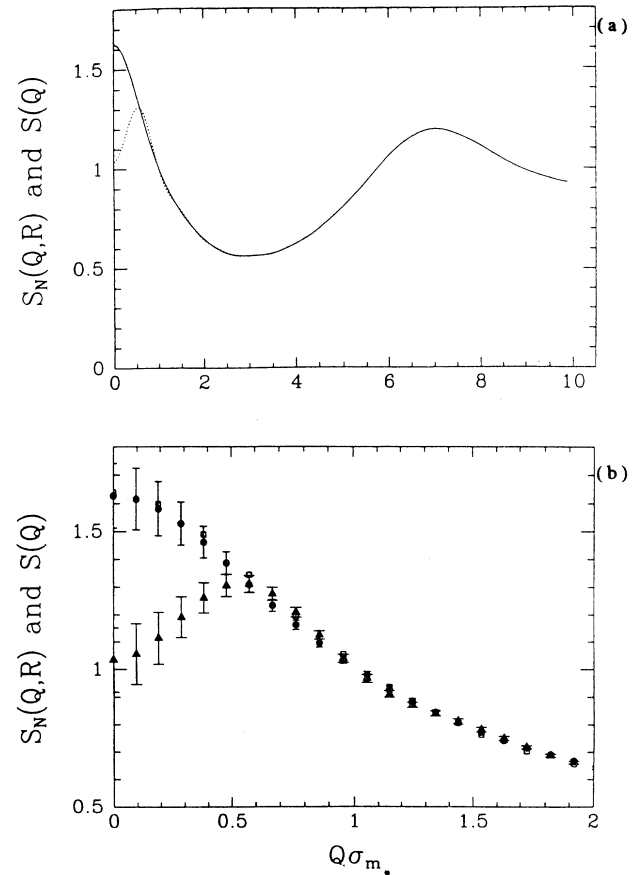FIG. 9. Same as Fig. 6, but for $\rho^*=0.25$ and $R=6$ in units of σ_m.

FIG. 10. Same as Fig. 6, but for $N=2048$, $\rho^*=0.25$, and $R=8$ in units of σ_m.

tion $g_N(r)$ at large r is found to have reached its asymptotic value to within the noise of the simulation. From theory we can compute the quantity $\Delta(Q,R)=S(Q)-S(Q,R)$, where R is half the size of the simulation box or larger. Then $\Delta(Q,R)$ for $R>L/2$ gives the contribution to $S(Q)$ due to $g(r)$ at distances that are not accessible to the simulation. Again the contribution is largest at $Q=0$. In the case of $N=706$, $\Delta(0,L/2)$ at $\rho^*=0.4$ is 10% of the $1/N$ correction, i.e., $F(0,L/2)/N$, and 5% at $\rho^*=0.25$.

A numerical comparison between $F(0,R)$ and the quantity $\frac{4}{3}\pi\rho R^3 S(0)$ shows it to be an excellent approximation to $F(0,R)$ (accurate to within 1%) for $R>4\sigma_m$. Given the size of the $1/N$ correction to $S_N(0,R)$, this means that correlations at distances beyond $L/2$ give a contribution to $S(0)$ that is within the statistical uncertainty of the simulation value and this gives support to the method of analysis of the simulation data.

IV. DISCUSSION

In evaluating the results of the preceding section, we first note that the close agreement between our MD data
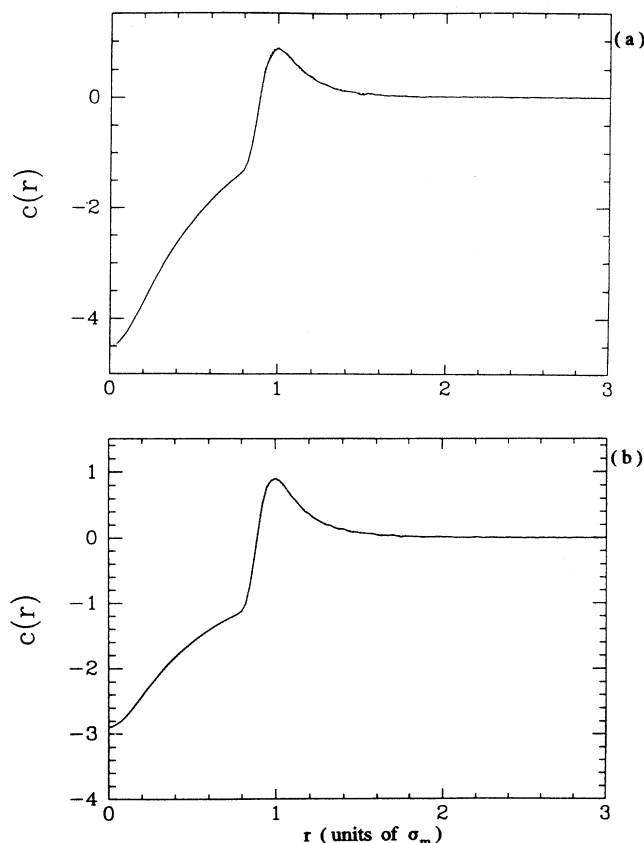

FIG. 11. Bulk direct correlation function $c(r)$ for the same states and notation in Figs. 6 and 9: (a) $N=706$, $\rho^*=0.4$ and (b) $N=706$, $\rho^*=0.25$. Fourier transforming our MD data for $S(Q)$ in the range $0<Q\sigma_m<100$ yields $c(r)$ accurately for r (in units of σ_m) $>2\pi/100$ (solid curves). For comparison, the corresponding MHNC results are indistinguishable from the above result on the scale shown.

and the MHNC predictions for $g_N(r)$, as seen in Figs. 1 and 7, indicates consistency at this level between MD simulation and MHNC theory. Our subsequent predictions of $S(0)$ for the $\rho^*=0.4$ state—0.89 ± 0.05 for the 706-particle simulation and 0.90 ± 0.07 from the 2048-particle simulation—are both in good agreement with the MHNC value of 0.9078. Similarly, for the $\rho^*=0.25$ state, the small- and large-system predictions of $S(0)$—1.65 ± 0.06 and 1.71 ± 0.15, respectively—compare favorably with the MHNC value of 1.6388. Predictions of $S(Q)$, as illustrated in Figs. 6, 9, and 10, are seen to be in excellent agreement with corresponding MHNC results, aside from a slight discrepancy in the range $0.5\leq Q\sigma_m\leq1.5$.

We have demonstrated that our method gives consistent results as N varies between 706 and 2048. There is a relationship, however, between the uncertainties associated with predictions of the method and the system size N, which we now briefly describe. Consider Eq. (23) of paper I, which, as previously noted, holds for R values at which $g_N(R)$ has reached its asymptotic limit. In the $\rho^*=0.4$, $N=706$ case, Fig. 1 shows that *to within the uncertainties of the simulation* $g_N(R)$ is roughly constant for $R/\sigma_m\geq4.0$. Although this statement is rather qualitative, note that for $R/\sigma_m\geq4.0$ an equal number of data points for $g_N(R)$ fall above and below the value $-S(0)/N$, with a majority falling within one standard deviation of $-S(0)/N$. Lengthening the simulation would increase the number of estimates of $g_N(R)$ and of $S_N(0,R)$, which in turn would reduce the uncertainties in these quantities and in the predicted value of $S(0)$. Reducing the uncertainty in $g_N(R)$, however, may expose finer structure in this quantity at larger-R values, thus shifting the asymptotic region to larger values of R. The process of reducing the uncertainty in $S(0)$ is therefore limited for a fixed system size N, since at some point the asymptotic region of $g_N(R)$ is shifted to $R\geq L/2$, rendering the method inapplicable. Put simply, for a given N there is an upper limit to the practical duration of a simulation beyond which it is not worthwhile to continue the simulation. To achieve any further reduction in uncertainty, N must be increased, necessitating simulations of a larger system.

Another reason for determining the complete $S(Q)$ for a bulk fluid from a MD simulation is that we may calculate the direct correlation function $c(r)$ from the Fourier transform of Eq. (A2), for example. In Fig. 11 we show the comparison of the function obtained from MHNC and simulation. Excellent agreement is found, comparable to that seen in Figs. 6, 9, and 10.

V. SUMMARY AND CONCLUSIONS

In this paper we have described in detail the numerical implementation of our method, presented in the preceding paper, for determining the static structure factor $S(Q)$ of a bulk fluid. To demonstrate and test the method, we have carried out a series of molecular dynamics computer simulations of a simple model of krypton atoms interacting via a truncated Aziz pair potential.

Simulation data (particle configurations) for two distinct system sizes N and two thermodynamic states (T^*, ρ^*) were analyzed to extract $S(Q)$, in particular in the challenging low-Q range (*including* $Q = 0$).

Independence of configurations was found to be crucial to the accuracy of the results and was ensured by separating configurations by a suitable time interval. The techniques of including a random element in the interval and of averaging over seeds of the random number generator permitted maximal use of the simulation data. For a given system size and state, the appropriate interval was determined empirically by examining the decay of time correlation function related to the intermediate scattering function. Longer intervals were required for the larger system due to a slower decay of correlations.

Estimation of the compressibility ratio $S(0)$ by three different approaches—from (i) $S_N(0, R)$ via Eq. (23) of paper I, (ii) the plateau level of $S(0, R)$, and (iii) the "difference formula" [Eq. (24) of paper I]—gave consistent results, although the first two approaches were found to be superior to the third. Furthermore, the results for $S(Q)$ were essentially independent of N. Consistency and N independence of the results represent important tests and give support to the three main approximations underlying the method, these being (i) truncation to $O(1/N)$ of the series expansion for $g_N(r)$, (ii) neglect of the density derivative term in the $O(1/N)$ coefficient of the series, and (iii) neglect of implicit size effects.

In comparison with corresponding predictions of MHNC theory, the results for $g_N(r)$ and $S(Q)$ were found to agree usually quite closely. These comparisons may be interpreted either as a test of the $O(1/N)$ finite-size correction formula (assuming the accuracy of MHNC) or as a test of MHNC theory (assuming the accuracy of the correction formula). The slight discrepancy between our results and MHNC predictions near $Q\sigma_m \approx 1$ is admittedly difficult to account for. Whether it reflects an artifact of one of our approximations or signals a slight inaccuracy in the theory is still an open question.

A significant consequence of our ability to calculate $S(Q)$ at all Q from our simulations is that we may compute the direct correlation function $c(r)$. Because it is a central quantity in a number of theories of fluids, more advanced tests of the usual theoretical assumptions will become possible.

It is important to emphasize that the practical utility of our method has been tested for certain thermodynamic states and, for a given state, is limited to a certain minimum system size. In particular, we have chosen states near the critical density but with a temperature sufficiently far from the critical point that $S(0)$ remains of order unity. $S_N(0, R)$ must pass through its maximum as a function of R and this imposes a lower limit on the size of the system and also implies that $R \approx L/2$ falls within the asymptotic tail of $g_N(r)$. Furthermore, the system must be large enough that implicit size effects due to periodic boundary conditions have a negligible effect on the structure.

In the preceding paper we also presented the time-dependent generalization of our method to the intermedi-ate scattering function $I(Q, t)$, which characterizes the dynamics of a system. Application of the method to calculations of $I(Q, t)$ will be the subject of future studies.

APPENDIX: SUMMARY OF THE MHNC INTEGRAL EQUATION SOLUTION

The radial distribution function $g(r)$ of the infinite system can be obtained using the method of the integral equation. The MHNC equation is known to be highly accurate [4] and it reads

$$g(r) = \exp[-\beta\Phi(r) + g(r) - 1 - c(r) + E_{HS}(r; d)] , \quad (A1)$$

where $\Phi(r)$ is the pair interatomic interaction, $c(r)$ is the direct correlation function, which can be expressed in terms of $S(Q)$ via the relation

$$c(Q) = [1 - S^{-1}(Q)]/\rho , \quad (A2)$$

and $E_{HS}(r; d)$ is the bridge function of hard spheres of diameter d and at density ρ. The diameter d is determined by a variational principle, and the relevant equation, the expression of E_{HS}, and details on the numerical method used to solve Eqs. (A1) and (A2) can be found in previous work [4]. In order to make contact with the result of a simulation we may compute the theoretical estimate of $g_N(r)$. From (18), (19a), and (20) of paper I we can write, to eliminate $O(1/N)$,

$$g_N(r) = g(r) - \frac{c_1}{N} , \quad (A3)$$

$$c_1 = S(0) + \frac{S(0)}{2}\frac{\partial^2}{\partial\rho^2}\{\rho^2[g(r) - 1]\} \quad (A4)$$

and from this we obtain $S_N(Q, R)$ via (3b) of paper I. The second density derivative in (A4) is computed by solving the MHNC equation at closely spaced values of the density.

For completeness we write Eqs. (21) and (3a) of paper I in the form

$$S(Q, R) = S_N(Q, R) + \frac{1}{N}F(Q, R) + O(1/N^2) , \quad (A5)$$

where

$$F(Q,R)=\rho S(0)\frac{4}{3}\pi R^3 u(QR)$$
$$+\rho\frac{S(0)}{2}4\pi\int_0^R dr\, r^2\frac{\sin(Qr)}{Qr}$$
$$\times\frac{\partial^2}{\partial\rho^2}\{\rho^2[g(r)-1]\}\ . \qquad \text{(A6)}$$

We have computed $g(r)$, $S(Q)$, $S(Q,R)$, and the correction term $F(Q,R)$ using the MHNC theory with the truncated Aziz potential for krypton under the same conditions as the simulation. The results are presented in the figures and tables and reviewed in Sec. IV.

---

[1] J. J. Salacuse, A. R. Denton, and P. A. Egelstaff, preceding paper, Phys. Rev. E **53**, 2382 (1996).

[2] R. A. Aziz, Mol. Phys. **38**, 177 (1979); in *Inert Gases,* edited by M. J. Klein (Springer-Verlag, Berlin, 1984), Vol. 34, Chap. 2.

[3] J. M. Haile, *Molecular Dynamics Simulation* (Wiley, New York, 1992).

[4] L. Reatto and M. Tau, J. Chem. Phys. **86**, 6474 (1987).