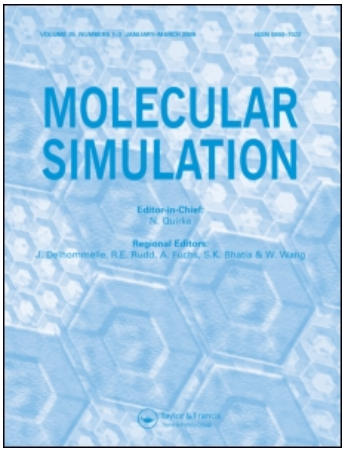

Molecular Simulation
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713644482

A Thermodynamic-scaling Study of Gibbs-ensemble Monte Carlo

John P. Valleau

Online publication date: 13 May 2010

To cite this Article Valleau, John P.(2003) 'A Thermodynamic-scaling Study of Gibbs-ensemble Monte Carlo', Molecular Simulation, 29: 10, 627 — 642
To link to this Article: DOI: 10.1080/0892702031000103167
URL: http://dx.doi.org/10.1080/0892702031000103167

Molecular Simulation, VOL. 29 (10–11), October–November 2003, pp. 627–642

Taylor & Francis
Taylor & Francis Group

A Thermodynamic-scaling Study of Gibbs-ensemble Monte Carlo

JOHN P. VALLEAU

Chemical Physics Theory Group, Department of Chemistry, University of Toronto, Toronto, Ont., Canada M5S 3H6

(Received October 2002; In final form February 2003)

This is meant to contribute to our understanding of Gibbs-ensemble Monte Carlo (GEMC) computations. We study the liquid–gas transition for a Lennard–Jones fluid, for which there exist extensive precise data on the canonical Helmholtz free energy. These data are used to *predict* the results of GEMC computation; the adequacy of these predictions is confirmed by comparison with actual GEMC results. We are then able to use such (very quick and economical) predictions to study the behaviour to be expected in actual GEMC computations, and their dependence on various parameters of the computations. In particular, the existence of a "cusp" maximum in the box density distributions is confirmed, and sensitivity of the apparent "coexistence" properties to the overall density is observed. Such effects are particularly important in the critical region, but extend to temperatures substantially lower. The results suggest certain limitations and precautions which should be kept in mind when using this method.

Keywords: Gibbs-ensemble Monte Carlo (GEMC) computation; Phase transitions; Thermodynamic scaling; Monte Carlo

INTRODUCTION

The so-called "Gibbs Ensemble" method [1,2] in the Monte Carlo study of phase transitions is one of particular conceptual beauty. In an ordinary Monte Carlo study of a finite system in a phase-separating region, it is impossible to observe the two co-existing phases, primarily due to the free-energy expense that would be involved in creating an interface between them. Panagiotopoulos sought to overcome this problem by separating the two phases into separate volumes, so that no surface between the phases was required. He showed that by a combination of ordinary particle moves within each box and transfers of both particles and volume between the two boxes (with total volume and number of particles fixed, in the usual, canonical, version) it would be possible to establish an equilibrium between the phases.

The method has proved very effective at locating the phase transitions of model systems, and has been widely applied. In spite of that there remain many ambiguities about the exact nature of the "Gibbs ensemble", although insight has been obtained progressively in a number of theoretical and computational investigations [3–7]. This article is meant to add to that understanding. We are concerned with the reliability of coexistence data from particular GEMC runs, and with the proper analysis of such data. Of particular interest is the behaviour of Gibbs-ensemble Monte Carlo (GEMC) computations as the critical region is approached, because it would be nice to know how, or whether, the technique can be used to obtain information on the critical behaviour. In that regime fluctuations become especially large, of course. It is crucial to know not only how GEMC should behave in its *thermodynamic limit*, but how that behaviour may depend on system size for the rather small systems feasible in Monte Carlo simulation.

In an earlier short article [8], it was pointed out that problems may occur in the interpretation of Monte Carlo studies of liquid–gas transitions by GEMC as the critical region is approached, due to the fact that the numbers of particles in the two "phases" will usually be unequal. It was shown that such problems could arise from the number-dependence of the thermodynamic properties of small systems. It was suggested that "thermodynamic equilibrium" between the two GEMC "boxes" would therefore be confined to temperatures well below the finite system's expected effective critical temperature, and also that the mean density of the two "phases" might

ISSN 0892-7022 print/ISSN 1029-0435 online © 2003 Taylor & Francis Ltd
DOI: 10.1080/0892702031000103167

be artificially shifted in a way depending on the parameters controlling the computation. The possibility of such effects was illustrated with the help of detailed information on the relative Helmholtz free energy $A(N,V,T)$ (with N particles in volume V at a temperature T) over a wide range of density and temperature, for (a particular version of) the Lennard–Jones (LJ) system, obtained from a "thermodynamic-scaling" MC study [9].

The discussion in that previous article was inadequate, however, because it neglected, as the article pointed out [8], the *fluctuations* in the partitioning of the particles and of the total volume between the two boxes of the GEMC computation. However, it is not very difficult to use the same information on relative free energies to estimate the expected behaviour of the Gibbs ensemble for finite systems, *including* these constrained fluctuations. That is what is done below. These predictions are tested by carrying out some actual GEMC computations, and are quite reliable. We can therefore use our predictive method to examine in a more comprehensive way the nature of the GEMC coexistence curve as the critical region is approached, and the way it depends on the parameters of the computation. (While the qualitative conclusions of the earlier article [8] quoted above are actually confirmed, they arise in a way rather different from that suggested by it. This is because it is the fluctuations themselves that play the decisive role in determining GEMC behaviour for finite systems; the number dependence of the expectation value of the finite-size specific free energy plays an interesting but secondary role, as we will see.)

The distribution function governing the box densities in a Gibbs-ensemble computation may be written (see Eq. (4), below) in terms of canonical partition functions. We make use of this to express the distribution in terms of the available precise data on the Helmholtz free energies $A(N,V,T)$; some approximations are necessary due to the limited density and N ranges of the available data, but these look fairly innocent, and we confirm this by comparing with actual GEMC computations using the same parameters. In a recent article, Bruce [7] instead wrote the distribution in terms of grand-canonical density distributions. This allowed him to discuss the Gibbs-ensemble behaviour, at least *at the bulk critical point*, in terms of the expected Ising-like universal behaviour of the critical grand-canonical density distribution. He was thus able to make very interesting observations about the so-called "restricted Gibbs ensemble" (in which particles but not volumes are transferred between the boxes), and about the form to be expected for the density distribution in the GE itself. Of course this approach has its own difficulties, as the author pointed out, and was unable to study the GE behaviour over a range of temperatures, as we do below.

Smit *et al.* [4] observed and discussed the unexpected appearance of a third, so-called "cusp", maximum in the box density distribution, near the overall density, in GEMC studies of the gas–liquid transition, and they proposed a phenomenological explanation in terms of the reduction of surface tension as criticality is approached. In some other studies [5,6], it has not been possible to observe such a peak, so that doubt has sometimes been cast as to its reality, and also about the adequacy of the explanation. Bruce's theoretical results [7] seem to confirm the expectations of such a "cusp" at the bulk T_c of a liquid–gas transition, however. The results we present confirm its existence and illustrate its appearance substantially below T_c and its evolution as the temperature increases through the critical region.

PREDICTIVE METHOD AND MODEL

Predicting GEMC Behaviour

In the most usual version of a Gibbs-ensemble computation there are two boxes 1, 2 (typically each periodic) with, say, volumes V_1, V_2 and containing N_1, N_2 particles, at a common temperature T. The V_i and N_i fluctuate at a fixed total volume V and number of particles N:

$$V_1 + V_2 = V, \quad N_1 + N_2 = N, \tag{1}$$

leading to box densities ρ_i (and specific volumes $\bar{v}_i$) given by

$$\rho_i \equiv \bar{v}_i^{-1} = N_i/V_i, \quad i = 1,2. \tag{2}$$

A particular microstate of the system corresponds to a particular value of V_1 and a particular assignment of N_1 (labelled) particles to V_1, as well as to particular configurations $\mathbf{r}^{N_1}$ and $\mathbf{r}^{N_2}$ and momenta $\mathbf{p}^{N_1}$ and $\mathbf{p}^{N_2}$ of the particles in the boxes 1 and 2. Its relative probability is given by the Boltzmann factor

$$\mathbf{p}^{\mathrm{GE}}\left(N_1,V_1,\mathbf{r}^{N_1},\mathbf{r}^{N_2},\mathbf{p}^{N_1},\mathbf{p}^{N_2}|N,V,T\right)$$
$$\propto \exp\left(-\beta\left[E(N_1,V_1,\mathbf{r}^{N_1},\mathbf{p}^{N_1})+E(N_2,V_2,\mathbf{r}^{N_2},\mathbf{p}^{N_2})\right]\right), \tag{3}$$

where $E(N_i,V_i,\mathbf{r}^{N_i},\mathbf{p}^{N_i})$ is the energy and $\beta \equiv 1/k_B T$. To find the distribution of N_1, V_1 we integrate over the particle configurations and momenta at fixed N_1, V_1, sum over the possible assignments of the N_1 particles to box 1, and obtain, after taking account of particle indistinguishability,

$$P^{\mathrm{GE}}(N_1,V_1|N,V,T) = \frac{Z(N_1,V_1,T)\,Z(N_2,V_2,T)}{Z^{\mathrm{GE}}(N,V,T)}, \tag{4}$$

in which $Z(N,V,T)$ is the canonical partition function

$$Z(N,V,T)=\frac{1}{\Lambda^{3N}N!}\int_V \mathrm{d}\mathbf{r}^N \exp\left(-\beta U(N,V,\mathbf{r}^N)\right) \equiv \frac{1}{\Lambda^{3N}N!}Q, \tag{5}$$

defining the "configuration integral" Q, where Λ is the "thermal wavelength" $h\sqrt{\beta/2\pi m}$ (h being Planck's constant and m the mass of each particle) and U is the potential energy of the configuration. Thus the Gibbs-ensemble partition function Z^{GE} is the sum of $Z(N_1,V_1,T)Z(N_2,V_2,T)$ over all N_1 (with $0\le N_1\le N$) and integrated over $\alpha=V_1/V$ (with $0\le\alpha\le 1$).

On approach to the *thermodynamic limit* we have of course

$$Z(N,V,T)=\exp(-\beta A(N,V,T)) \to \exp(-\beta N\bar{A}(\rho,T)), \tag{6}$$

where A is the Helmholtz free energy (which becomes strictly extensive in that limit, so we may write $A(N,V,T)/N\to\bar{A}(\rho,T)$, defining the *specific* free energy $\bar{A}(\rho,T)$). It follows that, for a Gibbs-ensemble simulation in the limit of large N and V (with the *overall* density $\rho_0\equiv N/V$ fixed) we have

$$P^{\mathrm{GE}}\left(N_1,V_1|N,V,T\right) \propto \exp\left(-\beta[N_1\bar{A}(\rho_1,T)+N_2\bar{A}(\rho_2,T)]\right). \tag{7}$$

Now it follows from Eq. (1) that $N_1=N(\bar{v}_0-\bar{v}_2)/(\bar{v}_1-\bar{v}_2)$ and $N_2=(\bar{v}_1-\bar{v}_0)/(\bar{v}_1-\bar{v}_2)$ where $\bar{v}_0=V/N$. Thus this probability (7) will have a *maximum* for exactly the densities $\rho_1\equiv\bar{v}_1^{-1}$, $\rho_2\equiv\bar{v}_2^{-1}$ given by the usual "double-tangent" construction on the isotherm $\bar{A}(\bar{v},T)$ at each temperature T, and, since the relative fluctuations of the densities will also vanish in that limit, the average properties of the coexisting phases in GEMC will match those of the canonical (or another) ensemble in the thermodynamic limit. This is satisfying. We will shortly note however that this agreement is far from persisting for computations on *finite* systems in the various ensembles.

As we have pointed out [8], the position and shape of GEMC coexistence curves for finite systems will depend on the fluctuating populations and volumes of the "phases" but also on the expected number dependences of the thermodynamics of the finite systems in each box. It is desirable, in examining this, to try to keep some account of these number-dependences.

In this connection, it is useful to recall the form of $Z(N,V,T)$ and the N-dependence of the corresponding free energy quantity $A(N,V,T)$ in finite systems. We have

$$\begin{aligned}\beta A(N,V,T)\equiv N\beta\bar{A}_N(\rho,T)&=-\log Z(N,V,T)\\ &=-\log(V^N/\Lambda^{3N}N!)-\log(Q(N,V,T)/V^N)\\ &=\beta A^{\mathrm{id}}(N,V,T)+\beta A^{\mathrm{ex}}(N,V,T)\end{aligned} \tag{8}$$

where the final lines describe separation into "ideal" and "excess" contributions. The leading identity defines a *finite-system* specific free energy $\bar{A}_N(\rho,T)$, which will however, unlike that in the thermodynamic limit, depend on N (at a given ρ and T). This is in part because $N!$, which occurs in the "ideal" contribution βA^{id} (see Eq. (8)), differs from N^N, an effect that can be well approximated by including the first three terms of Stirling's approximation $\log N!=N\log N-N+\log\sqrt{2\pi N}+\ldots$ In addition to this density-independent effect, $\log(Q)$ will not be strictly extensive; this will be reflected in a number-dependence of the dependence of $\bar{A}_N$ on ρ and T. Thus we can expect that, if $N\neq N'$,

$$\beta\bar{A}_N(\rho,T)-\beta\bar{A}_{N'}(\rho,T)\approx\frac{\log\sqrt{2\pi N}}{N}-\frac{\log\sqrt{2\pi N'}}{N'}+C_{\mathrm{A}}(N,N',\rho,T); \tag{9}$$

the "ideal" terms independent of density arise from the fact that $N!\neq N^N$ while the density-dependent term C_{A} describes additional effects arising from the (small) N-variation of $(1/N)\log Q$ (as manifested in the subtle changes of *shape* of the free energy isotherm with changing N (cf. Fig. 1)).

We wish to make use of rather precise data about finite-system free energies obtained for a particular version of the Lennard–Jones (LJ) fluid by a temperature-and-density-scaling Monte Carlo (TDSMC) study [9]. The LJ pair potential is defined by

$$\phi(r)=4\epsilon\left(\left(\frac{\sigma}{r}\right)^{12}-\left(\frac{\sigma}{r}\right)^{6}\right), \tag{10}$$

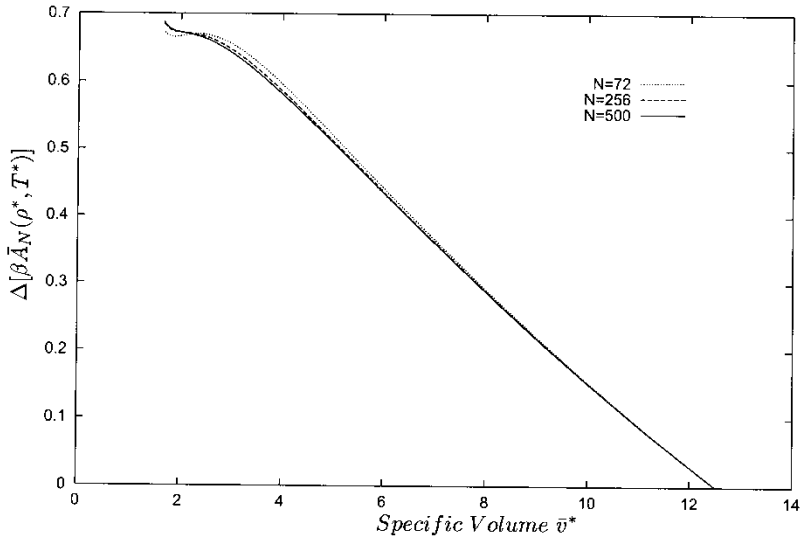

FIGURE 1 TDSMC results for the relative free energy $\Delta[\beta\bar{A}_N(\rho^*, T^*)]$, plotted against the specific volume $\bar{v}^* = 1/\rho^*$ along the isotherm at $T^* = 1.20$. The reference state is $T^* = 1.20$, $\rho^* = 0.08$. Results are shown for the three different system sizes studied, so the N-dependence of the density-dependences, to be expected in small systems, is apparent. These "curves" actually consist of line segments joining adjacent raw data points, of which there are 53 along the isotherm; isotherms generated from the global (ρ^*, T^*) fits mentioned in the text are completely indistinguishable on this scale.

where r is the distance between the centres of the pair of particles; this defines the parameters ϵ and σ characteristic of the model. By "particular version" we mean that in which the LJ potential is truncated at half the (current) box length, and a "tail correction" is added to approximate the interactions at larger distances, under the usual approximation that the pair correlation function can be taken as unity for such distances (cf. Ref. [10] for details). (As we will see, this is not actually a particularly attractive approximation for use in Gibbs-ensemble computations[†]). The Lennard–Jones parameters are used to define "reduced" thermodynamic variables used in what follows: the reduced temperature $T^* \equiv k_B T/\epsilon$, the reduced density $\rho^* \equiv \rho\sigma^3$ and the reduced volume and specific volume, $V^* \equiv V/\sigma^3$ and $\bar{v}^* \equiv V^*/N$.

In the TDSMC study, data were obtained for the relative reduced specific free energy, i.e. for

$$\Delta\left[\beta\bar{A}_{N'}(\rho^*, T^*)\right] \equiv \beta\bar{A}_{N'}(\rho^*, T^*) - \beta_{\text{ref}}\bar{A}_{N'}\left(\rho^*_{\text{ref}}, T^*_{\text{ref}}\right), \qquad (11)$$

where the final term refers to some fixed reference state $\left(\rho^*_{\text{ref}}, T^*_{\text{ref}}\right)$. This was done for three sizes N' (72, 256 and 500), and for each size data were collected over a broad range of both ρ^* and T^*, covering a region including the critical region, and on a fairly fine grid: for each N' data were obtained for at least 901 thermodynamic states (ρ^*, T^*). Figure 1 shows results along a single isotherm (of the 17 isotherms explored) for each of the three system sizes studied. The TDSMC data on $\Delta[\beta\bar{A}_{N'}]$ for each N' over the whole region of density and temperature were fitted by a 50-parameter Padé function, in which numerator and denominator were double polynomials of 5th-order in $\sqrt{\rho^*}$ and $\sqrt{1/T^*}$. This fit is everywhere excellent (when compared to the apparent uncertainties of the data points), and tests confirmed that using somewhat more or less flexible Padé functions had negligible effect on the result. This fitted function was used in what follows.

Using Eqs. (8), (9) and (11) and recalling Eq. (4), we can write $P^{\text{GE}}(N_1, V_1^*|N, V^*, T^*)$ for finite systems, in terms of these *relative* specific free energies, in the form

$$\begin{aligned} P^{\text{GE}}&(N_1, V_1^*|N, V^*, T^*) \\ &\propto \exp\{ - N_1\Delta[\beta\bar{A}_{N'}(\rho_1^*, T^*)] - N_2\Delta[\beta\bar{A}_{N'}(\rho_2^*, T^*)] \\ &\quad - N_1 C_{\text{A}}(N_1, N', \rho_1^*, T^*) \\ &\quad - N_2 C_{\text{A}}(N_2, N', \rho_2^*, T^*) - \log\sqrt{N_1 N_2}\}, \end{aligned} \qquad (12)$$

where we have discarded irrelevant constants.

[†]This version of the model does not reflect the possibility that the pair correlations may become rather long-ranged, especially in the critical region, and it has the objectionable feature that the cutoff radius, and hence the pair potential, is made to depend on the simulation box volume. It is nevertheless popular, since it leads in a simple way to mean energies close to those of a complete, untruncated, Lennard–Jones potential. We have checked, using a version of the Lennard–Jones model with a fixed cutoff radius, that the qualitative GEMC behaviours described here are reproduced for such a model.

Unfortunately, the TDSMC data do not give continuous information about the N-behaviour of the C_A terms, only what can be inferred from the differences in the behaviour of $\Delta[\beta\bar{A}_{N'}]$ at the three discrete values of N′ studied, $N' = 72$,256 and 500; these differences are fairly modest, as we can see in Fig. 1 for a typical isotherm.

An obvious approximation is to neglect the C_A terms in Eq. (12), thus retaining only the "ideal" part of the number dependences of the relative specific free energies. We call this "Approximation 1". Many of the characteristics of GEMC calculations to which we shall draw attention are of a qualitative or only semi-quantitative nature, and for them this "Approximation 1" is perfectly adequate.

On the other hand, we can do better than this, including the density-dependent number-effects to a reasonably good approximation (at least at the lower temperatures), if we take advantage of the fact that the density dependence of the number dependence of the specific free energy remains unimportant at the lower end of the density range covered by the TDSMC study. We can do this by interpolating the $\Delta[\beta\bar{A}_{N'}]$ results for the three system sizes studied by TDSMC. To see this, it is convenient to use Eqs. (8) and (11) to rewrite $P^{GE}(N_1, V_1^*|N, V^*, T^*)$ as

$$\begin{aligned} P^{GE}&\left(N_1, V_1^*|N, V^*, T^*\right) \\ &= a\left(N, V^*, T^*\right)\exp\left(-N_1\beta\bar{A}_{N1}\left(\rho_1^*, T^*\right)\right. \\ &\quad \left.-N_2\beta\bar{A}_{N_2}\left(\rho_2^*, T^*\right)\right) \\ &= a\left(N, V^*, T^*\right)\exp\Big(-N_1\Delta\left[\beta\bar{A}_{N1}\left(\rho_1^*, T^*\right)\right] \\ &\quad -N_2\Delta\left[\beta\bar{A}_{N_2}(\rho_2^*, T)\right] - N_1\beta_{\text{ref}}\bar{A}_{N1}(\rho_{\text{ref}}^*, T_{\text{ref}}^*) \\ &\quad -N_2\beta_{\text{ref}}\bar{A}_{N_2}(\rho_{ref}^*, T_{ref}^*)\Big). \end{aligned} \tag{13}$$

Here the $\Delta[\beta\bar{A}_{N_i}]$ terms (cf. Eq. (11)) refer to the *actual* sizes N_1, N_2 rather than to the sizes N' for which we happen to have detailed TDSMC results (as in Approximation 1). However, we can estimate these quantities well by interpolating (or mildly extrapolating) those TDSMC results. The remaining, reference state, terms in Eq. (13) also in principle depend on the N_i. But as we can see in Fig. 1 the density-dependence of the number-dependence is nearly absent at the lower densities of the TDSMC range. Thus if we choose a low reference density ρ_{ref}^*, and relate all the *reference state* terms to those for some arbitrary fixed size N'', we can expect that dependence on N_1 and N_2, of the sum of the reference state terms in Eq. (13), will be well approximated by taking account only of the density-independent "ideal" contribution (cf. Eq. (12)). We have used the reference state $\rho_{\text{ref}}^* = 0.08$, $T_{\text{ref}}^* = 1.20$. We call the result "Approximation 2", and it has the form

$$\begin{aligned} P^{GE}&\left(N_1, V_1^*|N, V^*, T^*\right) \\ &= a'\exp\left(-N_1\Delta\left[\beta\bar{A}_{N_1}(\rho_1^*, T^*)\right]\right. \\ &\quad \left.-N_2\Delta\left[\beta\bar{A}_{N_2}(\rho_2^*, T^*)\right] - \log\sqrt{N_1N_2}\right), \end{aligned} \tag{14}$$

where irrelevant constants have been absorbed in $a' = a'(N, V^*, T^*, N'', \rho_{\text{ref}}^*, T_{\text{ref}}^*)$, and the $\Delta(\beta\bar{A}_{N_i})$ terms are estimated by interpolating or mildly extrapolating the TDSMC results.

This will be an excellent approximation as long as the interpolation or extrapolation is reliable. We fitted, to the available TDSMC results for $\Delta[\beta\bar{A}_{N'}(\rho_i^*, T^*)]$ (those for $N' = 72$,256 and 500), a quadratic in $1/N'$, at every temperature T^* and every box density ρ_i^* relevant in the calculation; these functions were used to approximate $\Delta[\beta\bar{A}_{N_i}(\rho_i^*, T^*)]$ in Eq. (14). Because the TDSMC results vary little and regularly with system size (cf. Fig. 1) this should be quite satisfactory where *interpolation* is involved; indeed, other simple fitting functions were also tried and made little difference. However, we can place less confidence in *extrapolation*. Now very often such extrapolation turns out to be inconsequential; it will become important only for conditions in which the population in either box makes substantial fluctuations to very small N_i, well below that of the smallest TDSMC system studied ($N' = 72$), when Approximation 2 may cease to be reliable. What makes this case worse is that the fitting function, a quadratic in $1/N_i$, may well go wild for very small N_i. So, as a further approximation we neglected the N_i-dependence for N_i less than some $N_{\min}$, by setting $\Delta[\beta\bar{A}_{N_i}] = \Delta[\beta\bar{A}_{N_{\min}}]$ for $N_i < N_{\min}$. The value of $N_{\min}$ is irrelevant for small T^* (roughly $T^* \leq 1.26$), because then the N_i do not anyway fluctuate to small values; the predictions should then be quite accurate. At higher temperatures, and particularly near criticality, there are larger N_i fluctuations, however. A value $N_{\min} = 40$ was chosen, since choosing any value greater than this had little effect on the predictions; this extrapolation approximation must to some extent degrade the accuracy of the predictions at the higher temperatures, however.

With either approximation, the distribution $P^{GE}(N_1, V_1^*|N, V^*, T^*)$ can be evaluated at every possible value of N_1 and each point of an arbitrarily fine grid of the possible values of V_1^*. To find from this distribution, instead, the joint distribution of the densities ρ_1^* and ρ_2^*, i.e. $P^{GE}(\rho_1^*, \rho_2^*|N, V^*, T^*)$, one needs only the Jacobian $(J = (N - \rho_1^*V^*)\times (N - \rho_2^*V^*)/(\rho_1^* - \rho_2^*)^3)$; the result can be summed or integrated to estimate reduced (and normalised) distributions or averages, such as $P^{GE}(\rho_1^*|N, V^*, T^*)$ or $\langle\rho_1^*\rangle$.

One matter should be readdressed [4]. Suppose that, instantaneously, $\rho_1^* \leq \rho_2^*$; it is then necessarily

the case that $\rho_1^* \leq \rho_0^* \leq \rho_2^*$ (where ρ_0^* is, as before, the *overall* density N/V^*). At temperatures far below criticality, ρ_1^* and ρ_2^* will tend to be very different, and also the fluctuations of density in each box will then be small, so the box densities will have a negligible probability of approaching ρ_0^*; in that case the "gas" or "liquid" identity of each box will persist throughout a GE simulation. At temperatures such that the probability of ρ_0^* in each box is *not* negligible, however, interchanges of those identities will sometimes occur and we will argue that no reliable information on the phases will then be available from the individual box averages. Our predictions will imply that this will occur even well below the actual critical temperature, and the GEMC computations confirm this. (For example for a 512-particle system such interchanges are frequent at $T^* = 1.27$ and occur occasionally at $T^* = 1.26$, while the corresponding ($N = 256$) *canonical* critical temperature is near 1.34 and the *bulk* T_c^* is thought to be about 1.31.) A response to this situation, apparently adopted in some studies, is simply to use data only for temperatures so low that no such "box interchanges" are observed. A more common alternative has been to collect averages for the low-density ("gas") and high-density ("liquid") boxes throughout the run, regardless of their instantaneous locations as box 1 or 2. This procedure leads to well-defined so-called "gas" and "liquid" box averages at *every* temperature, although the physical interpretation of such results now evidently requires some examination, as we discuss below. The GEMC calculations we report are of this kind; that is we allow interchanges of the relative density, but keep averages separately for the lower-density ("gas") box and the higher-density ("liquid") box throughout; the resulting graph of the distributions of both box densities is what is termed the "overall density distribution" by Smit *et al.* [4].

Reliability of the Predictions

The behaviour to be expected from GEMC computations can be predicted very quickly by these techniques. To verify their relevance, some predictions based on the above approximations are compared here with some *actual* GEMC computations on the same system. The Gibbs-ensemble computations were carried out in the usual way. The two boxes were cubic and periodic. The runs reported were of 48×10^6 Monte Carlo steps or longer.

We predict the behaviour of GEMC runs involving $N = 512$ particles. For a finite, N-particle, system in the canonical ensemble (unlike the situation in the grand canonical or isothermal–isobaric ensembles) there is a well-defined "finite-system critical point", at which, for a particular temperature, the density gap vanishes and the compressibility diverges; (this is the point at which the double tangent points converge in a plot of $\bar{A}_N$ as a function of $\bar{v}$). The resulting canonical critical density ρ_c^* for $N' = 256$ is not far from 0.325; we consider at first such a choice for ρ_0^* in the Gibbs ensemble, by setting $V^* = 1577$ at all temperatures. The joint distribution of the gas and liquid densities, ρ_g^* and ρ_l^*, is then easily estimated as described above.

The TDSMC data extend over the range $0.08 \leq \rho^* \leq 0.60$, and the fit can presumably not be relied upon to extrapolate correctly far beyond this range (which is that most crucial to the problem). As a second, merely practical, approximation we therefore neglected states with a box density outside the range 0.07–0.65. (In most of the states we have studied the effect of this truncation is negligible, because box densities outside this range will never occur. It could in principle affect the prediction of the gas density distribution at the lowest temperature we study, where the mean gas density is itself small; we will see that this error is small in practice. It could also in principle affect the results at higher temperatures, where the larger density fluctuations might carry the box densities into the neglected

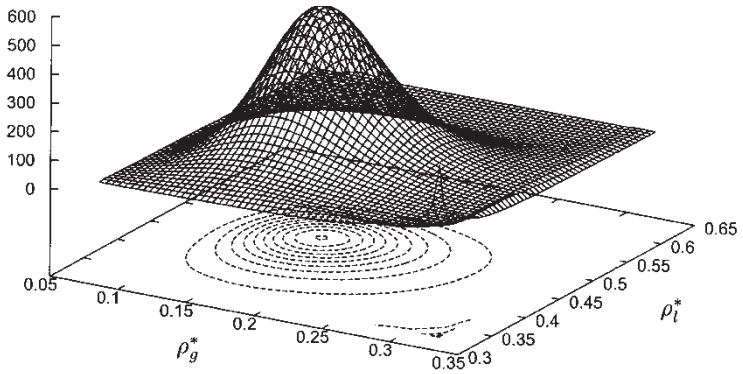

FIGURE 2 Predicted joint distribution of the densities in the two boxes, with $N = 512$ and $V^* = 1577$, at $T^* = 1.28$, using Approximation 1. (The boxes are ordered with respect to their densities, as discussed in the text.)

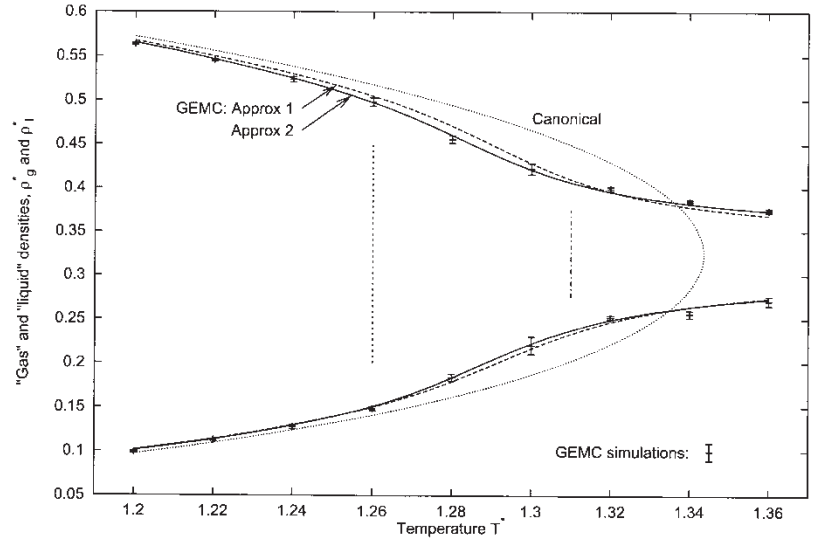

FIGURE 3 Mean densities of the low and high density boxes, as a function of temperature. At the lower temperatures these densities correspond to those coexisting phases, but this interpretation cannot of course be made at high temperatures. The dashed curves show our predictions, for fixed N and V^*, in Approximation 1; the solid curves are in Approximation 2. The points with error bars show results from actual GEMC simulations with the same parameters. The dotted curve shows the canonical coexistence behaviour for a system of 256 particles (which corresponds to the mean box population in the GEMC results). The dashed vertical line indicates the temperature region above which the low-density and high-density boxes will exchange identities during the run; the dash-dotted vertical line is roughly at the bulk critical temperature of the system.

ranges, but we will see that these effects are negligible in the present application.)

Figure 2 shows the resulting predicted joint probability distribution for the low and high density ("gas" and "liquid") boxes, in this case for $T^* = 1.28$ (using Approximation 2, but the results are similar in Approximation 1). By appropriate sums over such joint distributions we can find the predicted box density distributions at each temperature; it is then straightforward to average such density distributions to find the predicted mean densities in the low-density ("gas") and high-density ("liquid") boxes, for various temperatures. The resulting "coexistence curves" can be seen in Fig. 3, as the upper and lower solid curves for Approximation 2 and as dashed curves for Approximation 1. (For the $N = 512$ GEMC calculations the average number of particles per box is of course 256, and it is then natural to use the TDSMC data for this size, $N' = 256$, in Approximation 1.)

We defer any discussion of the interpretation and implications of Figs. 2–4 to the following section; here we examine merely the adequacy of the predictive method. To that end Fig. 3 also displays observed mean densities obtained in actual Gibbs-ensemble computations, for the same N and V^*, and

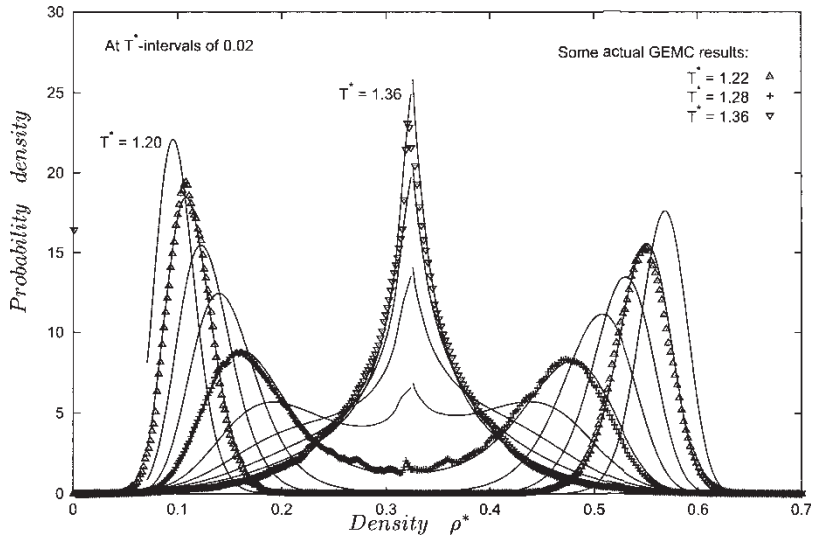

FIGURE 4 Predicted density distributions for both boxes at intervals of 0.02 in T^*, between $T^* = 1.20$ and 1.36, in Approximation 2. At three temperatures, 1.22, 1.28 and 1.36, distributions from actual GEMC computations are shown for comparison.

at several temperatures. Both approximations are fairly good. Approximation 1 seems however to overestimate the "liquid" densities, slightly but systematically. This should be due to the approximation of neglecting altogether the density-dependent part of the number-dependence of the free energy (described by the C_A terms in Eq. (12)). And indeed the approximate inclusion of these effects, by Approximation 2, appears to overcome that deficiency, which is satisfying. We have mentioned that the *interpolation* of the relative canonical specific free energies, with respect to system size, is likely to be fairly good, but that *extrapolation* may well lead to more serious errors. It turns out however that at the lower temperatures the box populations do not fluctuate to values of N_i sufficiently small to require any serious extrapolation of the TDSMC relative free energy results; we would therefore expect these predictions to be excellent, as is the case. As the critical region is approached, however, this is no longer the case and the predictions may become less reliable (although we can see that in fact the *average* densities are not at all unsatisfactory, in practice). Any inadequacies in the predictions of Approximation 2 will arise primarily from inadequacy of the N-dependence information available from the actual TDSMC data set available. If these free energy data were collected over a wider range of N, and for more values, the GEMC predictions would become thoroughly reliable and quantitative; the present accuracy is however entirely adequate to our present purpose.

Figure 4 shows some predicted (Approximation 2) density distributions, at intervals of 0.02 in T^*, from $T^* = 1.20$ to $T^* = 1.36$, for both the box with $\rho^* < \rho_0^*$ and that with $\rho^* > \rho_0^*$. The points show, for three of these temperatures, the corresponding distributions from actual GEMC computations. In accord with the above remarks, the agreement of the predictions with GEMC is entirely satisfactory at the lower temperatures. At higher temperatures the accuracy does degrade slightly. As discussed above, this is due primarily to the necessity to extrapolate the TDSMC data to small N; the problem seems less severe than one might have feared. (One can also see that, as expected, the neglect of densities below 0.07 (due to the limited density range of the TDSMC data) is not entirely adequate for the gas at $T^* = 1.20$; however the resulting error in the mean density will be very small, and this is confirmed in Fig. 3. This neglect, and also that of densities above 0:65, is seen to be inconsequential at higher temperatures.)

Overall, the predictions appear very satisfying. Were more complete data on N-dependence included in the TDSMC study, the procedures we have outlined would provide completely quantitative predictions of GEMC results. As it is, they are successful in giving reliable semiquantitative predictions of this behaviour. We can therefore use them with some confidence to explore various further aspects of such computations, as we do in what follows.

UNDERSTANDING GEMC BEHAVIOUR

Basic Behaviour

We begin by examining the nature of GEMC results and noting some problems likely to occur in their interpretation. For this purpose it is convenient to revisit Figs. 2–4. In the joint distribution of box densities (Fig. 2) one sees the expected large and broad maximum corresponding to coexistent gas and liquid, but there is also a small sharp peak near (ρ_0^*, ρ_0^*), and this must correspond to a nascent "cusp" maximum of the sort reported by Smit *et al.* [4], already present although this joint distribution is for a temperature well below criticality.

This is confirmed by the density distributions in Fig. 4. As expected, the two peaks of the box density distributions at the lowest temperatures are sharp and well-separated, and it is natural to identify them with the gas and liquid. As the temperatures increases one sees the emergence of the "cusp" maximum near ρ_0^*, as observed by Smit *et al.* [4]. Eventually, at high temperatures, *only* the "cusp" remains, and the remaining gap between the box densities then represents merely the random fluctuations of the densities in the boxes away from the *mean* density, associated with random fluctuations of N_1 and V_1; in the limit of large N these fluctuations would be negligible and the density distribution would show delta-functions at ρ_0^* for both boxes. At intermediate temperatures (in these small systems) the states described by the "cusp" and those described by the peaks corresponding to separate phases appear to *coexist*, as a result of the predicted fluctuations; there is therefore nothing which marks a sharp changeover between the "liquid–gas coexistence" regime at low temperatures and the "supercritical" regime at high temperatures; instead they subsist together over a range of temperatures. From another standpoint the "cusp" arises in a purely statistical way due to the large number of available states with densities close to ρ_0^* in both boxes, and is intrinsic to the GEMC computation in finite systems. It is difficult to make any direct contact with the original phenomenological discussion of the cusp [4] in terms of diminished surface tension.

In interpreting GEMC computations, averages over such box density distributions, namely the mean densities of the low- and high-density boxes, are ordinarily plotted as a function of temperature,

over some range of temperature, to provide a "coexistence curve". Figure 3 shows such averages, both as predicted (continuous curves for Approximation 2), and obtained from actual GEMC computations (points with error bars), over the range of temperatures available to us. "Coexistence curve" could be an appropriate description only at the lowest temperatures, evidently: at high temperatures its "width" (i.e. the "density gap") merely reflects the random fluctuations of density in the boxes away from the mean density of the system, as is to be expected in the small finite systems, and has nothing to do with any phase separation. Thus there can be no GEMC critical point for finite systems, in the sense of convergence of the "gas" and "liquid" densities.

The dotted curve maps the coexistence densities found (for $N = 256$) in the *canonical* ensemble; it is worth recalling that these GEMC and canonical coexistence curves are derived from the *same* TDSMC free energy data. It is clear that for *finite* systems the coexistence behaviours of the two ensembles are predicted to be very substantially different from each other, notwithstanding their convergence in the thermodynamic limit.

We noted that the "top" of the coexistence curve (near the critical point) is not directly accessible in a GEMC simulation. At low temperatures ($T^* \ll T_c^*$) the GEMC mean box densities correspond to coexisting phases and at high temperatures ($T^* \gg T_c^*$) to random density fluctuations. At intermediate temperatures they evidently correspond to some mix of contributions of both types, and have no simple physical interpretation. Nevertheless the mean box densities retain precise values, with nothing obvious to distinguish them from the more meaningful densities at lower temperatures; there will be no obvious way to decide the temperature above which such data will become misleading regarding coexistence properties. It is of interest to examine this issue.

It is clear in the literature that most researchers have accepted GEMC "coexistence" data for temperatures high enough that frequent interchanges of "low" and "high" density character between the boxes will have occurred. Others seem to have rejected this, however, reporting "coexistence" data only for lower temperatures where such interchanges likely did not occur. Such a restriction is surely appropriate: absence of interchange will ordinarily be reflected in an almost vanishing probability of either box density's approaching the mean density ρ_0^*; where, instead, the box density distribution is no longer negligible at that limit (so that interchanges are likely), it is clear that the density distribution from which the box average density is computed (and which is truncated at the limit) *cannot* be that corresponding to a gas or liquid. For the present case, this would imply that only data below $T^* = 1.26$ would correspond to acceptable coexistence data.

It is not certain, however, that even this restriction by itself is always sufficient, because as a precursor to the actual appearance of a "cusp" peak, we see enhanced probability of densities in the region between the coexistence peaks, made obvious by the striking asymmetry of the separate box density distributions which results (cf. Fig. 4). This does actually occur, for the present system (at least for the "liquid" box), at temperatures somewhat *less than* those where "interchanges" would be observed in a GEMC run of practical length. Where this effect is significant, the density distribution may be thought to include extra contributions (those which eventually lead to the "cusp") in addition to those of the separate phases. The extent of narrowing of the apparent GEMC mean-density gap may then be exaggerated, and no longer represent the shape of the coexistence curve. It is not easy to see how to decide in an objective way the temperature range over which this may be invalidating "coexistence" properties. (In the following section, we note a further concern regarding the temperature range over which one may regard GEMC simulations as providing coexistence information.)

Changing the Overall Density

In carrying out a Gibbs-ensemble computation (in the most common version), one chooses the temperature T^*, the total number of particles N and the total volume V^*, and thus fixes the overall density $\rho_0^* \equiv N/V^*$. If one has in mind approaching the critical point as the temperature is raised, with fixed N, it might be natural to choose ρ_0^* to be near the critical density ρ_c^* at each temperature, or at least to try to arrange to approach ρ_c^* as the critical temperature T_c is approached. However ordinarily one would not actually *know* ρ_c^* (or T_c^*) in advance, so in practice a certain arbitrariness would still exist in the choice of ρ_0^* even if such a strategy were attempted.

In fact, as long as the overall density ρ_0^* lies within the coexistence envelope, one expects to get phase separation, and then, *in the thermodynamic limit*, the properties of the two separate phases will be independent of this overall density. The impression is often given that the particular choice of ρ_0^* is equally unimportant in GEMC simulations of small systems; in fact usually values of ρ_0^* have not even been reported. Yet in a finite system there is no guarantee that the apparent coexistence properties will be independent of its value.

We wish to examine this matter, by asking whether variation of the choice of ρ_0^* does influence GEMC

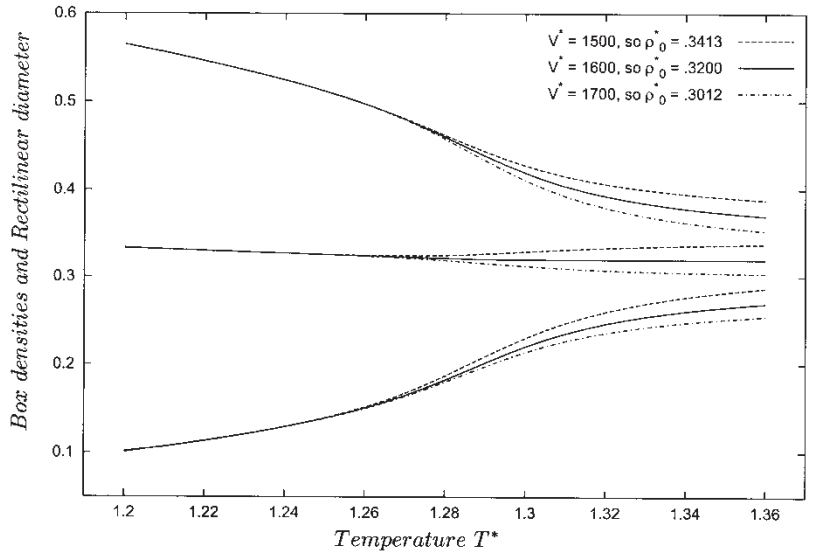

FIGURE 5 The effect of changing the total volume V^* on the apparent "coexistence" curves, based on Approximation 2. There are significant shifts of the densities of the "phases" extending to temperatures far below critical.

results. In general, we may expect it will, since, as noted above, at sufficiently high temperatures the predicted box density distributions just show fluctuations around the *mean* density ρ_0^* (not around ρ_c^*); so evidently the shape of the box density curves cannot be independent of ρ_0^*; the question is whether this is of concern at subcritical temperatures.

For the present LJ model the temperature and density at which the gas and liquid phases converge in a canonical ensemble (the canonical "finite-system critical point") is known fairly accurately for certain system sizes, and the N and V^* were chosen above to give an overall mean density ρ_0^* close to the critical density ρ_c^*. In general one will *not know* the critical density in advance, of course, and for a chosen N the volume V^* is therefore arbitrary. We wish to examine the effect of altering this arbitrary choice of V^*, and hence of ρ_0^*.

Figure 5 shows "coexistence" curves as predicted (using Approximation 2) for $N = 512$ and for three choices of V^*, differing by several per cent: 1500, 1600 and 1700. (Recall that the critical density corresponds to a V^*, slightly below 1600.) The high-temperature portions are of course shifted dramatically by varying the choice of V^*. The mean box densities for the various values of V^* do approach each other as the temperature decreases, no doubt converging asymptotically; however noticeable shifts persist over a quite large temperature range, including a substantial sub-critical region, persisting indeed throughout at least all the temperatures at which box interchanges occur. This means that not only is the apparent coexistence gap artificially *narrowed* even at temperatures far below critical, as we noted previously, but the apparent densities of both phases (and their mean) may also be *shifted* according to the arbitrary choice of ρ_0^* at such temperatures.

There is a further consideration. Critical densities are often estimated, in GEMC computations, by linearly extrapolating the "rectilinear diameter" (i.e. the locus of the mean of the densities of the two "phases") up to whatever is estimated to be the critical temperature (ignoring the small extra shift of the critical density associated with divergence of the slope of the diameter very near the critical point). These rectilinear diameters are shown in Fig. 5. It is clear that the choice of V^* could affect critical densities estimated by such an extrapolation. The extent of the resulting dependence of the apparent critical density on the choice of ρ_0^* depends of course on the range of temperature for which one accepts the validity of the mean box densities as representing coexistence data, as well as on the estimated critical temperature. (Linear extrapolation of the rectilinear diameters shown in Fig. 5, up to the best guess of the bulk critical temperature, $T^* = 1.31$, and using data up to $T^* = 1.27$, gives estimated critical densities ranging over a few per cent.)

These examples are concerned with a ρ_0^*-shift of only several per cent, and they correspond to keeping V^* (as well as N) fixed as the temperature is altered. It is not always clear in reported GEMC results whether the total volume has been chosen in a way such as this, or indeed in any systematic way whatever. (It is even tempting to ascribe part of the apparent "noise" in some reported GEMC coexistence curves as having been due not to statistical error but to an unsystematic choice of V^* as runs were done at a succession of temperatures.)

Overall Size Dependence

It is of interest to examine the effect of the overall system size on the observations to be expected in

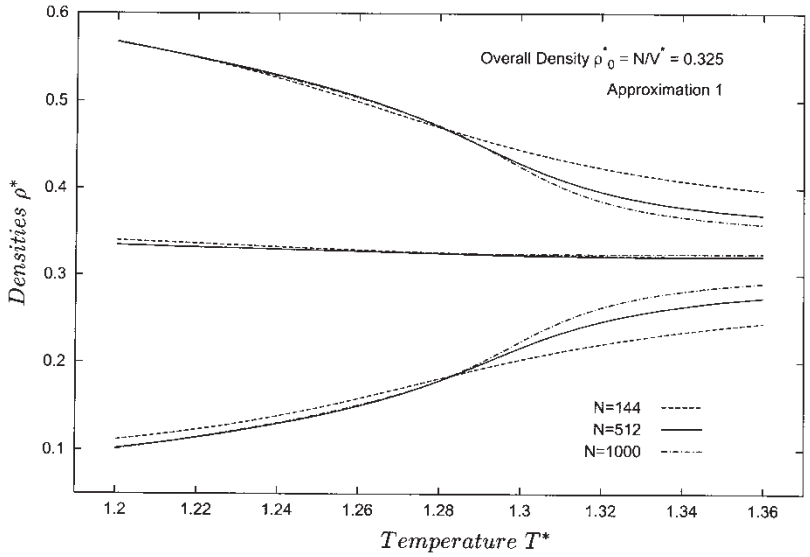

FIGURE 6 The mean box densities for systems of various sizes, all at the same overall density ρ_0^*, in Approximation 1.

GEMC computations. We report in Fig. 6 the predicted "coexistence" curves for $N = 144$, 512 and 1000, using Approximation 1. (The TDSMC data does not extend over a sufficient range of system size to derive these curves reliably in Approximation 2. The particular values of N used are convenient in that detailed TDSMC results for the model are available for systems of size $N' = N/2$ particles in each case, corresponding to the mean box populations, and these values of N' were accordingly used (cf. Eq. (12)) The observed canonical critical densities for these systems are essentially equal near $\rho_c^* = 0.325$, and V^* in each case was chosen to give ρ_0^* very near this value.

We can see in Fig. 6 that at high temperatures the mean-density gap between the boxes is larger for the smaller system; this is as expected, since the smaller the system the larger will be the relative fluctuations from the mean density. By contrast, at lower temperatures the density-gap instead *increases* with the system size: the larger fluctuations in the smallest system allow it to explore more fully the densities between those of the canonical double tangent points. (It is true that the effect of the number-dependence corrections neglected in Approximation 1 may well be especially important for the small system with $N = 144$, making this "coexistence curve" especially uncertain. However their inclusion is expected to narrow the density gap still more, thus enhancing the qualitative effects we describe.) This coexistence behaviour contrasts sharply with the behaviour in the canonical ensemble, where a larger system displays a narrower gap of mean density, at every subcritical temperature. The effect is also quite asymmetrical, and this leads to a noticeably different slope in the rectilinear diameter for the smallest system, at least in the present approximation; if this remains true with full account of the number-dependences, it will mean a noticeable size dependence of estimates of the critical density obtained by linear extrapolations of the rectilinear diameter for small systems (again in contrast to canonical estimates).

What may seem surprising is the remarkably small difference between the mean box densities for the two larger systems ($N = 512$ and 1000) at low temperatures; we would not expect this behaviour to change qualitatively as a result of including the full number dependences (Approximation 2). The fluctuations of box populations and volumes in GEMC have the effect, as we have seen, of "narrowing" the observed mean density gap, relative to the canonical ensemble; this effect increases of course for smaller systems. At the same time the canonical mean density gap itself "widens" as the system becomes smaller. It seems that in going from $N = 512$ to $N = 1000$ these opposing effects essentially cancel each other. Presumably in going to still larger systems the gap would once again narrow, therefore. Indeed it will do so, because the density gaps expected for the Gibbs ensemble must approach those in the canonical ensemble, as the thermodynamic limit is approached, and in that limit the gaps will probably be narrower, at each subcritical temperature, than those for either ensemble as displayed by the small-N systems we have been discussing.

Concerning Critical Parameters

For the liquid–gas transition the density gap between the phases is described as the "order

parameter" OP of the transition. For sufficiently large systems, and sufficiently close to criticality, the order parameter is expected to depend on temperature according to

$$\mathrm{OP} \equiv \rho_l^* - \rho_g^* \propto A_1(-\epsilon)^{\tilde{\beta}} + \text{Corrections},$$

where $\epsilon \equiv (T - T_c)/T_c$ and the leading term dominates the behaviour as ϵ approaches 0^- (at least in the thermodynamic limit). This "scaling" behaviour of the order parameter is described by the exponent $\tilde{\beta}$ (where the tilde has been added to distinguish the exponent from β defined previously); for the Ising model in three dimensions this is thought to have a value near 0.326, and simple fluids with short-range forces, such as the Lennard–Jones fluid, are thought to belong to the Ising "universality class" and thus to share its value of $\tilde{\beta}$. The constant A_1 is by contrast model-dependent.

A common approach to estimating the critical temperature from simulation results is to compare the observed density-gaps to the above scaling relation. A difficulty lies in the fact that the density fluctuations, increasing in range as T_c^* is approached, and which are responsible for the universal scaling described by the exponent $\tilde{\beta}$, are clearly limited in range, for finite systems, by the box size (finite-size effect); as a result the order parameter is expected to deviate from this scaling behaviour at the higher subcritical temperatures in such systems ("finite-size crossover") and only data for temperatures well below criticality can be used in estimating T_c^* and $\tilde{\beta}$. This expectation seems consistent with the behaviour observed in MC study of finite-system fluids in other ensembles. GEMC results, for temperatures well below criticality, have often been used in this way. Sometimes fits have included some correction terms (e.g. Wegner corrections), but usually the apparent precision of the data does not justify this. Simple plots of $(\mathrm{OP})^{1/\tilde{\beta}}$ as a function of temperature, assuming an Ising-like value of $\tilde{\beta}$, have often seemed plausibly linear at temperatures well below critical; the intercept with the temperature axis then constitutes an estimate of T_c^*. Alternatively the value of $\tilde{\beta}$ giving the best linear fit may be sought simultaneously. The resulting T_c^* are invariably "low", in the sense of having values not only far below those estimated for canonical systems of similar size, but also below estimates of the bulk-system critical temperature.

In the light of the GEMC behaviour we have explored above, this procedure looks perhaps a bit mysterious. As we have seen, the actual shapes of the coexistence curves are quite *different* for GEMC and canonical systems of comparable small size (see Fig. 3), in spite of the fact that they must converge on approaching the thermodynamic limit. Furthermore the shapes of the GEMC curves are controlled as well by the choice of V^*, and hence of ρ_0^*, which is essentially arbitrary (see Fig. 5). At the same time, changing the system size also changes the shapes of the coexistence curves (see Fig. 6). There seems little reason, then, to expect the above scaling behaviour for small GEMC systems. It is surprising to discover that it does seem roughly to hold. Figure 7 shows plots of the kind described, as predicted for GEMC with $N = 512$ (based on inclusion of both types of number correction, i.e. in Approximation 2), but with different overall densities ρ_0^*, and assuming $\tilde{\beta} = 0.326$. It is surprising that the GEMC plots for different values of ρ_0^* are so nearly coincident, since they arise from quite discrepant coexistence curves

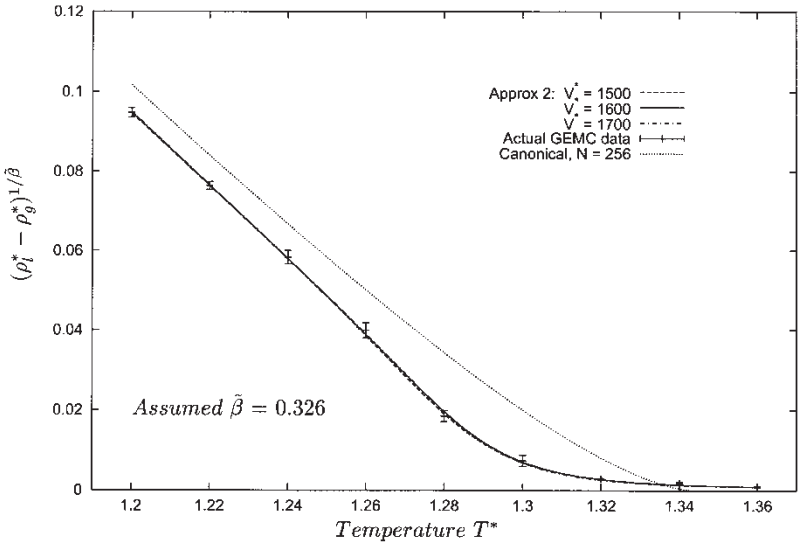

FIGURE 7 Simple-minded "scaling" plot, based on roughly "Ising-like" behaviour (i.e. exponent $\tilde{\beta} = 0.326$ and supposing one may neglect the corrections). The GEMC data and predictions are plausibly linear at the lower temperatures (and remarkably insensitive to V^*, which acts to shift the whole "coexistence" curve with only a small change in the density gap—cf. Fig. 5).

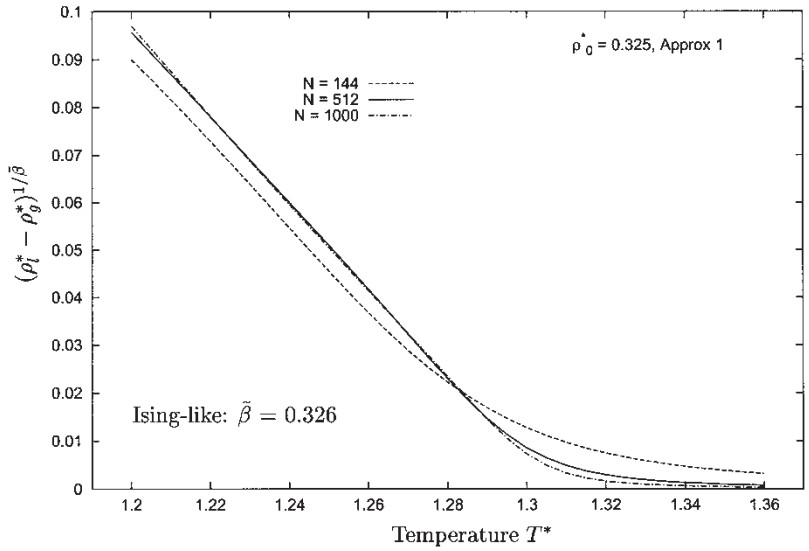

FIGURE 8 Similar "scaling" plots for systems of different sizes N, at $\rho_0^* = 0.325$, using Approximation 1 (cf. Fig. 6).

(cf. Fig. 5). It is the case that noisy data consistent with these predictions would seem plausibly linear at low temperatures. There remains however the problem, one theme of the present analysis, that one must decide on the range of temperature over which the observed average box densities may validly be accepted as representing those of coexisting phases. We pointed out that this decision influences the apparent critical density. In the same way, it will evidently influence the estimate of the critical temperature (and also, strongly, estimates of $\tilde{\beta}$ or of scaling corrections, should these be attempted).

Figure 8 shows similar plots predicted for GEMC systems with $N = 144$, 512 and 1000 in Approximation 1, all at $\rho_0^* = 0.325$. (As explained, the available data is inadequate to use Approximation 2 for the large and small systems; however Approximation 1 is expected to capture the size effects semi-quantitatively.) At high temperatures the N-dependence reflects merely the expected dependence on N of the random box density fluctuations around ρ_0^*. Estimates of T_c depend on extrapolation of the low-temperature results. In going from $N = 144$ to $N = 512$ the resulting apparent T_c clearly shifts upward, in contrast to the corresponding *downward* shift with increasing N for the canonical case; on the other hand, the GEMC shift on going from $N = 512$ to $N = 1000$ is negligible. (These facts are already obvious from the low-temperature coexistence curves shown in Fig. 9). The bulk value of T_c^* is known to be quite near 1.31 for this Lennard–Jones model, so the GEMC intercept

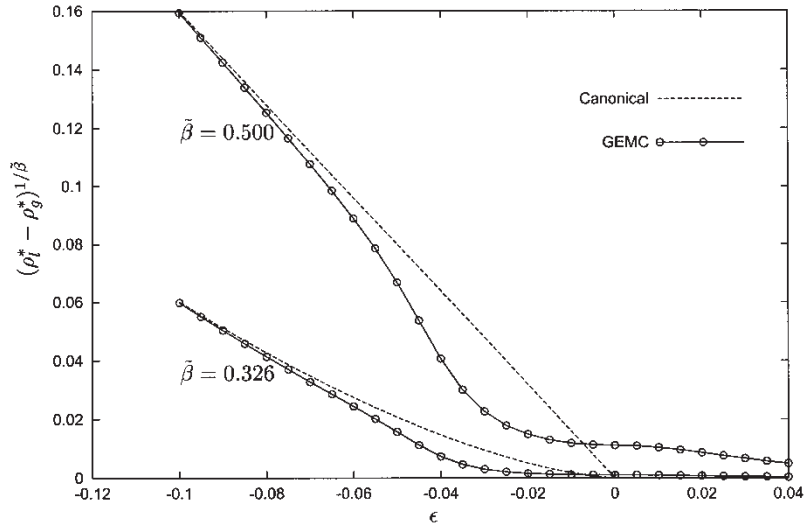

FIGURE 9 Scaling plots for the "fake" system of Eq. (15), for $N = 512$ and $\rho_0^* = \rho_c^*$, in both the canonical and GEMC ensembles. The upper curves propose mean-field-like behaviour ($\tilde{\beta} = 0.500$), the lower Ising-like behaviour ($\tilde{\beta} \approx 0.326$).

must presumably, as N is further increased, begin once again to increase. This corresponds to the further "narrowing" of the density gap mentioned earlier, and will be accompanied by a decrease of the slope of the linear portion (therefore a decrease of A_1) to approach the canonical behaviour. It seems to mean that the number-dependence of this apparent finite-system GEMC T_c must be quite complicated!

The apparent conformity of the GEMC order parameter to Ising-like scaling behaviour, in spite of the peculiarities of the coexistence curves, seems rather surprising; furthermore it is characteristic of GEMC results for coexistence behaviour. The *canonical* gas and liquid coexistence densities correspond to the double-tangent points on free energy curves; the very different "coexistence" curves in the case of finite-system GEMC arise from the fluctuations of ρ_1^* and ρ_2^* with respect to those densities, and depend on the asymmetry of the free energy curves around the tangent points. It is natural to ask whether the scaling of the GEMC order parameter we have noted is closely associated with that of the tangent points (as in the canonical case), or arises rather from the nature of the fluctuations.

To be more precise, suppose that the transition in question were not Ising-like, but, say mean-field-like (i.e. $\tilde{\beta} = 1/2$). Would the GEMC order-parameter scaling plot reflect this? As a test of this, we have looked at a completely fake system in which the free energy takes the form

$$\beta\tilde{A} = a(\epsilon)x^2 + bx^6, \tag{15}$$

where $x \equiv \rho^* - \rho_0^*$, b is constant, and $a(\epsilon) = -3bc^4\epsilon^2$; this ensures that the canonical order parameter (i.e. the gap between the tangent points) will obey $\rho_l^* - \rho_g^* = 2c\sqrt{-\epsilon}$. (The 6th power was used, rather than for example the 4th, merely to contain the density fluctuations within convenient ranges.) It is straightforward to predict the resulting GEMC behaviour, taking into account the fluctuations, in a finite system; the results reported are for $N = 512$ and $\rho_0^* = \rho_c^*$. The "coexistence curves" for the "fake" system, in the GEMC and the canonical ensembles, are qualitatively just like those shown in Fig. 3. It is easy to explore the possible scaling behaviour of the system in both ensembles, as we do in Fig. 9. As we can see, mean-field scaling ($\tilde{\beta} = 0.500$), which has been *forced* for the "canonical" case, certainly could not plausibly be concluded for the GEMC case. If instead the plots are made assuming an Ising-like exponent ($\tilde{\beta} \approx 0.326$; lower curves in Fig. 9), the canonical plot is necessarily strongly curved; however the GEMC result could plausibly be regarded as more-or-less linear over a substantial range of temperature, seeming to imply Ising-like behaviour!

(Another way of imposing mean-field canonical behaviour would be to take the actual TDSMC free energy isotherms, but ascribe them to temperatures on a distorted scale such that the canonical critical region became "quadratic". If this is done, it leads once more to behaviour similar to that in Fig. 9, with the GEMC results again appearing to scale in roughly an Ising-like way.)

The meaning of such exercises is not terribly clear, because of course the shapes of the free energy isotherms themselves reflect the model, and they also control the GEMC fluctuations. But it does seem to mean that the Ising-like order-parameter scaling noted in many GEMC results is not very directly connected with that observed in other ensembles, and raises the question of whether the apparent GEMC order-parameter scaling gives any useful information about the critical behaviour of a model. Apparent Ising-like behaviour in GEMC studies of transitions in various models, including for example the Restricted Primitive Model, has been offered as evidence that this universality class was characteristic of the model; is this valid, or should the behaviour be regarded merely as a characteristic of the GEMC technique? Other questions are raised (even where the system does actually belong to the Ising-like class): Does the slope of the GEMC scaling plot for a small system actually reflect a useful estimate of the corresponding pre-exponential factor A_1? Does the temperature at which deviation from near-linearity occurs in such a scaling plot in fact reflect that for which fluctuations become limited by box size, and thus that of the conventional finite-size crossover?

Connected with this is the question of seeking the bulk system critical temperature using GEMC data. In the canonical case the apparent T_c of finite systems is expected, and appears, to decrease smoothly as the system size N increases. It is possible to predict from the "finite-size-scaling" hypothesis the dominant term of the mathematical form of this decrease, and thus extrapolate to the bulk T_c. It is not clear what would in principle be possible along these lines for GEMC results. As we have seen, our predictions for the Lennard–Jones model indicate that the apparent T_c will increase with N at small N, but this increase seems essentially to cease in going from $N = 512$ to $N = 1000$; it is not clear whether a decrease then occurs for immediately larger systems, but eventually the apparent T_c must increase once again. Evidently no very simple extrapolation law would be consistent with this.

Bruce [7] tries to predict density distributions at the critical temperature (after assuming Ising-like character for the system). He proposes locating the critical point by adjusting the parameters to bring a match of the GEMC simulation results to such predictions. (This is along the lines of work

pioneered by Wilding and Bruce (see e.g. Ref. [11]) that has been applied to other ensembles; see however the warning by Fisher and co-workers [12,13] that this approach may not always be justified.) Indeed Bruce predicts a GEMC density distribution at criticality quite like that we find (Fig. 3) for temperatures near the expected $T_c^* \approx 1.31$ (although it is not certain that the origins of the two predictions are the same, in view of the highly localised "cusp" peak we find in the joint box density distribution (Fig. 2)). As far as we know there have been no serious attempts to use this approach on GEMC data.

Volume Fluctuations and Integrity of the Model

It may be worthwhile drawing attention to one further concern about GEMC computations. The particular LJ model under study, with a "cutoff" at half the box length, is in certain ways convenient, but it is somewhat pathological in that the interparticle force law depends on the box size. As long as one is always dealing with box volumes sufficiently large, the tail corrections are in any case tiny and this does not seem like a serious worry. In GEMC computations, however, the box volumes are of course allowed to fluctuate freely, and under some conditions unacceptably small cutoff radii may therefore occur. We can examine this possibility using our predictions. It is easy, for example, to predict the effect of restricting the box volumes never to fall below some minimum value $V_{\min}^*$. The minimum cutoff radius thought acceptable in LJ studies seems to have been 2.5σ, which corresponds to $V_{\min}^* = 125$. As would be expected, at quite low temperatures (and for $N = 512$) the discrepancies in the results arising from this volume restriction are negligible, reflecting the low probability of fluctuations to such small volumes at temperatures well below criticality. At higher temperatures the effect is however noticeable, which means that in GEMC runs fluctuations to worryingly small volumes (and cutoffs) may be expected. For the present model, it actually turns out that the discrepancies are at most a few per cent (for $N = 512$), usually smaller than the expected statistical uncertainty of GEMC results. (As a check on this we performed GEMC calculations embodying the same restriction $V_{\min}^*$, and confirmed that they gave mean densities differing from those of GEMC runs without the restriction only by amounts within their uncertainties.) So the problem turned out to be quantitatively less drastic than feared, for the present N and model, although disturbingly small cutoff distances were indeed observed.

The importance of such effects will be model-dependent and size-dependent, however, and it arises not only for a model with a fluctuating cutoff, such as that studied here, but wherever the potentials have a finite range; for example, with a fixed cutoff the interaction will simply be truncated (and anisotropic) when the volume fluctuates below a certain volume. Such effects will be worse for a system of smaller N, of course. It would seem prudent at least to monitor the volume fluctuations occurring in GEMC runs, or to carry out parallel GEMC runs, with and without a $V_{\min}$ restriction, to ascertain whether conflict of the volume fluctuations with integrity of the model may be causing unreliable results.

CONCLUSION

We have predicted the results of GEMC computations, using detailed data on the Helmholtz free energy. It is confirmed that to explain fully the behaviour of such computations the size-dependence of the thermodynamics of the fluctuating subsystems must be taken into account. The available free energy data allow this to be done to a very good approximation at sufficiently low temperatures, when the predictions are quantitative; at higher temperatures, near criticality for the system, we had to be satisfied with good semi-quantitative data. This is adequate, however, to allow exploration of the nature of GEMC results and their dependence on the parameters controlling the runs.

We see that, while the behaviour of systems in the canonical and Gibbs ensembles will converge in the thermodynamic limit, they are very discrepant for finite systems. In GEMC the subsystem (box) densities change smoothly, as the temperature increases, from representing those of coexisting phases to representing random fluctuations from the system mean density. At no temperature is there classic "critical" behaviour in the form of convergence of the "phase" properties, or indeed anything to mark in any obvious way the transition from two-phase behaviour to mere random fluctuations, since this occurs continuously. Care must therefore be exercised not to misinterpret apparent density-gap data as phase-density information at temperatures for which this is not valid. In practice, only data substantially below criticality contains useful physical information about coexistence.

At a given temperature T^* and for a fixed system population N, one must choose as well a total volume V^* before carrying out a GEMC computation. One might guess, on the basis of naive explanations of the GEMC method, that over a wide range this choice would be arbitrary, in the sense of having no effect on the properties of the two separate phases. However it is easy to show that this is not the case: changing V^*, and hence the overall density ρ_0^*, has the effect of shifting

both box densities, and hence the "coexistence curve". This shift is large only at high temperatures, but unfortunately appreciable shifts extend down into the low-temperature region from which one might have hoped to extract coexistence data, and they could lead to misleading conclusions. In particular, estimates of critical density obtained by extrapolation of the apparent rectilinear diameter may in practice depend on the choice of V^*, and so not be completely "physical".

The details of the number dependence of the apparent GEMC coexistence depend, from one point of view, both on that of the canonical behaviour (the shifts of the double-tangent points) and on the changing magnitude of the asymmetric GEMC fluctuations around that behaviour. These effects tend to oppose each other in some ways, as we have pointed out. The result seems to be a quite complicated size-dependence for GEMC coexistence behaviour, which probably makes extrapolations towards asymptotic behaviour problematic.

An unresolved problem concerns the extraction of estimates of critical temperatures and of the critical order-parameter exponent $\tilde{\beta}$ from GEMC box density data, even using only temperatures well below critical. In published studies, the density-gap data has often looked plausibly Ising-like, and T_c has been extracted on that basis. In view of the complications of GEMC behaviour and its sensitivity to N and V^* the question arises whether that behaviour is "physical", or an artefact of using the GEMC method on small finite systems. We have suggested that the latter is a possibility, but the full answer awaits further study.

Acknowledgements

The author enjoyed useful conversations with Athanassios Panagiotopoulos and Ilja Siepmann. The work was supported by the Natural Sciences and Engineering Research Council of Canada.

References

[1] Panagiotopoulos, A.Z. (1987) "Direct determination of phase coexistence properties of fluids by Monte Carlo simulation in a new ensemble", *Mol. Phys.* **61**, 813.
[2] Panagiotopoulos, A.Z. (1992) "Direct determination of fluid phase equilibria by simulation in the Gibbs ensemble", *Mol. Simul.* **9**, 1.
[3] Panagiotopoulos, A.Z., Quirke, N., Stapleton, M. and Tildesley, D. (1988) "Phase equilibria by simulation in the Gibbs ensemble: alternative derivation, generalization and application to mixture and membrane equilibria", *Mol. Phys.* **63**, 527.
[4] Smit, B., de Smedt, Ph. and Frenkel, D. (1989) "Computer simulation in the Gibbs ensemble", *Mol. Phys.* **68**, 931.
[5] Mon, K.K. and Binder, K. (1992) "Finite size effects for the simulation of phase coexistence in the Gibbs ensemble near the critical point", *J. Chem. Phys.* **96**, 6989.
[6] Recht, J.R. and Panagiotopoulos, A.Z. (1993) "Finite-size effects and the approach to criticality in Gibbs ensemble sinulations", *Mol. Phys.* **80**, 843.
[7] Bruce, A.D. (1997) "Finite-size critical behaviour in the Gibbs ensemble", *Phys. Rev. E* **55**, 2315.
[8] Valleau, J.P. (1998) "Number-dependence concerns in Gibbs ensemble Monte Carlo", *J. Chem. Phys.* **108**, 2962.
[9] Valleau, J.P. (1999) "Thermodynamic-scaling methods in Monte Carlo and their application to phase equilibria", In: Ferguson, D., Siepmann, I. and Truhlar, D.G., eds, *Advances in Chemical Physics*, (Wiley, New York) **105**, pp 369–404. A more comprehensive review is forthcoming.
[10] Valleau, J.P. (1993) "Density-scaling Monte Carlo study of subcritical Lennard-Jonesium", *J. Chem. Phys.* **99**, 4718.
[11] Wilding, N.B. and Bruce, A.D. (1992) "Density fluctuations and field mixing in the critical fluid", *J. Phys.: Condens. Matter* **4**, 3087.
[12] Orkoulas, G., Fisher, M.E. and Panagiotopoulos, A.Z. (2001) "Precise simulation of criticality in asymmetric fluids", *Phys. Rev. E* **63**, 051507.
[13] Fisher, M.E. and Orkoulas, G. (2000) "The Yang-Yang anomaly in fluid criticality: experiment and scaling theory", *Phys. Rev. Lett.* **85**, 696.