Characterization of solvent clusters in a supercritical LennardJones fluid

Hernan L. Martinez, R. Ravi, and Susan C. Tucker

Citation: The Journal of Chemical Physics **104**, 1067 (1996); doi: 10.1063/1.470762
View online: http://dx.doi.org/10.1063/1.470762
View Table of Contents: http://scitation.aip.org/content/aip/journal/jcp/104/3?ver=pdfcov
Published by the AIP Publishing

Characterization of solvent clusters in a supercritical Lennard-Jones fluid

Hernan L. Martinez, R. Ravi, and Susan C. Tucker
Department of Chemistry, University of California, Davis, California 95616

(Received 1 August 1995; accepted 12 October 1995)

We have developed a new methodology for characterizing the solvent cluster structures which occur in a pure supercritical fluid in its compressible regime. This methodology takes advantage of the time scale separation which exists between collective-cluster and individual-solvent–atom motions in order to classify atoms according to their ''instantaneous'' local environments. The resultant picture is of a fluid having density inhomogeneities on a mesoscopic length scale—i.e., clusters and cavities. Calculation of partial radial distribution functions shows that atoms residing in different density domains have very different equilibrium structural properties, information which is not available from the usual total radial distribution function. For example, for a 2-dimensional Lennard-Jones fluid at a reduced temperature $T_r=1.06$ the nearest neighbor coordination number in a high density domain is 4.2, whereas in a low density domain it is only 1.0. We have also found that, for such clustering fluids that in a finite volume system there is an ensemble independent [within terms of $\mathcal{O}(1/N)$], nonstructural long-range correlation which arises from an excluded volume effect. This long range correlation enables us to determine the average domain size, volume fraction and density for both the high and low density domains. © *1996 American Institute of Physics.* [S0021-9606(96)50803-X]

I. INTRODUCTION

Recent studies of solute reactivity in supercritical fluid (SCF) solvents have shown that these solvents often promote unusual reactivity, with both reaction rates and mechanisms being sensitive to the overall thermodynamics conditions.[1–10] While this unusual reactivity is not well understood, it is clear that it is due in part to the elevated temperatures of supercritical (SC) conditions.[11,12] The most obvious effect of these elevated temperatures is an exponential increase in the rates of all activated processes and the concomitant increase in competitiveness of mechanisms which could be ignored at lower temperatures. A more subtle, but equally critical, effect results from changes in the relative stabilities of reactant, product, and intermediate species of each mechanism that are caused by changes in the fluids' solvating properties in the SC regime.[2,12–16] An SCF's solvent properties are very sensitive to both temperature (T) and pressure (P) in the region of high compressibility in the vicinity of its critical point.[14,17] This sensitivity results primarily from the large changes in solvent density with T and P that accompanies such high compressibility. Note that the bulk of this compressible region of the phase diagram is in general not near enough to the critical point (T_c,ρ_c) for universality and critical phenomena theories to be valid. For example, in the present work we consider states in this region having temperature $T=0.75$ and density $\rho=0.3$ which corresponds to the reduced values ($T_r=T/T_c$) $T_r=1.59$ and ($\rho_r=\rho/\rho_c$) $\rho_r=0.86$ and also the state $T_r=1.06$ ($T=0.5$), $\rho_r=0.86$. Because small changes in the T or P can alter a fluid's solvent properties dramatically in this region, it is here that the greatest control over chemical reactivity is expected, and it is on this region that much of the recent work on reactivity in SCF's has focused.

The goal of the present work is to gain insight into the solvent properties of a pure SCF in the compressible region, in order to provide a backdrop against which the solvation of reactive solutes may be understood. A few recent studies have focused upon the characterization of the microscopic structure of pure SC water in this regime.[18–22] These studies have considered only short ranged density correlations on a molecular (microscopic) length scale, in order to understand the hydrogen bonding structure. Yet, it is well known that as the critical point is approached, the correlation length of the order parameter—in this case density—approaches infinity, as does the compressibility. This result implies that in the compressible region of interest, the correlation length, though finite, may still be much longer than the molecular length scale. Thus in the region of interest, density correlations (or fluctuations) are expected to occur on a mesoscopic length scale—much shorter than macroscopic but longer than microscopic. Previous studies of solvent structure in the SC regime have not considered these mesoscopic density fluctuations. We address the issue of these fluctuations here.

Mesoscopic length scale density fluctuations correspond to regions of high and low density in the pure SCF solvent, i.e., they signal the presence of solvent–solvent clusters and cavities. Such solvent–solvent clustering is the pure solvent corollary of the well established long-range[23] solvent–solute clustering observed in dilute attractive SCF solutions,[2,18,24–32] in which a large number of solvent molecules ''cluster'' around a single solute molecule. Such solvent–solute clustering has recently been shown to be an important factor in chemical reactivity in SCF solvents.[2,12,15,33] A solid understanding of the tendency of the solvent to form clusters even in the absence of the solute will be necessary for the development of theoretical treatment of solvent–solute clustering.[23] We have thus performed simulation studies of mesoscopic length scale density fluctuations (clusters) in a pure SCF. In Sec. II, we describe the fluid studied and our

simulation techniques. In Sec. III, we address the issues of equilibrium versus nonequilibrium density fluctuations (and equilibration criteria in SCF systems), the separation of time scales present in systems with mesoscopic density fluctuations, and partial radial distribution functions. These functions will enable us to examine more closely the properties of the solvent clusters and cavities. We also briefly address the temperature dependence of these density fluctuations; the density dependency will be addressed in future work.

II. METHODS

We performed molecular dynamics simulations on a 2-dimensional Lennard-Jones (LJ) fluid with a truncated, unshifted potential,

$$V(r)=\begin{cases} 4\epsilon\left(\left(\frac{\sigma}{r}\right)^{12}-\left(\frac{\sigma}{r}\right)^{6}\right) & r<2.5\sigma \\ 0 & r>2.5\sigma \end{cases}. \quad (1)$$

The critical point for this potential has previously been only roughly determined,[34] with $T_c \approx 0.5 \pm 0.02$ and $\rho_c \approx 0.3$. This estimate of ρ_c is too imprecise to be useful. However, the critical point has been determined much more accurately, for the shifted version of this potential, $V_s(r) = V(r) - V(r=2.5\sigma)$, with $T_c = 0.47 \pm 0.01$ and $\rho_c = 0.35 \pm 0.01$.[35,36] Although the critical point parameters of the shifted potential are not directly transferable to the case of the unshifted potential,[36,37] we use these values when computing reduced variables for the unshifted potential studied here. Lennard-Jones reduced units are used throughout, with $\sigma=1$ and $\epsilon=1$. An effective NVT ensemble of $N=1152$ atoms with periodic boundary conditions was considered. Integration was performed using the Verlet algorithm[38] with time steps of 4.0×10^{-3}. The desired temperature was achieved during equilibration of an NVE ensemble through simple velocity scaling. Because maintaining a constant temperature is important near the critical point, velocity scaling was performed every 33rd time step throughout the simulation in order to damp the temperature fluctuations and mimic an NVT ensemble. The simulations were started from an FCC lattice and were equilibrated for 80 000 time steps, well beyond the time required to achieve an equilibrium cluster structure. We describe below the function we use to determine the attainment of this equilibrium condition. Because solvent structural reorganizations involve long time scale components as the critical point is approached (see Sec. III B), sufficient sampling of solvent configuration space for the evaluation of ensemble averages can be problematic (see Sec. III C 2). Thus the results presented here were averaged for 750 000 time steps in order to ensure convergence. Finally, all radial distribution functions were evaluated by standard methods[39] using a bin length of 2.5×10^{-2} and a sampling frequency of 40 time steps.

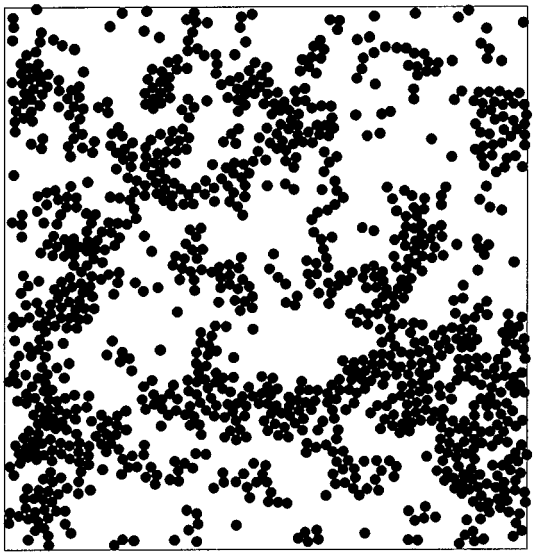

FIG. 1. Snapshot of a Lennard-Jones SCF at $T_r=1.06$ and $\rho_r=0.86$ in two dimensions after 80 000 time steps.

III. RESULTS AND DISCUSSION

A. Equilibrium density fluctuations

Although solvent–solvent clustering in SCF's has not been discussed to any degree in the literature, some evidence for the existence of such clusters is available, primarily from theoretical work. We briefly review these previous observations in order to motivate the present study of solvent–solvent clustering in SCF's. Perhaps the most convincing evidence for the existence of possibly transient solvent–solvent clusters comes from computer simulation snapshots of SCF's.[34,40,41] For example, solvent–solvent clusters are evident in Gao's simulation of the benzene dimer in SC water.[40] Retention of the pure solvent structure in the presence of the solute is expected in this case because benzene is only a weakly attractive solute, and thus does not greatly alter the pure solvent structure. Solvent–solvent clusters are also clearly visible in snapshots of the 2-dimensional LJ fluid studied by Binder and co-workers[34] and by ourselves, see Fig. 1.

The observation of solvent–solvent clusters at single molecular dynamics time steps or single Monte Carlo configurations provides evidence only that these clusters exist as large length scale nonequilibrium fluctuations. These observations can say nothing about whether such density variations are an equilibrium structural feature of these fluids. Evidence that these structures are equilibrium entities, in the same sense that the microscopic structure of a liquid represented by peaks in the radial distribution function is an equilibrium structure, comes from two previous works. In the first of these, Debenedetti and co-workers[30,42] have studied a 3-dimensional LJ fluid in which they use an energetic criterion to define physical clusters.[43] They evaluated the equilibrium distribution of cluster sizes under supercritical conditions, $T_r=1.037$ and $\rho_r=1.0$, and observed an extremely broad distribution. While small clusters of size 1–2 atoms

make up 50% of the distribution, clusters of more than 5 atoms make up to 36% of the distribution and clusters greater than 60 atoms make up 9% of the distribution. As will become clear below, atoms in the small size clusters must correspond to our low density atoms. In addition, the probability distribution for finding any given atom in a cluster of a certain size would be much more heavily weighted towards large clusters than is the cluster distribution itself, because the large clusters contain many atoms. In checking to ensure that the simulation cell was longer than the range of correlations, Debenedetti and co-workers took the decay of the short range microscopic structure in the usual radial distribution function to indicate this range. Yet, as we discuss in Sec. III C 2, an effective density correlation can extend beyond the range of the microscopic structure, and such a correlation could affect the observed cluster sizes. Hence, it is possible that the cluster distribution observed in this previews work is distorted towards smaller clusters by the limited box size, especially given the proximity of the simulated state point to the critical point.

In the second work indicating the presence of equilibrium solvent–solvent clusters, Rovere *et al.* studied a 2-dimensional LJ fluid.[34] They used a block density distribution function to examine finite size effects on simulated critical parameters. Their block distribution function gives the equilibrium probability of finding n_l particles in a square box of area l^2, i.e., $P(n_l)$. These authors found that, for certain block sizes, the distribution $P(n_l)$ is not peaked at the average particle number $\langle n_l \rangle = l^2\rho$, where ρ is the bulk density, but instead exhibits a distinct peak at number densities below the bulk average and a shoulder at number densities above the average. This behavior is observed *above* the critical point, indicating that even in this macroscopically one-phase region the equilibrium distribution contains higher than bulk density and lower than bulk density regions on length scales of order l. The block distribution function of Rovere *et al.* is intuitively appealing because observation of the density behavior on different length scales (block sizes l^2) is reminiscent of the Kadanoff transformations of renormalization group theory.[44,45] The disadvantage of this block method is that it imposes a lattice onto a nonlattice fluid, thus weighting all the results according to this lattice filter. Such filtering complicates interpretation, especially if the clusters are similar in size to the lattice area, l^2. Additionally, the observed distributions are sensitive to the block size, with the two peaks coalescing as the block size exceeds some average cluster size. In principle this sensitivity could be used to extract this ''average'' cluster size, but again one would obtain a result filtered through a square box when the clusters can be of any shape. Thus we take a different, atom-centered approach to study these equilibrium density fluctuations, which avoids the need to impose a lattice structure on the fluid.

In our atom-centered approach, we use a time-dependent classification of atoms according to the instantaneous local density around each atom. As will become clear below, this classification provides us with the necessary information for studying both cluster formation (equilibration) and equilibrium cluster properties. The criterion we use to classify the local environment of each atom is as follows: First, we evaluate n_{ck}, the integer number of atoms, including the test atom, in a circle of radius r_{ck} around that test atom. This value of n_{ck} is then compared to n_0 the average number of atoms expected in a circle of this radius according to the bulk density ρ, i.e., $n_0 = \pi r_{ck}^2 \rho$. If the radius r_{ck} is chosen such that n_0 is an integer, then the local number density around a test atom n_{ck} may be lower, higher or exactly equal to the average number density n_0, yielding classification of the atom as low, high or average, respectively. This criterion is excessively strict in that small fluctuations in an average atom's local solvent structure will cause the classification of that atom to be switched from average to high or low even if only one atom moves across the r_{ck} boundary. In order to relax this criterion, we classify test atoms having a range of number densities, given by $n_0 \pm \delta$ (δ an integer), as average atoms. Thus the classification of each atom is given according to the following rule:

$$\begin{aligned} &\text{high} && \text{if} && n_{ck} > n_0 + \delta, \\ &\text{average} && \text{if} && n_0 - \delta \leqslant n_{ck} \leqslant n_0 + \delta, \\ &\text{low} && \text{if} && n_{ck} < n_0 - \delta. \end{aligned} \tag{2}$$

The fluctuation we have allowed in our calculations is 30% of n_0, such that the ''average'' window centered at n_0 is of size $2\delta + 1 = 0.3 n_0$. For example, for $\rho = 0.3$ ($\rho_r = 0.86$) and $r_{ck} = 3.09$, $n_0 = 9$ atoms, so $2\delta + 1 = 3$ atoms giving $\delta = 1$. The inclusion of fluctuations makes our results less sensitive to the exact choice of r_{ck}. Choosing r_{ck} such that n_0 is an integer guarantees that, with integer counting of n_{ck}, the average value n_0 is realizable and falls in the center of the fluctuation region.

A configuration snapshot for the conditions $T_r = 1.59$, $\rho_r = 0.86$ with $r_{ck} = 3.09$, 3.99, and 4.72 ($n_0 = 9$, 15, and 21) is shown in Figs. 2(a), 2(b), and 2(c) respectively. With each atom labeled according to its designation from Eq. (2). First note the presence of distinct clusters containing high density atoms and cavity regions containing low density atoms. The average density atoms appear primarily along the cluster–cavity borders [see especially Fig. 2(a)]. These average or border atoms are separated from the high density domain atoms because they are radially asymmetric even at short ranges, feeling a high density domain on one side and a low density domain on the other side. Separating these border atoms out from the high density atoms turns out to be important when the equilibrium properties of the high and low density atoms are studied. This is because the spatial extent of the high density regions is not large and edge effects due to the border atoms would make a significant contribution to the properties of the cluster atoms if these atoms were included as part of the high density region. This point will be discussed in more detail when we discuss the partial radial distribution functions.

In order to study the equilibrium structure of an SCF's density fluctuations, it is necessary to determine that an equilibrium configuration of these mesoscopic length scale fluctuations (clusters) has been reached prior to the evaluation of

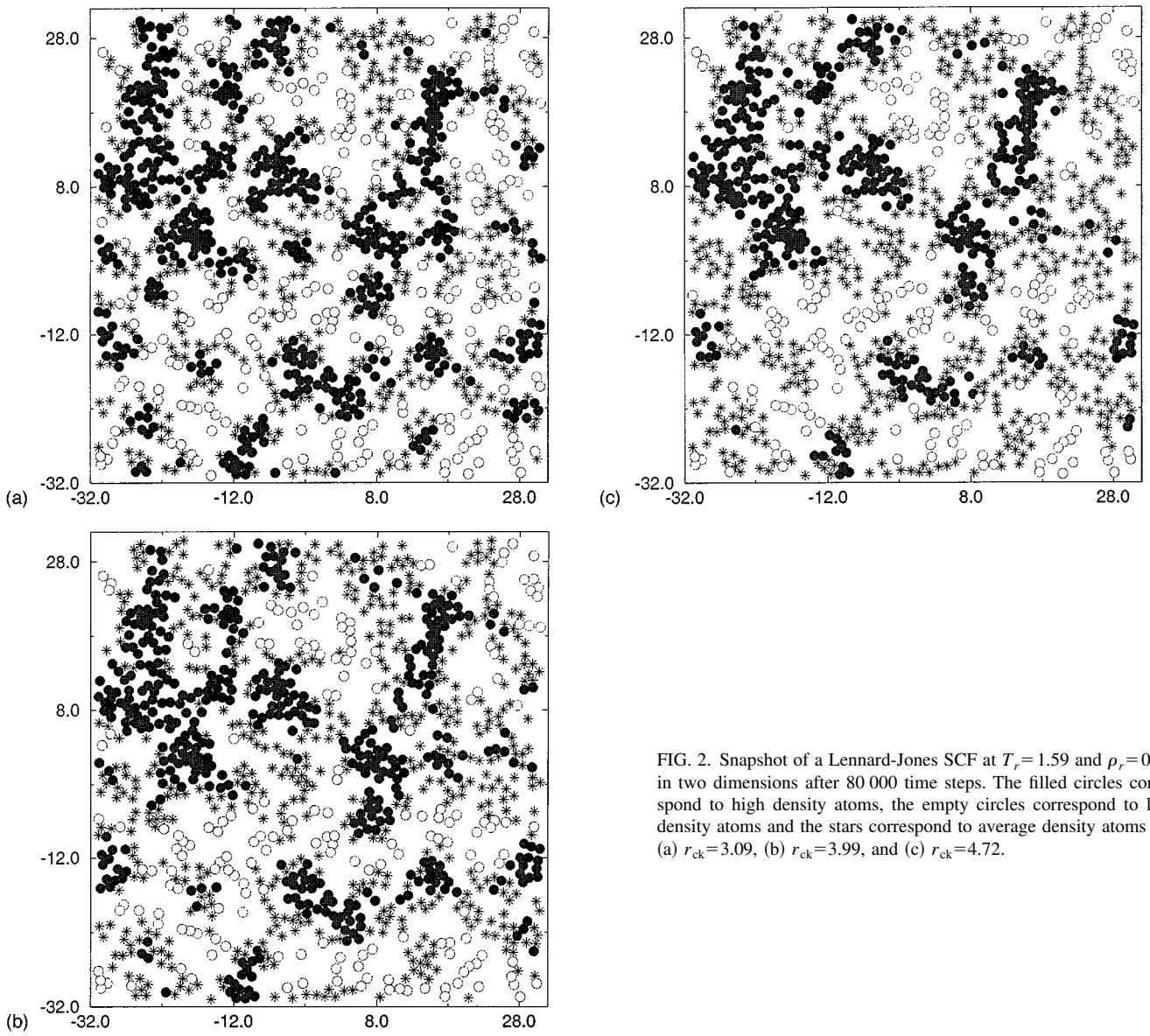

FIG. 2. Snapshot of a Lennard-Jones SCF at $T_r=1.59$ and $\rho_r=0.86$ in two dimensions after 80 000 time steps. The filled circles correspond to high density atoms, the empty circles correspond to low density atoms and the stars correspond to average density atoms for (a) $r_{ck}=3.09$, (b) $r_{ck}=3.99$, and (c) $r_{ck}=4.72$.

equilibrium averages. The usual functions for establishing equilibrium are not useful indicators of cluster formation. For example, consider the translational order parameter λ[38,46] for monitoring the breakup of an initial crystal lattice structure, i.e.,

$$\lambda = 0.5[\lambda_x + \lambda_y], \tag{3}$$

where

$$\lambda_x = \frac{1}{N}\sum_i^N \cos\left(\frac{4\pi x_i}{a}\right) \tag{4}$$

and a is the length of the FCC unit cell. This order parameter gives only microscopic, atomic length-scale information about the structure of the system and cannot be used to identify the growth of clusters to some equilibrium configuration (see below).

In order to monitor the cluster equilibration, we develop a new set of functions based on the atom classifications in Eq. (2). Specifically, we define a function $n_x(t)$, x=hi (high), lo (low), or av (average), which counts the total number of atoms of type x at time t, i.e., $n_x(t)$ counts the total number of atoms in *all* existing cluster (or cavity or border) regions at time t

$$n_x(t) = \sum_i^N \Theta_x^i(t), \tag{5}$$

with

$$\Theta_x^i(t) = \begin{cases} 1 & \text{if atom } i \text{ is of type } x \text{ at } t \\ 0 & \text{if atom } i \text{ is not of type } x \text{ at } t, \end{cases} \tag{6}$$

and N the total number of atoms. Simply by checking the density of each atom's local environment we are able to determine the total number of atoms that are "in" clusters, without ever providing an explicit, and necessarily arbitrary,

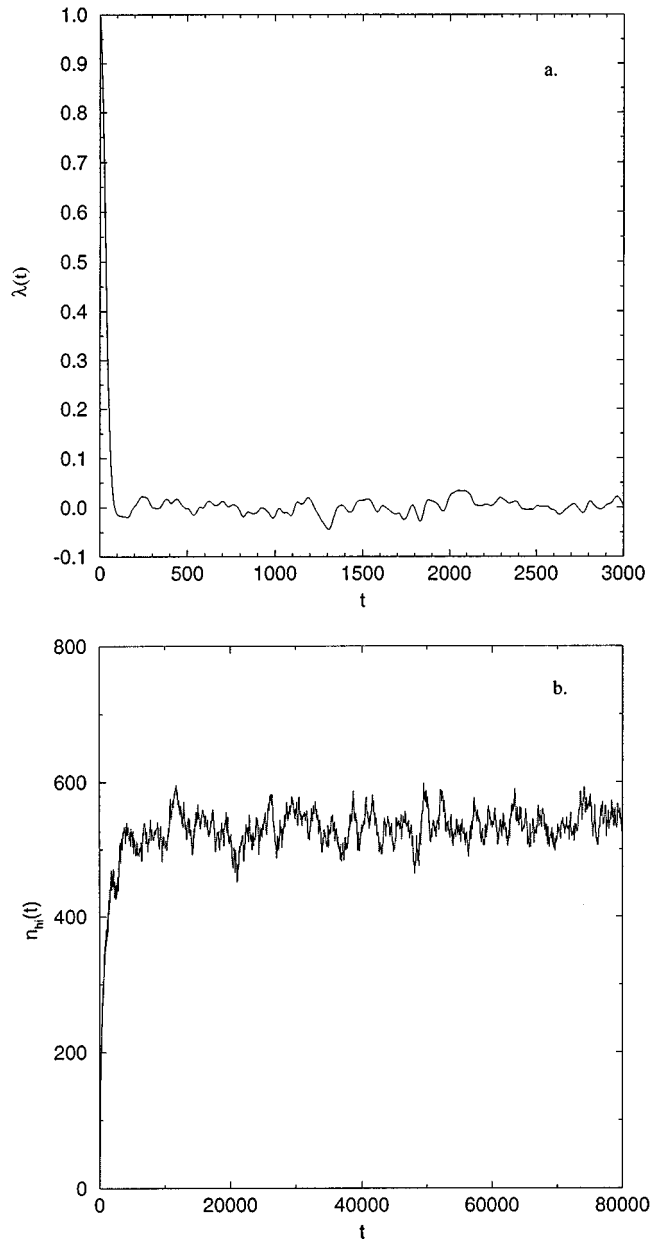

FIG. 3. (a) λ function for a two dimensional supercritical Lennard-Jones fluid at $T_r=1.59$ and $\rho_r=0.86$. (b) $n_{\rm hi}(t)$ for the same fluid.

definition of a cluster. Admittedly, our method is still somewhat arbitrary, as the definition of the local environment depends on $r_{\rm ck}$ and δ.

The functions $n_x(t)$ have the advantage that they provide simply one number per time step, making them easy to use for determining the attainment of equilibrium. In Fig. 3 the usual λ function is shown along with $n_{\rm hi}(t)$ for $T_r=1.59$, $\rho_r=0.86$. The λ function decays to zero, showing the breakup of the initial lattice structure, at around 100 time steps. Similarly, the $n_{\rm hi}$ function begins at zero and rises to a plateau value; however, the rise time for this function is significantly longer than the time for the decay of the λ function. Since $n_{\rm hi}(t)$ provides information about the creation of high density regions, i.e., cluster growth, the initial transient rise can be equated with the growth of high density regions—clusters. The plateau in $n_{\rm hi}(t)$ thus indicates equilibration of cluster growth. From Fig. 3(b). It is clear that the cluster formation takes nearly 10 000 time steps. This much longer time scale for equilibration of cluster formation relative to the breakup of the initial crystal lattice is consistent with the longer length scale and collective nature of the solvent–solvent clusters, and illustrates why the usual microscopic measures of equilibration fail to denote the attainment of equilibrium in these systems.

In Fig. 4, the three functions $n_x(t)$ are shown for the same conditions as are shown in Fig. 2. While the absolute value of the plateau for $n_{\rm hi}$ varies with the choice of cutoff radius, the time scale for cluster formation, i.e., the time to reach this plateau, does not vary noticeably.[47] Also note that the function $n_{\rm hi}(t)$ appears to be the best indicator of the slowest equilibration time. Finally, we have evaluated the functions $n_x(t)$ for a nonclustering fluid. Specifically, we considered the same 2-dimensional Lennard-Jones fluid, but at a liquid density ($T_r=0.957$, $\rho_r=2.04$). For this case there is no rise in $n_{\rm hi}(t)$; instead, the number of average density atoms, $n_{\rm av}(t)$, remains large for all times, accounting for 97% of the total number of atoms. Thus the high and low density functions, $n_{\rm hi}(t)$ and $n_{\rm lo}(t)$, clearly distinguish between cluster formation and the maintenance of a more uniform density on mesoscopic length scales.

Analysis of the plateau value reached by $n_{\rm hi}$ can provide an indication of the behavior of the cluster structure at equilibrium. While $n_{\rm hi}$, the number of atoms in high density regions, fluctuates, the average fluctuation is not large, i.e., $\Delta n_{\rm hi}=(\langle n_{\rm hi}^2\rangle-\langle n_{\rm hi}\rangle^2)^{1/2}=17.9$, which is 3.4% of $\langle n_{\rm hi}\rangle$, for the case presented in Fig. 2 with $r_{\rm ck}=3.09$. The maintenance of a large number of high density atoms indicates both that (i) high density regions (clusters) are present in the equilibrium structure (although it says nothing about the spatial extent of these high density regions) and (ii) once the clusters have formed, cluster dynamics do not involve substantial loss of cluster structure at any instant of time. This conclusion is consistent with sequences of configurational snapshots which suggest that clusters rearrange by large collective motions, e.g., fracturing, rather than by complete decimation or evaporative mechanisms. These results are also consistent with Debenedetti's conclusion that in pure SCF solvent's clusters are continually being reformed.[42]

B. Time scales

It can be inferred from Fig. 3 that there are two time scales of solvent motion for an SCF in its compressible region. As discussed in the previous section, the time scale for cluster formation is much longer than the time scale for the breakup of the initial lattice. Both of these time scales may also be observed in the subsequent equilibrium dynamics represented by the fluctuations in the plateau value of $n_{\rm hi}(t)$ (Fig. 4). The fluctuations in $n_{\rm hi}(t)$ occur on both a short time scale, $t_m\sim 50$ time steps, which corresponds to molecular motion of single solvent atoms[48] and a longer time scale, $t_c\gtrsim 1000$ time steps, which corresponds to collective solvent

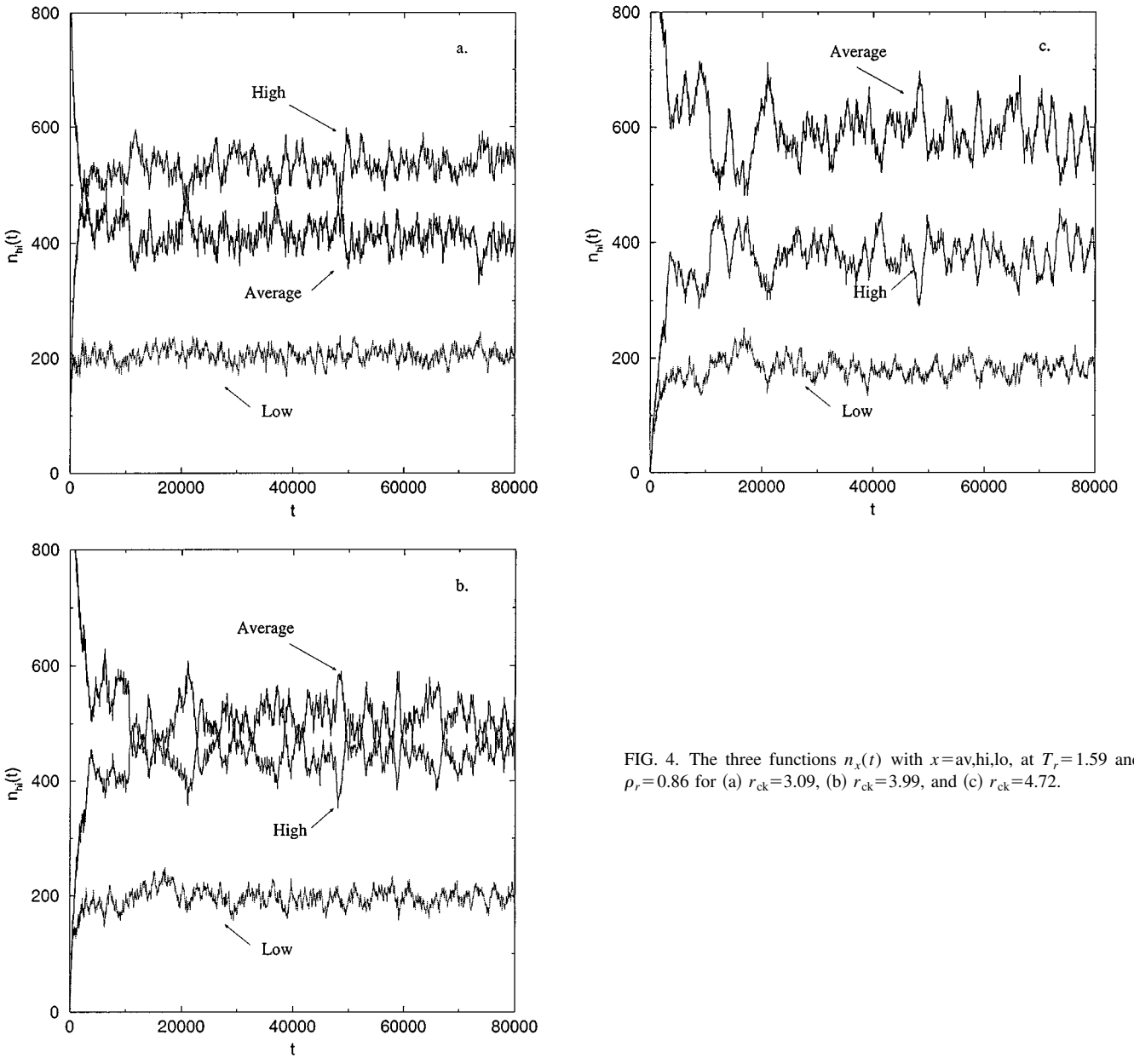

FIG. 4. The three functions $n_x(t)$ with x=av,hi,lo, at $T_r=1.59$ and $\rho_r=0.86$ for (a) $r_{ck}=3.09$, (b) $r_{ck}=3.99$, and (c) $r_{ck}=4.72$.

motions associated with the cluster dynamics. Additionally the time scale for molecular motion, t_m, appears in the fluctuations of the λ function around zero, while the fluctuations on the time scale t_c do not appear in this function. Since the λ function does not pick up the cluster structure, this absence of t_c provides additional evidence that t_c is a collective solvent, cluster dynamics time scale, whereas t_m is a molecular (atomic) time scale.

The time scale separation between molecular and cluster dynamics allows a description of the system in terms of high and low density domains (clusters, cavities) which interconvert only rarely compared to the time scale for molecular motion. Thus on time scales shorter than t_c, the cluster dynamics time scale, it is reasonable to expect that atoms in high (low) density domains will remain in high (low) density domains. It follows that on these time scales solvent atoms may be classified according to the type of domain in which they reside [Eq. (2) for $t \ll t_c$], i.e., they can be thought of as distinct species. Considering high and low density domain atoms to be separate species is much the same as treating the reactants (A) and products (B) of a unimolecular rearrangement reaction ($A \leftrightarrow B$) having a double well potential of mean force as distinct species.[49] By this analogy, we are *not* suggesting that the thermodynamic free energy versus density for an SCF fluid above its critical point has a double well character. The thermodynamic free energy is a macroscopic quantity, and on a macroscopic length scale these systems appear homogeneous, with a uniform density, consistent with the single (though perhaps still very flat) free energy minimum expected above the critical point. What we are suggesting is that the ''mesoscopic'' free energies of activation along some set of collective cluster fragmentation coordinates may

have double well character, consistent with both the slow cluster rearrangement time scale and the fluid's inhomogeneous appearance on a mesoscopic "domain" length scale (cf. Fig. 2).

As suggested above, atoms in different types of domains may be thought of as different species. It follows that these different species atoms may exhibit distinct domain-specific properties. Equilibrium averages of these domain-specific properties may be calculated on times of order $t_m \ll t < t_c$. Over times longer than t_c the domains interconvert and atoms will change their identities. After sufficiently long times (many t_c), each atom will have sampled each domain, and all atoms will be equivalent. The fraction of time spent by each atom in each domain is given by $\langle n'_{\rm hi}\rangle/N$ and $\langle n_{\rm lo}\rangle/N$, for the high and low density domains, respectively.[50] The properties of each of these equivalent atoms will then be simply a sum of the atom properties associated with each domain, weighted according to the fraction of time each atom spends in each domain. Taking such a long time average washes out available information about the individual domains. Thus domain specific averages ($t < t_c$) are required if one wishes to study the details of the mesoscopic density fluctuations, as we do here.

C. Partial radial distribution functions

In order to characterize the details of the solvent structure in the low and high density domains we take advantage of the time scale separation between atomic and cluster dynamics and treat atoms in different density domains as separate "species." We can then evaluate partial radial distribution functions[51] between these different domain-type atoms, i.e., we evaluate $g_{xx'}(r)$, where x, x'=av, hi, or lo type atoms [Eq. (2)]. These partial radial distribution functions provide information about the distribution of like (i.e., same domain type, $x'=x$) and nonlike (i.e., different domain type, $x' \neq x$) atoms around a given atom of type x.

In order to evaluate the partial radial distribution functions, we would like to average over short times $t < t_c$ for which most atoms retain their domain identities. However, our system size ($N=1152$) is much too small to sample a significant portion of the distribution of cluster structures in one sampling. Consequently, we must sample the different cluster structures over time by averaging for times $t \gg t_c$. In order to evaluate domain-specific equilibrium averages on this long time scale over which atom identities interconvert, we recheck the identity of each atom at every time step at which the radial distribution functions are evaluated. Thus at each sampling, the local instantaneous environment for each atom pair is checked to determine to which "species" each atom belongs and, consequently, towards which partial distribution functions the pair will contribute. Specifically, $g_{xx'}(r) = \langle \rho^{(2)}_{xx'}\rangle/\langle \rho_{x'}\rangle$, where $\rho^{(2)}_{xx'}$ is the usual conditional probability of finding an x'-type atom at a distance r given that an x-type atom is at the origin, and $\langle \rho_{x'}\rangle$ is the bulk density for atom type x' averaged over the length of the simulation. This corresponds to the usual normalization choice for an open system.[51] Because atoms may change their identity over time, we effectively have an open system for $n_{\rm hi}$, $n_{\rm lo}$, and $n_{\rm av}$, but subjected to the constraint that the total number of atoms (N) remains fixed. Note that since the numbers of $x \to x'$ and $x' \to x$ pairs at any interparticle distance must be identical, $g_{xx'} = g_{x'x}$.[52] The partial distribution functions provide a more direct view of the equilibrium domain structure than does the usual homogeneous fluid radial distribution function, $g(r)$. Specifically, the usual, or total, $g(r)$ treats all atoms equivalently and thus provides only a weighted average of the high density, low density, and average density distributions. Similarly, the domain specific functions provide more information about the long-range (mesoscopic length scale) correlations than does the total $g(r)$.

1. Short range correlations

The partial radial distribution functions, $g_{\rm hi-hi}(r)$ and $g_{\rm lo-lo}(r)$, are plotted in Fig. 5 for the conditions $T_r=1.59$, $\rho_r=0.86$, and $r_{\rm ck}=3.09$. The short range microscopic structure characteristic of $g_{\rm hi-hi}(r)$ can be observed in Fig. 5(a), and that for $g_{\rm lo-lo}(r)$ can be observed in Fig. 5(c). The details of the long range behavior can be seen more clearly in Figs. 5(b) and 5(d), respectively, in which the vertical axis has been expanded. By looking at the structure over short range one can see that the peak positions of $g_{\rm hi-hi}(r)$ and $g_{\rm lo-lo}(r)$ are nearly identical. This similarity suggests that the average nearest-neighbor and next-nearest-neighbor distances are the same in the high and low density domains, despite the density differences observed between these two regions. This result is explained by a configurational snapshot as shown in Fig. 2. There it can be seen that atoms in the low density regions are not spread out uniformly, but appear in clusters (or chains) of two's and three's, similar to the ones observed in gas phase.[53–55] Thus in both domains the nearest neighbor distance is controlled primarily by the intermolecular potential rather than by the local density, and similar peak positions result. The small 2–3 atoms cluster behavior of the low density domain is also confirmed by the very short range of the structured part of $g_{\rm lo-lo}$ in which only 2 distinct neighbor peaks are observed, as compared to the more extended structure of $g_{\rm hi-hi}$ in which 3 distinct neighbor peaks are observed. In addition to a comparison of peak positions for the low and high density domain atoms, it is also informative to make a comparison of the peak heights. However, because the partial distribution function g_{xx} is normalized according to the "bulk" density of species x, $g_{\rm hi-hi}$ and $g_{\rm lo-lo}$ have very different normalization constants ($\langle n_{\rm hi}\rangle=529.5$ whereas $\langle n_{\rm lo}\rangle=204.9$). As a result, a direct comparison of peak heights is misleading, and instead we compare the average number of nearest neighbors, $\langle n^{(i)}_{x'}\rangle$, in the two domains. The average number of x' atoms at a distance r_i around an x atom is given by

$$\langle n^{(i)}_{x'}\rangle \equiv 2\pi\langle \rho_{x'}\rangle \int_0^{r_i} dr\, r g_{xx'}(r), \tag{7}$$

where r_i is taken as the position of the ith minimum in $g_{xx'}$. Here, we are interested in the case $x'=x$. For the high density domain, the average number of atoms in the first

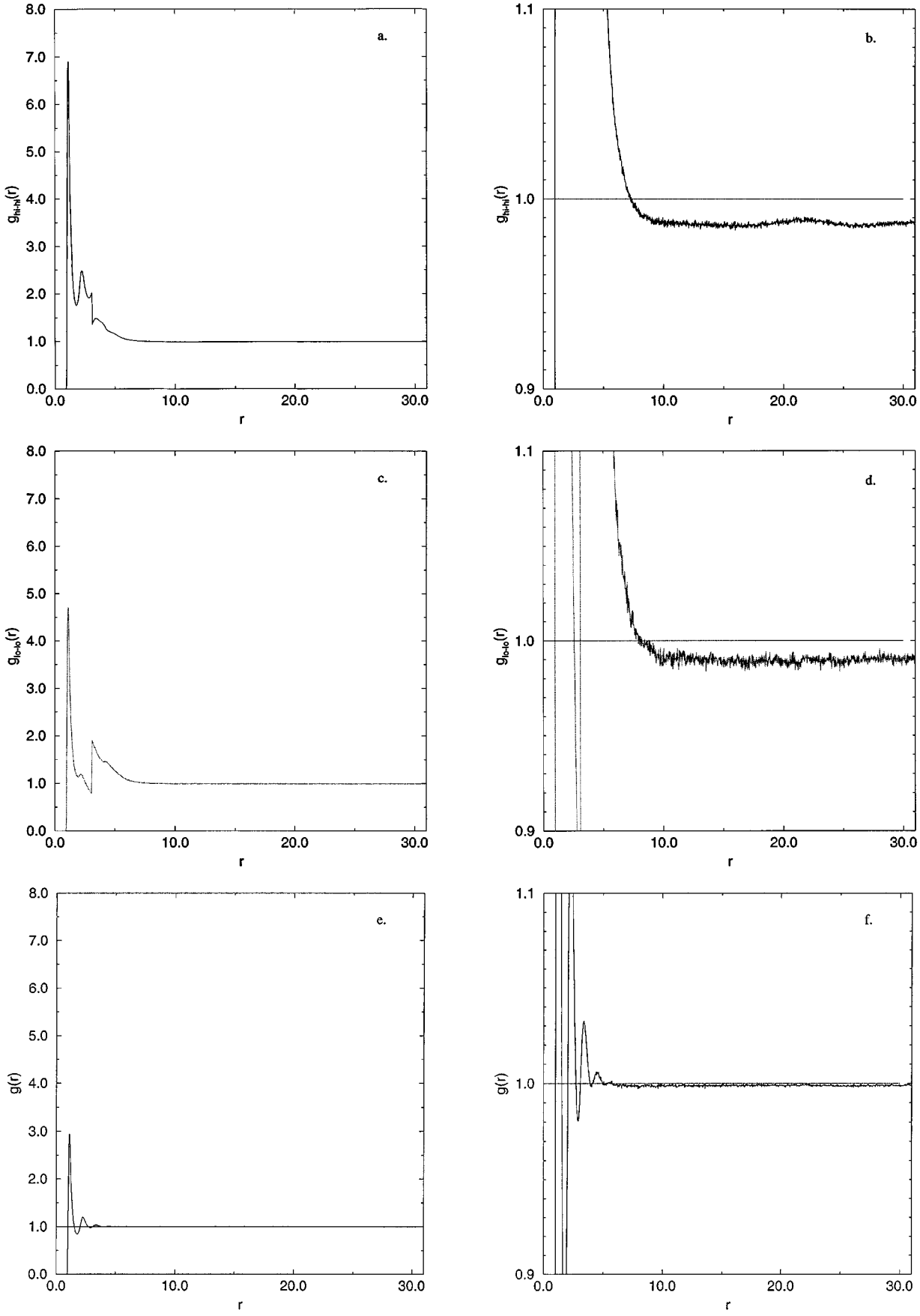

FIG. 5. For the conditions $T_r = 1.59$ and $\rho_r = 0.86$ and compared under the same scale: (a) Partial radial distribution function $g_{\text{hi-hi}}$, (b) long range part of $g_{\text{hi-hi}}$, (c) partial radial distribution function $g_{\text{lo-lo}}$, (d) long range part of $g_{\text{lo-lo}}$, (e) total radial distribution function $g(r)$, and (f) long range part of $g(r)$.

solvation shell is $\langle n_{\rm hi}^{(1)}\rangle \approx 3.9$, while for the low density domain the first solvation shell contains on average only $\langle n_{\rm lo}^{(1)}\rangle \approx 0.86$ atoms. The contrast within the first plus second shell is greater, with $\langle n_{\rm hi}^{(2)}\rangle \approx 7.6$ and $\langle n_{\rm lo}^{(2)}\rangle \approx 1.2$, indicating that low density atoms only rarely have a next nearest neighbor. We note that for the average particles $\langle n_{\rm av}^{(1)}\rangle \approx 1.6$ and $\langle n_{\rm av}^{(2)}\rangle \approx 2.2$ which fall in between the high and low values as expected for atoms on the edges of the clusters ($g_{\rm av-av}$ is not shown). The difference between domains is even greater in the case of the lower temperature $T_r = 1.06$. In this case, the average number of atoms in the first solvation shell for the high density domain, is $\langle n_{\rm hi}^{(1)}\rangle \approx 4.2$, while for the low density domain the first solvation shell contains on average only $\langle n_{\rm lo}^{(1)}\rangle \approx 1.0$. The contrast within the first plus second shell is now $\langle n_{\rm hi}^{(2)}\rangle \approx 11.0$ and $\langle n_{\rm lo}^{(2)}\rangle \approx 1.9$. The average domain coordination numbers again fall between the high and low values, with $\langle n_{\rm av}^{(1)}\rangle \approx 1.2$ and $\langle n_{\rm av}^{(2)}\rangle \approx 2.5$. The discontinuity apparent in both $g_{\rm hi-hi}(r)$ and $g_{\rm lo-lo}(r)$ at the cutoff radius $r_{\rm ck}$ appears because the density is constrained for radii less than $r_{\rm ck}$ but not for radii greater than $r_{\rm ck}$. A study of the related discontinuity at $r = r_{\rm ck}$ for the distribution function $g_{\rm lo-(lo+hi+av)}$ and $g_{\rm hi-(lo+hi+av)}$ (not shown) reveals that if the density around an atom is constrained to be below (above) the average value at small radii ($r < r_{\rm ck}$), it will remain so over the entire range of the short-range same-domain correlations. This behavior suggests that the domain (or cluster) dimension is as large or larger than the range of these microscopic correlations. It also confirms the equilibrium nature of the low and high density domains.

Both the short- and long-ranged parts of the total radial distribution function ($g(r)$) for $T_r = 1.59$ and $\rho_r = 0.86$ are pictured for comparison in Figs. 5(e) and 5(f). Since the differences in the hi–hi and lo–lo partial distribution functions confirm the existence of distinct domain "species" having different distribution properties, it is appropriate to interpret the total radial distribution function as the weighted average of all of the partial distribution functions, i.e.,

$$g(r) = \frac{1}{N^2}\left[\sum_i^3 \langle n_i\rangle^2 g_{ii}(r) + 2\sum_{i<j}^3 \langle n_i\rangle\langle n_j\rangle g_{ij}(r)\right], \quad (8)$$

where the indices correspond to the species low, average and high. In order to consider the effect of this averaging on the observed total $g(r)$, consider first the short range part of $g(r)$. Taking into account both the weight factors and the peak height, we find that the hi–hi, hi–av, and av–av functions dominate the short range part of the total $g(r)$. For example, at the first peak maximum the hi–hi function contribution is 2.5 times the hi–av and av–av contributions, but 9 times the lo–lo contribution and nearly 80 times the lo–hi contribution. The retention of sharp peaks in the short range part of the total distribution function [Fig. 5(e)] despite this averaging is a result of the similarity of the peak positions in both the high and low density domains and for the border atoms ($g_{\rm av-av}$, not shown). It is interesting to note that recent simulations[20,21,27] show the short-range part of the total radial distribution function for SC water in the compressible region to exhibit a broad, washed out peak structure in which only the nearest neighbor peak is clearly visible. A similar result was observed in a 3-dimensional LJ fluid at $T_r = 1.037$ and $\rho_r = 1.0$.[42] It will be interesting to see if, in either case, the observed broadening is due in part to the averaging of high- and low-density-domain radial distribution functions having differing peak positions, or whether it is due entirely to high temperature entropic effects as has been suggested. In addition, since the partial distribution functions indicate that high-density domain atoms have a more complete nearest neighbor shell than do low-density domain atoms (while the border atoms have an intermediate value), it follows that the number of nearest neighbors evaluated from the total radial distribution function should also be interpreted as an average of high and low density domain values. Hence, the coordination number evaluated from the total $g(r)$ does not provide a complete structural picture of this type of SCF.

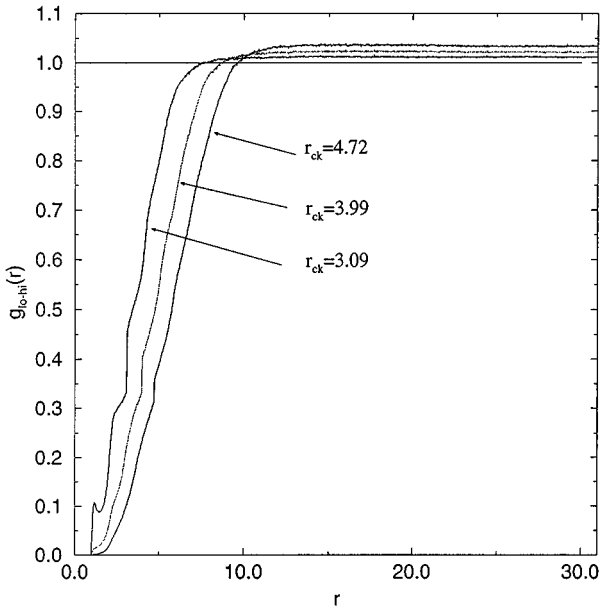

FIG. 6. The crossed lo–hi crossed partial distribution function as calculated using three different check radii, $r_{\rm ck} = 3.09$, 3.99, and 4.70.

In Fig. 6 we have plotted the "crossed" partial radial distribution function, $g_{\rm lo-hi}$, for the same values of $T_r = 1.59$, $\rho_r = 0.86$ and $r_{\rm ck} = 3.09$. The short range behavior of these curves show how the number atoms of the opposite species increases with increasing distance until the value corresponding to the bulk density is reached. For example, if the origin is placed on a low density domain atom, then no high density domain atoms will be observed until one has moved far enough away to have crossed the boundary into a high density domain. However, the "tagged" low density domain atom will not necessarily be situated at the center of the low density domain, and it is thus possible to find high and low density atoms at fairly close distances. Still, the average number of high density atoms within a low density atom's first solvation shell is only $\langle n_{\rm hi}^{(1)}\rangle \approx 0.04$, and within the second solvation shell is only $\langle n_{\rm hi}^{(2)}\rangle \approx 0.24$, where we used Eq. (7) with x = lo and x' = hi. The variation in distances between noncentrally located atoms in different domains causes a

gradual rise in the high–low partial radial distribution function with increasing distance instead of a sharp rise at the distance of the domain boundary. In addition, large fluctuations in the domain sizes contribute to this blurring of the domain boundary. At distances which are sufficiently large that, on average, one has crossed from the low to high density domain, the crossed partial radial distribution function rises above unity and high density domain atoms are found with higher than average likelihood (see Sec. II C 2 below). Thus we can very roughly equate the position at which $g_{\mathrm{lo-hi}}$ crosses from anti- to positively correlated with the average domain size, i.e., the average cluster size. In fact, on average, only $\langle n_{\mathrm{hi}}^{(\mathrm{cross})}\rangle = 18.0$ high density atoms are found at radii less that the cross-over radius, i.e., within the low density "domain," which is significantly less than in the case of a homogeneous fluid where the number of atoms would be 53. For $T_r = 1.59$, the cross-over takes place at $\sim 7.5\sigma$, which is also where the cross-overs of $g_{\mathrm{hi-hi}}(r)$ and $g_{\mathrm{lo-lo}}(r)$ from positively to anticorrelated takes place for the same conditions.

In Fig. 6 we also show the lo–hi crossed partial distribution function as calculated using two other different check radii, $r_{\mathrm{ck}} = 3.99$ and 4.72. Note that the discontinuity in $g_{\mathrm{lo-hi}}$ occurs at $r = r_{\mathrm{ck}}$, for the same reason as discussed in the $g_{\mathrm{hi-hi}}$ and $g_{\mathrm{lo-lo}}$ cases. From Fig. 6 it can be seen that as the check radius is increased, the number of high density atoms found close to the tagged atom decreases, and the structure in the short-ranged part of $g_{\mathrm{lo-hi}}$ is almost completely lost. In addition, the cross point moves to larger distances with increasing check radii, indicating an apparent increase in the low-density domain size. This check radius dependence, while non-negligible, can be understood in terms of the check radius dependence of the average density or border atoms. As the check radius is increased, more atoms are classified as average density, at the expense of the number of high density atoms. For example, for a check radius of $r_{\mathrm{ck}} = 3.09$, $\langle n_{\mathrm{hi}}\rangle = 529.5$ and $\langle n_{\mathrm{av}}\rangle = 417.7$, whereas for $r_{\mathrm{ck}} = 4.70$, $\langle n_{\mathrm{hi}}\rangle = 374.0$ and $\langle n_{\mathrm{av}}\rangle = 595.4$. Recall that the average atoms tend to reside on the border between high and low density domains because the circle drawn out around these atoms by the check radius encompasses both high and low density regions. If the check radius is increased, some atoms previously containing only high density regions within their check circles may now contain both high and low density regions within their larger check circles. These high density atoms will then be reclassified as average density atoms. As a result, the width of the band of border atoms between the domains will increase, and it will be much less likely for a low density atom to have nearby high density atoms, in agreement with the trends observed in $g_{\mathrm{lo-hi}}$. Also, as the width of the band of border atoms is increased, one must go further away from a low density atom to reach the high density domain, consistent with the apparent increase in the size of the low-density domain with increasing check radius.

2. Long range "correlations"

An important consideration as one approaches the critical point of a fluid from above is the rate at which the correlation length ξ approaches infinity. If the fluid is very close to its critical point, the correlation length will diverge according to universal critical exponents.[44] Further above the critical point, where the correlation may be large but still smaller than macroscopic, details of the molecular interactions may become important, and universality probably does not hold. In this region it is of interest to be able to determine the correlation length from simulation.

In discussions of critical phenomena, the correlation length is often equated with the "largest cluster size" in the fluid, and it is thus relevant to the present discussion of domain (or cluster) sizes. However, the rigorous definition of the correlation length of a fluid is the exponential decay constant for the larger r behavior of the total radial distribution function in the grand canonical ensemble,[56] $\bar{g}(r)$,

$$\bar{g}(r) \sim e^{-r/\xi}. \tag{9}$$

Thus although ξ is closely related to cluster size, it can not generally be directly equated to either a large or an average cluster size. Here, we extend the idea of correlation length to partial correlation lengths, $\xi_{xx'}$, defined by the partial distribution functions in an open system, i.e.,

$$\bar{g}_{xx'}^{\infty}(r) \sim e^{-r/\xi_{xx'}}. \tag{10}$$

A superscript ∞ has been added to $\bar{g}_{xx'}$ to denote that, as we shall show below, this definition will only be valid for systems of macroscopic volume.

All of the partial radial distribution functions we have considered exhibit a significant deviation from unity over very long ranges—in fact these deviations persist for the full half box length of 30.9σ [see Figs. 5(b), 5(d), 5(f), and 6]. Although these functions all exhibit "real" long range correlations [since $g_{xx'}(r) \neq 1$ for large r], the behavior of these long range correlations is unusual in that no exponential decay is observed. Instead, each of the distribution functions rises (or falls) to a plateau value. The lack of structure in these long range correlations implies that there are correlated density fluctuations at long range which are *structurally* uncorrelated with respect to the central particle. In order to make this statement more precise, we examine the behavior of an idealized, canonical two-component system—described by fixed N_{hi} and fixed N_{lo}—in a box of volume V. In this system, we assume that the atoms are completely uncorrelated for distances $r > r_+$. For distances $r < r_+$, we assume that there is only an average density fluctuation, i.e., our idealized fluid has no microscopic structure (or equivalently, we consider the microscopic structure to have been averaged out). Finally, we assume clustering behavior. Thus for $r < r_+$ around a low atom we assume that there will be no high atoms present. The crossed partial distribution function for this fluid will then be 0 for $r < r_+$, where r_+ is defined as the point at which the probability of finding a high atom is equal to the bulk density of the high atoms, ρ_{hi}. For $r > r_+$, the (uncorrelated) probability of finding a high atom is simply given by the density of high atoms in this region. However, since all of the high atoms fall outside of the region $r < r_+$, i.e., in a volume of size $V - V_+$ where $V_+ = \pi r_+^2$ in

two dimensions, this density is $N_{\rm hi}/(V-V_+)>\rho_{\rm hi}$. Hence, the crossed distribution function for this system in a canonical ensemble (superscript c) is

$$g^c_{\rm lo-hi}(r)=\begin{cases} 0 & \text{if } r<r_+ \\ \dfrac{N_{\rm hi}}{V-V_+}\left(\dfrac{V}{N_{\rm hi}}\right)=\dfrac{V}{V-V_+}>1 & \text{if } r>r_+ \end{cases}. \tag{11}$$

The self-correlation functions are found by realizing that in a clustering fluid the average density of species x surrounding an atom of type x at distances $r<r_+$, $\overline{\rho}_{x+}$,[57] will be greater than the "expected" density of species x, ρ_x. For a canonical ensemble it follows that, for $r<r_+$ one finds the open system result $\rho_x\int_0^{r_+} g^c_{xx}(r)\,d\mathbf{r}=\overline{\rho}_{x+}V_+$ while for $r>r_+$ one has $\rho_x\int_V g^c_{xx}(r)\,d\mathbf{r}=N_x-1$. These two expressions can be solved for the constants[58] $g^c_{xx}(r<r_+)$ and $g^c_{xx}(r>r_+)$ to find

$$g^c_{xx}(r)=\begin{cases} \alpha_x & \text{if } r<r_+ \\ \left(\dfrac{V-\alpha_xV_+}{V-V_+}\right)-\left(\dfrac{V}{V-V_+}\right)\dfrac{1}{N_x} & \text{if } r>r_+ \end{cases}, \tag{12}$$

where $\alpha_x\equiv\overline{\rho}_+/\rho_x$ and $x=$hi or lo. Note that in the limit of no clustering, $\alpha_x\to 1$ and $g_{xx}(r>r_+)\to 1-[V/(V-V_+)]\times(1/N_x)$, which is just the usual fixed N_x result of $1-(1/N_x)$ for an uncorrelated system, except that by treating $r<r_+$ as an open system, we have constrained the canonical $1/N_x$ correlation to be taken up over the volume $V-V_+$ rather than over the full volume; hence the factor $[V/(V-V_+)]$.

Our simplifying assumption that no high density atoms are present at distances $r<r_+$ away from a low density atom is easily relaxed by defining an average density of high atoms found in this region, $\overline{\rho}_{\rm lo-hi+}$. One then finds

$$g^c_{\rm lo-hi}(r)=\begin{cases} \alpha_{\rm lo-hi} & \text{if } r<r_+ \\ \left(\dfrac{V-\alpha_{\rm lo-hi}V_+}{V-V_+}\right) & \text{if } r>r_+ \end{cases}, \tag{13}$$

where $\alpha_{\rm lo-hi}\equiv\overline{\rho}_{\rm lo-hi+}/\rho_{\rm hi}$. For $T_r=1.59$, $\alpha_{\rm lo-hi}$ is nonnegligible, and we use Eq. (13) instead of Eq. (11) in the remaining analysis.

The present calculations were not performed in an ensemble with $N_{\rm hi}$ and $N_{\rm lo}$ fixed, however, but were performed in a constrained grand canonical ensemble in which $N_{\rm hi}$ and $N_{\rm lo}$ may vary subject to the constrain $N=N_{\rm hi}+N_{\rm lo}$.[59] The canonical results for g^c above, Eqs. (12) and (13), can be easily extended to the distribution function for this constrained case, $\tilde{g}$, by using that

$$\tilde{g}_{xx}=\frac{1}{\bar{N}_x^2}\langle N_x^2 g^c_{xx}\rangle \tag{14}$$

and

$$\tilde{g}_{xx'}=\frac{1}{\bar{N}_x(N-\bar{N}_x)}\langle N_x(N-N_x)g_{xx}\rangle, \tag{15}$$

and that $\langle\delta N^2\rangle=\bar{N}_x(N-\bar{N}_x)/N$ as a result of the constraint. One finds for the long range part of interest

$$\tilde{g}_{xx'}(r>r_+)=\left(\frac{V-\bar{\alpha}_{xx'}V_+}{V-V_+}\right)-\left(\frac{V}{V-V_+}\right)\frac{1}{N}, \tag{16}$$

where $\bar{\alpha}_{xx'}\equiv\bar{\rho}_{xx'_+}/\bar{\rho}_{x'}$ or when $x=x'$, $\bar{\alpha}_{xx}\equiv\bar{\rho}_{x+}/\bar{\rho}_x$.[60] Considering first the low–high crossed partial distribution function, one sees that the observed long range correlation is an excluded volume effect that results even when $N_{\rm hi}$ is allowed to vary. For $N=1152$, $1/N$ is too small to observe and can be neglected. This leaves $(V-\bar{\alpha}_{xx'}V_+)/(V-V_+)$ as the observed plateau value. For $r_{\rm ck}=3.09$, Fig. 6, we take $r_+=7.5\sigma$, the cross point of $\tilde{g}_{\rm lo-hi}$, and we find $\overline{\alpha}_{\rm lo-hi}=(\langle n_{\rm hi}^{\rm cross}\rangle/V+)(V/\langle n_{\rm hi}\rangle)=0.739$ (see Sec. III C 1) and $(V-\bar{\alpha}_{xx'}V_+)/(V-V_+)=1.012$, in excellent agreement with the observed plateau value of 1.012. Also the variation in the plateau values with check radii follows directly from the associated variation in the cross points with check radii. (For $r_{\rm ck}=3.97$, $r_+=8.74$ which predicts a plateau value of 1.021 and for $r_{\rm ck}=4.72$, $r_+=9.61$ which predicts a plateau value of 1.032; these predicted plateau values are in excellent agreement with the observed values of 1.021 and 1.033 respectively, see Fig. 6.) Finally, note that as the system size $V\to\infty$ for a fixed cluster size, V_+, the plateau value will approach unity.

Now consider the self-partial distribution functions $\tilde{g}_{xx}$. The value of $\overline{\alpha}_x$, which gives the average density of domains of type x relative to $\overline{\rho}_x$, is the single unknown. This value should agree with the integration of g_{xx} out to r_+. Again neglecting the $1/N$ term, we use the observed plateau value of 0.987 for $\tilde{g}_{\rm hi-hi}$ and the observed cross point of 7.5σ [Fig. 5(b)] to solve for $\overline{\alpha}_{\rm hi}$. We find $\overline{\alpha}_{\rm hi}=1.27$, which gives $\bar{\rho}_{\rm hi+}=0.18$ so $\overline{\rho}_{\rm hi+}/\rho_c=0.52$. This value of $\overline{\rho}_{\rm hi+}$ agrees with the one obtained by integrating $g_{\rm hi-hi}$ up to r_+ and then dividing by the volume V_+. Accounting for the also average and low density atoms residing in "high" density domains, we find a total average density in this domain of $\rho=0.32$, or $\rho/\rho_c=0.91$. Similarly, for the low density domains we find from $\tilde{g}_{\rm lo-lo}$ that $\overline{\alpha}_{\rm lo}=1.23$, $\overline{\rho}_{\rm lo+}=0.066$ so $\overline{\rho}_{\rm lo+}/\rho_c=0.19$. Finally we can also find the volume fractions of high and low density domains from $V_x^f=(\bar{\alpha}_{x'}-\bar{\alpha}_{xx'})/(\bar{\alpha}_{x'}\bar{\alpha}_x-\bar{\alpha}_{xx'}^2)$. We find $V_{\rm hi}=0.48$ and $V_{\rm lo}=0.52$.

While we have successfully explained the origin of the apparent long-range correlations in our simulations, we cannot discuss correlation lengths unless we extend our results to the grand canonical ensemble. This is easily done by using Eqs. (14) and (15) with $N-N_x$ replaced by $N_{x'}$ and realizing that $\langle\delta N_x^2\rangle=\langle N_x\rangle$ for the uncorrelated long-range results. One finds that $\bar{g}_{xx'}=\lim_{N\to\infty}\tilde{g}_{xx'}$ for both $x'=x$ and $x'\neq x$, as might be expected, i.e.,

$$\bar{g}_{xx'}(r>r_+)=(V-\bar{\alpha}_{xx'}V_+)/(V-V_+). \tag{17}$$

What is important about this result is that it proves that the excluded volume correlation which gives rise to the long range plateau is a result of the clustering behavior at short range and is retained in the open system result (i.e., it is *not* a fixed N or fixed N_x effect). If the cluster size, V_+, is much

smaller than macroscopic, then as the total volume of the system is taken to the macroscopic, thermodynamic limit, the plateau value will become unity and the correlation length, $\xi_{xx'}$, will be given by the decay of $\bar{g}^{\infty}_{xx'}$ to its cross point r_+. If the simulated finite volume ensemble is sufficiently large that $r<r_+$ behaves like the open system result, then the cross point will be independent of system size. Thus, we take the cross points in our simulations as a rough measure of the true correlation lengths. Since we find all of the partial correlation lengths to be equal in a given system, we equate them to the overall correlation length, ξ. The most accurate function for determination of this domain size is the crossed partial radial distribution function, see Fig. 6.

The nonunit plateau values do represent real correlations in finite systems, and hence, if the system size is made too small these correlations could effect the observed cluster sizes. Specifically, such effects are expected if V is not sufficiently large, compared to V_+ such that the area V_+ behaves like an open system. Such variation in the observed cluster sizes with simulation cell size may be more difficult to ascertain from the total $g(r)$ than from the partials, because the domain size is not easily read off from the total $g(r)$ [Figs. 5(e) and 5(f)].

Specifically, the plateau value observed for the total $g^c(r)$ is just the fix N result of $1-1/N$ and $\bar{g}(r)\to 1$. The lack of a distinct cross point in $g(r)$ makes it more difficult to assess the domain size and its variation with box size. Additionally, from Figs. 5(e) and 5(f) it also appears that fitting $g(r)$ to an exponential at large r, in order to find ξ, would be problematic. Finally, from Eqs. (8) and (17) one finds (again neglecting the role of the average atoms)

$$\bar{g}(r>r_+)=\frac{V-\bar{\alpha}_{\rm tot}V_+}{V-V_+}, \tag{18}$$

where

$$\bar{\alpha}_{\rm tot}\equiv\left(\frac{\bar{N}_{\rm hi}}{N}\right)^2\bar{\alpha}_{\rm hi}+\left(\frac{\bar{N}_{\rm lo}}{N}\right)^2\bar{\alpha}_{\rm lo}+2\left(\frac{\bar{N}_{\rm hi}\bar{N}_{\rm lo}}{N^2}\right)\bar{\alpha}_{\rm lo-hi+}=1, \tag{19}$$

and Eq. (19) provides a constraint on the domain densities.

From the preceding analysis it is clear that the equilibrium structural correlations around a tagged atom do not extend beyond the range of that atom's domain. This lack of structure probably results from both the high compressibility of the low density domains which are largely vacuum—and the broad distribution of domain sizes and shapes. One might expect the domain–domain correlations to give one smeared-out nearest-neighbor peak on the domain length scale, such as is observed in gas phase correlation functions for attractive particles. However, when two self-similar domains (e.g., clusters) come together at the "nearest-neighbor" distance, they combine to form one large domain. Hence, the nearest-neighbor peak appears as part of the same domain, self-correlation component of the partial radial distribution functions. We note that if the variations in cluster structures (over times $t\gg t_c$) are not adequately sampled, one will observe structural correlations between domains, i.e., one observes

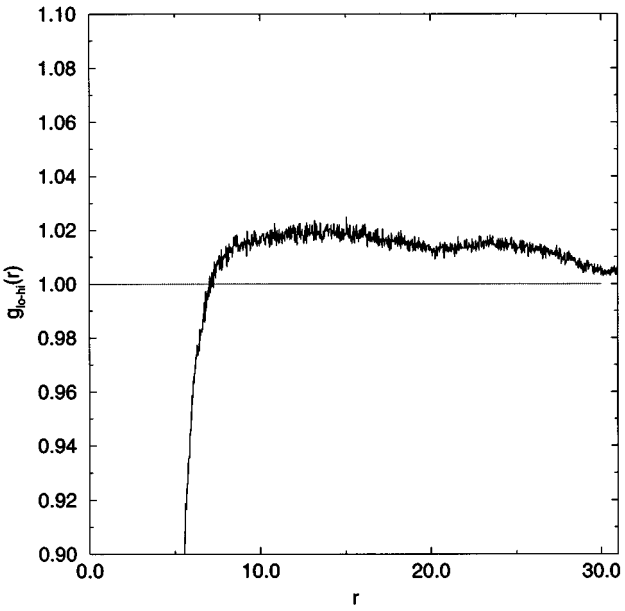

FIG. 7. Partial radial distribution function $g_{\rm lo-hi}$ after 220 000 time steps for $T_r=1.59$ and $\rho=0.86$, $r_c k=3.09$.

nonequilibrium "instantaneous" domain–domain correlations. We note that for $T_r=1.59$, such nonequilibrium structure was still observable in the partial radial distribution functions after 220 000 time steps following equilibration (Fig. 7). In particular, note that the dip in the long-range correlation occurs at $\sim 14\sigma\approx 2r_+$, or one domain diameter, beyond the cross point.

In addition to $T_r=1.59$, we have also studied $T_r=1.06$ at the same reduced density. As expected, the cross point for the partial radial distribution functions moves to a larger distance at the lower temperature, consistent with an increased cluster size. However, we have been unable to obtain a converged result at this temperature, either for the value of the cross point or for the long range tails. Because the converged $T_r=1.59$ calculations took 17 days of CPU time on an SGI R8000, and because a longer time scale is expected to accompany the larger clusters at the lower temperatures, we are in the process of implementing our code on a parallel machine in order to be able to converge the lower temperature results. In future work we will address both the temperature and density dependence of the cluster structure of supercritical fluids.

IV. CONCLUSIONS

We have developed a new methodology for characterizing the solvent cluster structures which occur in pure SCF's. For such clustering fluids, the usual simulation equilibration tests fail because they do not detect the completion of cluster growth. We have defined a new function, based on the classification of atoms according to their local densities, which predicts the onset of equilibrium in these systems. Specifically, at every time step each atom is classified as residing in a high, low or average density domain. We are able to iden-

tify from this function two characteristic time scales for motion in these fluids—a fast molecular time scale and a much slower cluster rearrangement time scale. This separation of time scales means that atoms residing in different domains interconvert only rarely compared with their molecular motion, and can thus be treated as separate ''species.'' It is then possible to use partial radial distribution functions for these different domain-type atoms to determine domain-specific structural properties in the clustering fluid. For example, in a 2-dimensional Lennard-Jones fluid at $T_r = 1.59$ and $\rho_r = 0.86$ we have found that the average coordination number is 3.2 in a high density domain, but only 0.86 in a low density domain. For $T_r = 1.06$ and the same density, these domain-specific numbers are 4.2 and 1.0, respectively. We have also found that the structural correlations, while somewhat longer than the range of the microscopic structure—i.e., the usual nearest and next-nearest neighbor, etc., solvation shell peaks in the total radial distribution function—do not extend beyond the range of the domain in which the tagged atom resides. However, in clustering fluids there will be a non-structural long-range correlation which results from an excluded-volume effect. We have shown that this excluded-volume correlation will become negligible in the macroscopic, thermodynamics limit, but that it will be present in open systems if they are of finite size. In a finite size system, we are also able to determine the average density of both the low and high density domains from this long-range correlation in the self partial distribution functions. These long-range tails also provide a measure of the fraction of the total volume which is encompassed by low and by high density domains. Additionally, we have shown that the low-high crossed partial distribution function provides the clearest illustration of the average cluster (domain) size, and that this information will generally be more difficult to ascertain from the total radial distribution function. We have discussed the relationship between the observed measure of the domain sizes and the correlation length of the fluid. In future work we will use the present methodology to address questions of the temperature, density and dimensionality dependence of the cluster behavior in supercritical fluids.

ACKNOWLEDGMENTS

This material is based on work supported by the National Science Foundation under Grant No. CHE-9307679. S.C.T. acknowledges a NSF Young Investigator Award.

[1] B. Subramaniam and M. A. McHugh, Ind. Eng. Chem. Res. **25**, 1 (1986).
[2] K. P. Johnston and C. Haynes, AIChE. J. **33**, 2017 (1987).
[3] *Supercritical Fluids*, edited by T. G. Squires and M. E. Paulaitis, ACS Symposium Series 329 (ACS, Washington, DC, 1987).
[4] *Supercritical Fluid Science and Technology*, edited by K. P. Johnston and J. M. L. Penninger, ACS Symposium Series 406 (ACS, Washington, DC, 1989).
[5] M. T. Klein, Y. G. Mentha, and L. A. Torry, Ind. Eng. Chem. Res. **31**, 182 (1992).
[6] X. Xu, C. P. De Almeida, and M. J. Antal, Jr., Ind. Eng. Chem. Res. **30**, 1478 (1991).
[7] X. Xu and M. J. Antal, Jr., AIChE. J. **40**, 1529 (1994).
[8] T. G. Squires, S. G. Venier, and T. Aida, Fluid Phase Equil. **10**, 261 (1983).
[9] S. Kim and K. P. Johnston, Chem. Eng. Commun. **63**, 49 (1988).
[10] W. S. Mok and M. J. Antal, Jr., J. Org. Chem. **54**, 4596 (1989).
[11] R. Narayan and M. J. Antal, Jr., J. Am. Chem. Soc. **112**, 1927 (1990).
[12] S. C. Tucker and E. M. Gibbons, in *Structure and Reactivity in Aqueous Solution*, ACS Symposium Series, edited by D. G. Truhlar and C. J. Cramer (ACS, Washington, DC, 1994).
[13] M. J. Antal, Jr., A. Brittain, C. DeAleida, S. Ramayya, and J. C. Roy, in *Supercritical Fluids*, ACS Symposium Series 329, edited by T. G. Squires and M. E. Paulaitis (ACS, Washington, DC, 1987).
[14] B. C. Wu, M. T. Klein, and S. I. Sandler, Ind. Eng. Chem. Res. **30**, 822 (1991).
[15] P. B. Balbuena, K. P. Johnston, and P. J. Rossky, J. Phys. Chem. **99**, 5196 (1995); J. Am. Chem. Soc. **116**, 2689 (1994).
[16] L. W. Flanagin, P. B. Balbuena, K. P. Johnston, and P. J. Rossky, J. Phys. Chem. **99**, 1554 (1995).
[17] J. F. Brennecke and C. A. Eckert, AIChE. J. **35**, 1409 (1989).
[18] P. T. Cummings, H. D. Cochran, J. M. Simonson, R. E. Mesmer, and S. Karaborni, J. Chem. Phys. **94**, 5606 (1991).
[19] P. Postorino, R. H. Tromp, M. A. Ricci, A. K. Soper, and G. W. Neilson, Nature (London) **366**, 668 (1993).
[20] T. I. Mizan, P. E. Savage, and R. M. Ziff, J. Chem. Phys. **98**, 13067 (1994).
[21] R. Mountain, J. Chem. Phys. **90**, 1866 (1989).
[22] A. A. Chialvo and P. T. Cummings, J. Chem. Phys. **101**, 4466 (1994).
[23] J. W. Tom and P. G. Debenedetti, Ind. Eng. Chem. Res. **32**, 2118 (1993).
[24] C. A. Eckert, D. H. Ziger, K. P. Johnston, and T. K. Ellison, Fluid Phase Equil. **14**, 167 (1983).
[25] C. A. Eckert, D. H. Ziger, K. P. Johnston, and S. Kim, J. Phys. Chem. **90**, 2738 (1986).
[26] R.-S. Wu, L. L. Lee, and H. D. Cochran, Ind. Eng. Chem. Res. **29**, 977 (1990).
[27] H. D. Cochran, P. T. Cummings, and S. Karaborni, Fluid Phase Equil. **71**, 1 (1992).
[28] Y. Ikushima, N. Saito, and M. Arai, J. Phys. Chem. **96**, 2293 (1992).
[29] J. P. Blitz, C. R. Yonker, and R. D. Smith, J. Phys. Chem. **93**, 6661 (1989).
[30] P. G. Debenedetti, I. B. Petsche, and R. S. Mohamed, Fluid Phase Equil. **52**, 347 (1989).
[31] I. B. Petsche and P. G. Debenedetti, J. Phys. Chem. **95**, 386 (1991).
[32] J. R. Quint and R. H. Wood, J. Phys. Chem. **89**, 380 (1985).
[33] W. M. Flarsheim, A. J. Bard, and K. P. Johnston, J. Phys. Chem. **93**, 4234 (1989).
[34] M. Rovere, D. W. Heermann, and K. Binder, J. Phys. Condensed Matter **2**, 7009 (1990).
[35] M. Rovere, P. Nielaba, and K. Binder, Z. Phys. **90**, 215 (1993).
[36] M. Rovere, J. Phys. Condensed Matter B **5**, 193 (1993).
[37] B. Smit and D. Frenkel, J. Chem. Phys. **94**, 5663 (1991).
[38] L. Verlet, Phys. Rev. **159**, 98 (1967).
[39] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Oxford University, Oxford, 1989).
[40] J. Gao, J. Am. Chem. Soc. **115**, 6893 (1993).
[41] H. Q. Luo, G. Ciccotti, M. Meyer, and B. Zappoli, Phys. Rev. E **51**, 2013 (1995).
[42] I. B. Petsche and P. G. Debenedetti, J. Chem. Phys. **91**, 7075 (1989).
[43] T. L. Hill, *Statistical Mechanics Principles and Applications* (Dover, New York, 1987).
[44] S. K. Ma, *Modern Theory of Critical Phenomena* (Benjamin, Reading, 1976).
[45] K. Kaski, K. Binder, and J. D. Gunton, Phys. Rev. B **29**, 3996 (1984).
[46] J. M. Haile, *Molecular Dynamics Simulation Elementary Methods* (Wiley, New York, 1992).
[47] The next smaller radius, $r_{ck} = 1.78$, for which $n_0 = 3$ and $\delta = 0$, does not adequately separate low and high density atoms, possibly because n_0 is too small to distinguish between large clusters and the small clusters of 2–3 atoms which occur in the cavity regions.
[48] Note that the amplitude of the fluctuations in $n_{hi}(t)$ on the molecular time scale, t_m, are of much smaller magnitude than those occurring on the cluster time scale, t_c, consistent with the conclusion that the majority of cluster rearrangements occur by slow collective motions.
[49] D. Chandler, J. Chem. Phys. **68**, 2959 (1978).
[50] $n_{hi'} = n_{hi}(t) + n_{av}(t)$, and we have included the average atoms in the high density count for simplicity.
[51] A. Ben-Naim, *Statistical Thermodynamics for Chemists and Biochemists* (Plenum, New York, 1992).

[52] N. E. Cusack, *The Physics of Structurally Disordered Matter: An Introduction* (IOP, Bristol, 1987).
[53] W. C. Swope, H. C. Andersen, P. H. Berens, and K. R. Wilson, J. Chem. Phys. **76**, 637 (1982).
[54] J. K. Lee, J. A. Barker, and F. F. Abraham, J. Chem. Phys. **58**, 3166 (1973).
[55] M. R. Mruzik, F. F. Abraham, D. E. Schreiber, and G. M. Pound, J. Chem. Phys. **64**, 481 (1976).
[56] It is conventional to define ξ in terms of the grand canonical distribution function $\bar{g}(r)$, because it is only in this ensemble that $\bar{g}(r)$, and thus ξ, is related to the isothermal compressibility through the compressibility equation (Ref. 51).
[57] The system defined by $r<r_+$ is an open system since x atoms may cross the $r=r_+$ boundary. Hence the appropriate density is an average value, $\bar{\rho}_{x+}$.
[58] The radial distribution functions are all step functions as a result of the design of the idealized system.
[59] For simplicity in the present discussion we neglect the presence of the average atoms as they do not appear in the partial distribution functions of interest.
[60] In order to perform the ensemble average we assume that the cluster size, V_+, and the cluster (or cavity) density, $\overline{\rho}_{x+}$, defined within the canonical ensemble for $r<r_+$ are already the average open system values, and that they may therefore be taken as constants with respect to fluctuations in $N_{\rm hi}$, $N_{\rm lo}$, or N. This result will be valid given a sufficiently large initial canonical ensemble.