Computer simulation of fluid phase transitions

Nigel B. Wilding
Department of Mathematical Sciences, The University of Liverpool, Liverpool L69 7BZ, United Kingdom

(Received 9 March 2001; accepted 1 July 2001)

The goal of accurately locating fluid phase boundaries by means of computer simulation is hampered by difficulties associated with sampling both coexisting phases in a single simulation. We explain the background to these difficulties and describe how they can be tackled using a synthesis of biased Monte Carlo sampling and histogram extrapolation methods, in conjunction with a standard fluid simulation algorithm. The combined approach provides a powerful method for tracing fluid phase boundaries. © 2001 American Association of Physics Teachers. [DOI: 10.1119/1.1399044]

I. INTRODUCTION

One of the fundamental tasks of statistical mechanics is to link the microscopic (atomic-scale) description of a particular substance to its equilibrium macroscopic (thermodynamic) properties. Typically the former is prescribed in terms of a model that specifies the constituent molecules and their mutual interactions. Given this description, the challenge is to obtain the associated macroscopic properties—quantities such as the compressibility, heat capacity, and, above all, the phase behavior, that is, the conditions under which the substance forms a solid, liquid or gas. Computer simulation is one way to do so.

In many respects, a simulation can be viewed as an experiment performed not on a real substance, but on a model system. As in real experiments, we have to prepare an equilibrated sample under the desired thermodynamic conditions. As in real experiments, we can measure the physical properties of the sample. But a simulation can have important advantages over a real experiment; materials can be studied that are too expensive, too complicated, or too dangerous to be tackled by real experiments. Furthermore, because the simulator has access to complete information about the state of the model system, there are fewer restrictions on which properties can be measured. Accordingly, information and insight can be gleaned from a simulation that is not easily obtainable by experiment. 1

However, simulations have their limitations. The chief drawbacks are constraints on computer speed and memory; most contemporary computers can deal with systems comprising only a few thousand particles—many orders of magnitude smaller than the $~10^{23}$ typically found in experimental samples. Such restrictions lead to finite size effects, that is, spurious artifacts and systematic discrepancies compared to the bulk limit. It should also be borne in mind that irrespective of the sophistication of the simulation techniques employed, results for macroscopic equilibrium behavior will reflect reality only to the extent that the underlying model correctly captures the true microscopic nature of the substance under study.

Simulation strategies for obtaining the equilibrium phase behavior of classical fluid systems fall into two broad categories: molecular dynamics (MD) and Monte Carlo (MC). Both have been discussed extensively in a number of introductory texts and articles, see for example, Refs. 2–5. The MD approach involves computing the phase space trajectory of a system of mutually interacting particles by integrating Newton’s equations of motion. Physical properties are measured in terms of time averages as the simulation evolves. MD is an attractive method if we are interested in obtaining dynamical information and also can be employed to obtain single phase thermodynamic properties. However for studies of phase transitions, that is, the process by which one phase spontaneously transforms into another, it is rarely a suitable approach because of hysteretic behavior (superheating and supercooling) wherein the temperature and pressure at which the transition occurs in a simulation may differ significantly from that of the bulk system. The nature of hysteresis is described in Sec. II.

It is generally agreed that phase transitions are best treated by Monte Carlo (MC) simulation. Here we employ a stochastic Markov process to generate a sequence of equilibrium configurations of the model system; physical properties are measured as configurational averages over the sequence. 2–5 The actual way that the system evolves from a single equilibrium configuration to the next is artificial, so information about physical dynamical processes is limited. Nevertheless, this loss is compensated by potential gains. Specifically, the MC simulator is at liberty to construct any number of methods by which the simulation efficiently explores the space of possible configurations. In this way it becomes possible to bridge time and length scales that cannot be probed using MD.

In this article we shall focus on recently developed MC techniques that allow us to accurately obtain the phase behavior of model fluids. The approach is a synthesis of several simulation techniques (developed originally in the context of lattice spin models) which, when combined with a grand canonical or constant pressure ensemble fluid simulation algorithm, provides a powerful and efficient route to phase coexistence properties. Before embarking on this description, it is instructive to review some key aspects of phase transition phenomenology, in particular, the origins of the hysteresis effect that for many years plagued simulation studies of phase transitions.

II. HYSTERESIS, INTERFACES AND THE FREE ENERGY BARRIER

The phase diagram of a typical simple substance such as Argon, is depicted in schematic form in Fig. 1. Depending on the values of the temperature and pressure, the substance can exist in three phases: solid, liquid, or gas. The corresponding regions of the phase diagram are delineated by phase boundaries at which a transition occurs from one phase to another. Notable features of this phase diagram are the triple point,
Phase diagrams such as that in Fig. 1 can be interpreted within the framework of thermodynamics by an appeal to the behavior of the free energy. An equilibrium system will choose its state to be that for which the free energy is a minimum. Phase boundaries arise naturally as those sets of points for which two phases have the same free energies, thus being equally favored thermodynamically. In experiments, however, a transition from one stable phase to another will not always occur exactly on the phase boundary. Instead we generally encounter “overshoot” or hysteresis, whereby the actual transition point depends on the thermodynamic history of the sample.

Fig. 1. Phase diagram of a simple substance in the pressure–temperature plane.

The following gedanken experiment will help to explain the origin of hysteresis. Imagine a purified fluid such as water in a sealed piston–cylinder arrangement at constant atmospheric pressure. Now suppose that by continuously adding energy, we start to raise the water temperature very slowly so that it always remains close to equilibrium. As we add more energy, the water temperature will rise steadily until at some point it boils—transforming into steam with a concomitant large and abrupt increase in the system volume. This type of transformation is an example of a first-order phase transition. If, however, we stop adding energy before all the water has vaporized, we observe coexistence between the liquid and vapor phases. That is, a portion of the container will be occupied by the liquid phase, separated by an interface from the remainder which contains vapor (see Fig. 2). In transforming from one phase to another, a system always passes through such mixed-phase configurations, and it is these configurations that are responsible for hysteresis.

Fig. 2. Photograph of coexisting liquid water and steam in a closed container. The denser liquid occupies the lower portion of the container, separated by an interface from the vapor.

Mixed-phase configurations possess a higher free energy than pure phase states. Because this additional free energy is entirely associated with the interface, it is often referred to as the surface tension. Due to their surface tension, mixed-phase configurations are thermodynamically less favorable than pure phase states at the phase boundary. As the water is slowly heated, it eventually reaches the phase boundary temperature $T = 373.15\,\text{K}$, above which steam has a lower free energy than liquid water and becomes the thermodynamically favored state. Nevertheless, it is possible to superheat liquid water beyond this temperature without transforming it to steam because water must first pass through the mixed-phase states of higher free energy before becoming pure steam. A similar effect occurs when cooling steam—it is possible to supercool it below 373.15 K without liquefying it. Hence, the temperature at which the transition occurs is not that at which the free energies of the two phases are equal, but instead depends on the initial state of the sample and the direction and rate at which the boundary is traversed.

Similar behavior occurs in simulations of first-order phase transitions. But here the problem is also intimately bound up with issues of finite size effects and simulation time scales. To appreciate how a simulation behaves near a first-order phase transition, it is instructive to consider a concrete example, namely, the liquid–gas transition of a prototype model fluid—the Lennard-Jones fluid. In this model the interaction potential for two point particles separated by a linear distance $r$ is given by

$$U(r) = 4\epsilon \left( \frac{\sigma}{r} \right)^{12} - \left( \frac{\sigma}{r} \right)^{6}$$

where the parameters $\epsilon$ and $\sigma$ set the strength of the interaction and the length scale respectively.

One way of performing a MC simulation of the Lennard-Jones fluid is to employ the grand canonical ensemble in which, for a given volume $V$, the total configurational energy $E$, and particle number $N$ are permitted to fluctuate, but with average values determined by prescribed values of the temperature $T$ and chemical potential $\mu$. The latter two variables span the phase diagram and by tuning their values, transitions can be induced between the gas, liquid and solid phases (see Fig. 1).

An outline of a MC implementation of the grand canonical ensemble is given in Box 1. Fluctuations in the energy and particle number occur by means of particle insertions and deletions. MC updates consist of either an insertion or a deletion attempt, which are chosen with equal probability. For an insertion, we choose a random position in the system and calculate the energy change $\Delta E_I$ associated with placing a new particle at that position. The trial insertion is accepted with a probability given by a metropolis rule designed to maintain detailed balance—a necessary condition for attaining thermodynamic equilibrium:

$$P = \min \left( 1, \exp \left( \frac{\Delta E_I - \Delta G}{k_B T} \right) \right)$$

where $k_B$ is the Boltzmann constant and $T$ is the temperature.
where $z = \exp(\beta m)$ and $\beta = 1/k_B T$. 7

Similarly for a particle deletion, we choose a particle at random from those present, calculate the energy change $\Delta E_D$ associated with its removal, and then perform the removal with a probability given by

$$ P_{\text{acc}}(N \rightarrow N-1) = \min \left[ 1, \frac{N}{zV} \exp(-\beta \Delta E_D) \right].  \tag{3} $$

Notwithstanding its simplicity, a grand canonical simulation of this type can provide a highly efficient computational route to equilibrium fluid phase properties. The simulation is even more efficient if, as is customary, the interparticle potential $U(r)$ is truncated at a cutoff radius $r_c$. 2 Contributions to $\Delta E_I$ and $\Delta E_D$ then arise only via local interactions and, by partitioning the simulation box into cubic cells of side $r_c$, the subset of contributing particles can be readily identified. Note that for simulations of liquid–gas phase transitions, we gain nothing by implementing explicit particle displacement moves. Instead, these moves are realized implicitly by virtue of repeated particle transfers, the sole use of which not only constitutes a valid algorithm (it is clearly ergodic) but focuses the computational effort on the bottleneck for phase space evolution, namely, the fluctuations in the number density.

Let us now examine how the grand canonical simulation algorithm behaves in the vicinity of the liquid–gas phase boundary. The primary observables of interest are the particle number $N$ and the configurational energy $E(\{r\})$, where $\{r\}$ denotes the set of particle positions, that is, the configuration. In particular, the fluctuations of the density, $\rho = N/V$, provide much insight into the nature of phase coexistence in a finite sized system. In the simulation we can compute $p(\rho)$, the probability density function of $\rho$ as a histogram averaged over many independent samples. The form of such a distribution close to liquid–gas coexistence is shown in schematic form in Fig. 3.

The principal feature of this distribution is its bimodal (double-peaked) character. Each of the peaks corresponds to one of the pure phases—the low density peak to the gas phase and the high density peak to the liquid phase. The location of the liquid–gas phase boundary line in the $\mu - T$ plane is determined by the equal peak weight criterion, that is, by the set of values of $\mu$ and $T$ for which the integrated weights (areas) under the two peaks are equal. To locate liquid–gas coexistence, we must tune $\mu$ and $T$ until the measured form of $p(\rho)$ is doubly peaked with equal areas under each peak. The problem is that to obtain accurate estimates of the relative peak areas, we need many independent samples of each of the two phases, which necessitates that the system pass many times between the two phases.

Unfortunately, the interphase route necessarily traverses the mixed-phase configurations in which regions of both

![Fig. 3. A schematic diagram of the form of $p(\rho)$ on the liquid–gas phase boundary.](image)
phases coexist. On account of their surface tension, such configurations have an a priori probability that is intrinsically low in comparison to those of the pure phase states. For this reason the trough separating the two peaks of \( p(\rho) \) may be regarded as a probability (or free energy) barrier to interphase transitions.\(^8\) The height of this barrier is the ratio of the maximum (peak) value of \( p(\rho) \) to its minimum value in the trough. For large barrier heights, the free energy cost associated with forming mixed-phase configurations is so high that transitions between the two pure phases become very rare on the time scale accessible to simulation. Consequently, the correlation time becomes large, hindering the accumulation of independent statistics on the relative peak weights and hence the accurate location of the phase boundary.

The barrier height depends on the temperature. At the critical temperature \( T_c \), the surface tension vanishes and with it the barrier height.\(^9\) If, however, we follow the liquid–gas boundary to progressively lower (subcritical) temperatures, the surface tension grows and the thermal fluctuations diminish. The growth in the barrier height is accompanied by a narrowing of the peaks of \( p(\rho) \). These features are illustrated in Fig. 4, which shows the measured form of \( p(\rho) \) obtained using the grand canonical algorithm close to the phase boundary at \( T=0.985T_c \) and \( T=0.965T_c \) for the Lennard-Jones fluid. Also shown in Fig. 4(b) is the corresponding evolution of the density expressed as a function of the number of Monte Carlo insertion/deletion attempts. For \( T=0.965T_c \), the barrier height is approximately 5 and interphase transitions are fairly frequent. On decreasing the temperature to \( T=0.965T_c \), however, the barrier height increases by a factor of 6 and interphase crossings become comparatively much rarer. This reluctance to transform between the phases is none other than hysteresis, viewed not from the standpoint of thermodynamics, but in terms of the underlying statistical mechanics.\(^8\)

Qualitatively similar effects occur as the linear dimension of the simulation box \( L \) is increased. The barrier height grows because the surface tension of the interface increases proportionally to its area (\(-L^2\) for spatial dimension \( d=3\)). Additionally, the peaks of \( p(\rho) \) narrow because the fluctuations in \( \rho \) diminish when measured on larger length scales. Even for modest values of \( L \), the barrier can be prohibitively large for interphase crossings to occur during the accessible simulation time scales.

We see that we are caught between the “rock” of wishing to minimize finite size effects by employing a large value of \( L \), and the “hard place” of seeking to sample on time scales exceeding the correlation time. Evidently, a better approach than simple grand canonical simulation is needed if we are to simulate first-order phase transitions at temperatures significantly below criticality. Recently, considerable effort has been invested in developing new MC simulation methods for circumventing the problems that we have identified. In the next sections we describe one solution to the problem which is rapidly becoming the method of choice for high resolution studies of fluid phase equilibria.

### III. BEATING THE BARRIER

Approaches for dealing with the free energy barrier in simulations of phase transitions fall into two broad categories: simulations without interfaces and biased sampling techniques. Foremost in the first category are methods such as Gibbs ensemble Monte Carlo\(^10\) and Gibbs–Duhem integration,\(^11\) both of which have enjoyed widespread use in studies of phase transitions. Although distinct in character, both methods link the pure phases thermodynamically without traversing mixed-phase configurations. Both are versatile and fairly easy to use, but have significant limitations. Specifically, it is difficult to use the Gibbs ensemble Monte Carlo method to obtain coexistence data of high statistical quality without a large investment of computational effort. Gibbs–Duhem integration is more efficient, but potentially suffers from integration errors rendering it difficult to assess the accuracy of results for phase boundary properties.

More recently, an alternative approach has emerged that, although less straightforward to implement, offers the rewards of considerably greater efficiency, precision, and flexibility than Gibbs ensemble Monte Carlo or Gibbs–Duhem integration.\(^12\) This approach is based on a synthesis of two existing simulation techniques: multicanonical biased sampling and histogram reweighting. We now describe how this combined approach works and explain how to implement it.

### A. Multicanonical sampling

Multicanonical MC owes its origin to the biased sampling techniques introduced in the 1970s by Torrie and Valleau to...
calculate free energies. Recently, such techniques have gained fresh impetus with the realization that they permit the bridging of the free energy barrier at a first-order phase boundary. In this context the term multicanonical sampling was coined and the method has been applied successfully to the study of phase transitions and free energy landscapes in a variety of lattice-based spin systems.

The basic idea of multicanonical MC is to preweight the sampling of configuration space so as to artificially enhance the occurrence of the mixed-phase configurations of intrinsically low probability. By so doing, it is possible to overcome the occurrence of the mixed-phase configurations of intrinsically low probability. By so doing, it is possible to overcome the probability barrier separating the two pure phases, thereby allowing the simulation to pass unhindered between them. The result is a great reduction in the correlation time of the sampling process.

Within the grand canonical framework, the bias is achieved by using a preweighting function incorporated into the Metropolis acceptance criteria for particle insertions and deletions. This function modifies the probability of visiting the configurations of the various densities in such a way that the measured \( p(\rho) \) is approximately flat over the entire density range separating the two pure phases. Of course, the results of such biased simulations deviate from Boltzmann statistics and consequently lack direct physical significance. Nevertheless, it is possible to unfold from the simulation results the effects of the imposed bias and recover the physically relevant quantities that would have been obtained in an unbiased simulation. In general, the price we pay for this gain is the effort of finding a suitable form for the preweighting function. Fortunately, it happens that for the purposes of biasing our simulation procedure. It rests on the observation that histograms obtained by letting

\[
p(N|V,\beta,\mu) = e^{\eta(N)} \overline{p}(N|\beta,\mu,\eta(N)).
\]

The details of the implementation of this procedure are described in Sec. III C.

### B. Histogram reweighting

Histogram reweighting is the second ingredient in our simulation procedure. It rests on the observation that histograms of observables accumulated at one set of model parameters (in our case \( \beta \) and \( \mu \)) can provide estimates of histograms for other values of these parameters. Consider the joint probability distribution of energy and particle number at the particular parameter values \( \beta = \beta_0 \) and \( \mu = \mu_0 \). Formally \( p(N,E|V,\beta_0,\mu_0) \) is given by

\[
p(N,E|V,\beta_0,\mu_0) = \frac{1}{Z_0} \prod_{i} \int_{V} e^{-\beta_0 H_0}\delta(E-E(\{r\}))e^{-\beta_0 H_0},
\]

where \( H_0(\{r\},N) = E(\{r\}) + \mu_0 N \), and \( Z_0 = Z(\beta_0,\mu_0) \) is the partition function.

The multicanonical MC method operates by sampling not from a simple Boltzmann distribution with Hamiltonian \( H(\{r\},N) \), but from a modified distribution with effective Hamiltonian

\[
\tilde{H} = H + \eta(N),
\]

where \( \eta(N) \) is a preweighting function defined on the set of particle numbers \( N \). The associated number distribution is given by

\[
\overline{p}(N|\beta,\mu,\eta(N)) = \frac{1}{Z} \prod_{i} \int_{V} dV e^{-\beta \tilde{H}}.
\]

Suppose for the sake of argument that we are able to choose the preweighting function such that \( \eta(N) = \ln p(N) \), where \( p(N) \) is the desired Boltzmann density distribution. Inspection of Eqs. (4)–(6) reveals that this choice implies that \( \overline{p}(N) \) is a constant for all \( N \). To the extent that such a choice of preweighting function can actually be realized, the density performs a one-dimensional random walk over its entire domain, thereby allowing a very efficient accumulation of the preweighted histogram \( \overline{p}(N) \).

However, this happy state of affairs cannot be immediately achieved in general, because the preweighting function \( \eta(N) = \ln p(N) \) that flattens \( \overline{p}(N) \) is just the logarithm of the function we are trying to find. Therefore, we must find a way of obtaining a form of \( \eta(N) \) that approximates \( \ln p(N) \) sufficiently well that interphase transitions occur with an acceptably high frequency. More refined forms of \( \eta(N) \) can be obtained in the further course of the simulation.

If a suitable preweighting function has been found and a simulation performed to obtain good statistics for \( \overline{p}(N) \), the next step is to infer \( p(N) \) by unfolding the effects of the multicanonical preweighting. This unfolding is achievable because knowledge of the preweighting function \( \eta(N) \) tells us how the relative probabilities of the various \( N \) were altered. We therefore need only divide out the relative probability enhancements from \( \overline{p}(N) \) to yield \( p(N) \). This reweighting is done for each value of \( N \) in the range of interest by letting

\[
p(N|V,\beta,\mu) = e^{\eta(N)} \overline{p}(N|\beta,\mu,\eta(N)).
\]
must be close (in a sense we shall describe) to those at which the simulation was performed or the procedure loses accuracy. The problem is traceable to the fact that the reweighting represented by Eq. (9) may drastically modify the relative statistical weights of the various members of the set of configurations that contribute to the averages. Specifically, difficulties arise with the subset of configurations that have a very low Boltzmann weight at \( \beta_0, \mu_0 \), configurations that consequently occur only very rarely in the sample. For expectation values of observables calculated at \( \beta_0, \mu_0 \), contributions from this rare subset of configurations do not contribute disproportionately to statistical uncertainties. However, under the reweighting to \( \beta_1, \mu_1 \), these configurations may be assigned a much greater statistical weight that does not reflect their actual representation in the overall sample. The effect is to magnify the overall statistical error of the measured expectation values and is manifested as a reweighted histogram that appears “ragged” in its external regions.

One way of dealing with this problem is to perform a sequence of separate simulations at strategic intervals across the range of model parameters of interest. Typically the intervals are chosen so that the tails of the histogram of an observable (for example, the energy) accumulated at neighboring state points overlap. The role of histogram reweighting is to interpolate to the regions of parameter space between the simulation points. We note in passing that it is possible to combine in a self-consistent fashion the results of several different simulations at different model parameters and perform histogram reweighting on the aggregate data. For a description of this more sophisticated procedure, we refer the reader to Refs. 17 and 2.

C. Stitching together the pieces

How do we implement the above formalism? The task falls naturally into two parts: performing the actual simulation and the subsequent data analysis. We consider them in turn.

The implementation of the multicanonical preweighting in a grand canonical simulation is straightforward. Assuming an appropriate set of multicanonical weights has been found, the simulation proceeds as outlined in Box 1, except that the Metropolis acceptance probabilities for insertion and deletion [Eqs. (2) and (3)] are modified to read

\[
P_{\text{acc}}(N \rightarrow N + 1) = \min \left[ 1, \frac{\eta(N)}{\eta(N+1)} \frac{z^V \exp(-\beta \Delta E_I)}{(N+1)} \right],
\]

\[
P_{\text{acc}}(N \rightarrow N - 1) = \min \left[ 1, \frac{\eta(N)}{\eta(N-1)} \frac{N z^V \exp(-\beta \Delta E_D)}{z^V} \right].
\]

Now consider the quantities we seek to obtain, namely the probability distributions of \( N, E \) and any other observables of interest. It is tempting to accumulate these distributions in the form of histograms during the simulation by simply binning successive measurements into an array. But for continuous variables such as the energy, this strategy necessitates a prior choice for the bin width. If a choice be made that subsequently turns out to be unsatisfactory, we must repeat the simulation. A superior approach, retaining complete information, involves decoupling the data analysis from the simulation by recording the full history of raw data measurements. The data is then postprocessed by a separate analysis program. Although such an approach can make large output files, it has the overriding advantage of ensuring maximum flexibility for the data analysis.

To facilitate the post processing, the raw data should be accumulated in the form of a list. Suppose we perform a grand canonical MC simulation at \( \beta_0 \) and \( \mu_0 \), employing a weight function \( \eta(N) \). As the simulation proceeds, we make a succession of measurements at regularly spaced intervals of time of the observables \( N \) and \( E \), together with any other quantities of interest. Successive measurements of these observables are gathered into a list by appending them to a file, viz:

\[
\{E_0, N_0, O_0, \ldots\}
\]

\[
\{E_1, N_1, O_1, \ldots\}
\]

\[
\ldots
\]

\[
\{E_j, N_j, O_j, \ldots\}
\]

\[
\{E_M, N_M, O_M, \ldots\}
\]

where \( j \) indexes the series of \( M+1 \) measurements and \( O \) denotes an observable of interest. The post-processing program should first remove from the data the unwanted effects of the preweighting in order to recover the desired Boltzmann distributed statistical properties. Second, it should output probability density functions of the observables of interest. It then should (if desired) reweight the data to obtain estimates of histograms appropriate to parameter values different from those at which the simulation was performed.

The operations of histogram reweighting and bias removal can be accomplished simultaneously because their mathematical structures are very similar. We simply run through the data list assigning each entry \( j \) a statistical weight, \( w_j = e^{-[(\beta_1-\beta_0)E_j + (\beta_1 \mu_1-\beta_0 \mu_0)N_j + \gamma(\eta(N))]}, \) where \( \beta_1, \mu_1 \) are the parameters to which we wish to extrapolate. The complete set of \( M+1 \) weights \( w_0, w_1, \ldots, w_M \) is then used to construct the reweighted histogram for some measured observable of interest \( O \):

\[
H(O \mid \beta_1, \mu_1) = \sum_{j=0}^{M} w_j \delta(O - O_j).
\]

After normalization this histogram is a discrete estimate for the probability distribution \( p(O \mid \beta_1, \mu_1) \). Implicit in its construction is the specification of the bin width, which may need to be tuned to strike a balance between resolution and data smoothness.

IV. TRACING COEXISTENCE CURVES

We now turn to the actual simulation procedure by which a fluid phase boundary can be determined. For illustration we apply the procedure to the liquid–gas transition of the Lennard-Jones fluid, although the essential approach is rather general and can be applied to other phase transitions such as demixing transitions in fluid mixtures.

One of the key components of our procedure is multicanonical preweighting, the use of which generally entails some preliminary effort to determine a suitable preweighting function. It transpires, however, that if we begin tracing the phase boundary from near the critical point, no additional
effort needs to be expended to determine the preweighting functions. Instead they can be obtained for “free” by virtue of histogram reweighting.

We start by determining the approximate location of the critical point from a series of short runs on a small system. We choose a temperature $T$ and a large negative value of $\mu$ and do a short simulation without multicanonical preweighting. The fluctuating density will typically settle down rapidly, and its average value can be estimated visually from the data output. We then repeat this procedure for a succession of progressively larger (less negative) $\mu$ values. As $\mu$ is increased, the average particle number will increase steadily. However, on traversing the hysteresis-shifted phase boundary curve $\mu_b(\beta)$, a sudden jump will occur in the density. If this jump is large, for example, from $\rho = 0.05$ to $\rho = 0.6$ (in units of $\sigma^{-3}$), then the temperature is well below the critical temperature and we should increase $T$ somewhat and begin the procedure again. If the jump is smaller, for example, from $\rho = 0.2$ to $\rho = 0.4$, then we have obtained an estimate for a near-critical point on the phase boundary. If no jump in density is observed, then the temperature probably exceeds the critical temperature and should be reduced.

We next perform a longer run for a larger system at the estimated near-critical phase boundary point, $\beta_c, \mu_c$. Because the surface tension and the associated barrier to interphase crossings are low near criticality, it should be possible to determine $p(N,E)$ (including information on states “between the peaks”) without resort to multicanonical preweighting. The next step is to reweight the data accumulated from this run to obtain an estimate of $p(N)$ at some lower temperature point on the phase boundary. This estimate is achieved by first choosing an extrapolation temperature $\beta_1$ inside the range of reliable reweighting so that the reweighted distribution $p_{\text{ex}}(N|\beta_1, \mu_c)$ appears smooth. We then tune $\mu$ within the histogram reweighting method until $p_{\text{ex}}(N|\beta_1, \mu)$ is bimodal with equal areas under each peak. This tuning procedure can be easily automated within the analysis program to deliver precise values of the phase boundary chemical potential $\mu_1 = \mu_{b}(\beta_1)$.

The reweighted phase boundary histogram $p_{\text{ex}}(N|\beta_1, \mu_1)$ will, on account of the lower temperature, be more strongly peaked than that from which it derives. Thus, multicanonical preweighting will probably be necessary for a new simulation at $\beta_1, \mu_1$. However, a suitable preweighting function is already to hand—its just the extrapolated function $p_{\text{ex}}(N|\beta_1, \mu_1)$. All we need do is set $\eta(N) = p_{\text{ex}}(N|\beta_1, \mu_1)$ and perform a multicanonical simulation at $\beta_1, \mu_1$ to obtain the actual distribution $p(N,E|\beta_1, \mu_1)$.

We then simply iterate this procedure: histogram reweighting of $p(N,E|\beta_1, \mu_1)$ is used to estimate a phase boundary point $\mu_2 = \mu_{b}(\beta_2)$ at $\beta_2$, together with the extrapolated distribution $p_{\text{ex}}(N|\beta_2, \mu_2)$. The latter serves as a preweighting function for a further simulation at $\beta_2, \mu_2$ and so on. In this way we step down the coexistence curve, obtaining at the same time the locus of the phase boundary $\mu_{b}(\beta)$ and the associated set of density probability density functions. Clearly the maximum feasible step size is set by the range of reliable histogram extrapolation, which decreases with increasing system size. In practice, however, quite large steps can be made even for large systems. For example, in Ref. 12 a system of approximately 600 LJ particles was studied and 7 steps were required to reach a subcritical temperature of $T = 0.85T_c$.

The results of implementing this procedure for the Lennard-Jones fluid are depicted in Fig. 5. The forms for $p(N)$ have equal weight in each peak and hence lie on the phase boundary. The associated phase diagram $\mu_b(\beta)$ is shown in Fig. 6. The values of the coexisting densities can be read off from the peak positions in Fig. 5(a). The enormity of the probability barrier that multicanonical preweighting allows us to negotiate is revealed by plotting these distributions on a log scale, as shown in Fig. 5(b).

Finally, we point out that the measured coexistence form of the density distribution $p(\rho)$ permits an estimate of the surface tension $\gamma$. For a cubic system of volume $L^3$, $\gamma$ is found from the ratio of values of $p(\rho)$ at the peak and at the trough between the peaks.

$$\gamma = \frac{1}{2\beta L^2} \ln \left( \frac{p_{\text{max}}}{p_{\text{min}}} \right).$$ (14)

Reference 18 describes a recent application of this relation in a study of the surface tension of the Lennard-Jones fluid.
V. DISCUSSION AND CONCLUSIONS

We have described a method whereby information on the locus of a fluid phase boundary is obtainable via multicanonical preweighting and histogram reweighting, in conjunction with a standard grand canonical simulation algorithm. This use of multicanonical preweighting completely eliminates hysteresis at a first-order phase transition.

In cases where we do not wish to start tracing a coexistence curve from near the critical point, it is necessary to bootstrap the procedure by obtaining an initial phase boundary preweighting function. A variety of techniques exist, ranging from simple extrapolation of the weight function to the unsampled region, to more sophisticated analyses of MC transition probabilities. For further details we refer the reader to the literature.15,21

Although we have illustrated our approach in the context of a simple fluid model, it is equally applicable to complex fluids such as molecules or polymers. In these systems, the grand canonical insertion probability is often small, and it is necessary to supplement the standard algorithm with a more intelligent insertion scheme—one that performs a biased choice of a molecular orientation favorable for the insertion. Methods such as configurational bias Monte Carlo22 and recoil growth23 allow us to do this. Apart from this added complication, multicanonical preweighting and histogram reweighting are implemented as for a simple fluid.

The method we have described is generalizable to other simulation ensembles such as the constant NpT ensemble. This ensemble can be more efficient than the grand canonical when dealing with very dense fluids where the success rate of particle insertions and deletions is small. The natural variable in which to preweight a constant NpT ensemble simulation is the fluctuating volume, but otherwise the formalism is similar to that we have described. For other types of phase transitions, such as the liquid–liquid transitions occurring in binary fluid mixtures, a suitable variable in which to preweight is usually the order parameter for the transition, for example, the concentration of one species.

Despite their utility in dealing with phase equilibria involving fluids, the specific multicanonical techniques we have described do not permit us to tackle phase transitions involving solids. The problem is that when attempting to traverse mixed-phase states involving crystalline order, the simulation invariably gets caught in nonergodic traps identifiable with defective crystalline configurations. Recently, a new technique has been developed that circumvents this problem by linking the two coexisting phases without traversing mixed-phase states. The method can be thought of as leaping directly from the configuration space of one pure phase to that of the other. Again use of multicanonical sampling is necessary, but in this case its role is to encourage the simulation to visit a subset of configurations (in each pure phase) from which a leap to the other phase will be accepted. This new method has recently shown its worth in studies of solid-phase free energy differences and hard sphere freezing.24,25

VI. SUGGESTIONS FOR FURTHER STUDY

1. Using Box 1 as a guide, write a simple grand canonical ensemble MC program to simulate the Lennard-Jones fluid with potential cutoff at r_c = 2.5σ. Do not correct for the potential truncation. Have the program display the particle number and the energy in list form (see Sec. III C). Further programming details can be found in Ref. 2.

2. Using a volume not exceeding V = 1000σ^3, run your program at the near critical phase boundary parameters, βε = 1.1876 and βμ = 2.778, saving the output data to a file. Make sure you run the program for a sufficiently long time for the density fluctuations to be adequately sampled.

3. Write a post-processing program to construct the number histogram p(N) from the raw data list.

4. Modify your grand canonical acceptance probabilities to take advantage of multicanonical preweighting (see Secs. III A and III C). Use your measured p(N) as the preweighting function for a multicanonical simulation at the same values of β and μ used in Problem (2). At the extrema of small (large) N, there will be histogram bins in p(N) with zero entries. Before using p(N) as your preweighting function, set these entries to a constant equal to the smallest nonzero entry to avoid possible division by zero in the acceptance probabilities.

5. Extend your post-processing program to unfold the effects of the preweighting (as described in Secs. III A and III C) to find p(N). Check that the form of p(N) thus obtained agrees with that found without multicanonical preweighting. By plotting p against the number of MC steps, compare visually the correlation time for the sampling processes with and without multicanonical preweighting (cf. Fig. 4).

6. Further extend your post-processing program to implement histogram reweighting (as described in Secs. III B and III C). Extrapolate the data obtained at the critical temperature β_c = 1.1876 to find the location of the phase boundary and the form of p(N) at βε = 1.17. Use this extrapolation as a preweighting function in a new multicanonical simulation at the new phase boundary state point. Compare your results with those of Ref. 12.

7. Experiment to discover qualitatively how the range of reliable histogram extrapolation depends on the duration of the simulation and the volume V.
1A comparison of the model’s phase behavior with the results of experiment can be useful in helping to refine the model parameters.


7To make the probability dimensionless, one should include a factor of $\lambda^3$, where $\lambda$ is the thermal de Broglie wavelength. This formula assumes units such that $\hbar = 1$.

8It can be shown that the constrained free energy $F(\rho)$ is directly related to the density probability density function by $F(\rho) = -\ln p(\rho)$ [see, for example, A. D. Bruce, “Universality in the two-dimensional continuous spin model,” J. Phys. A 18, L873–L875 (1985)]. This relation links the thermodynamic interpretation of phase stability in which a system occupies the state of minimum free energy with the statistical mechanics viewpoint that it is most likely to be found in a state of maximum a priori probability.

9The critical point is characterized by strong fluctuations in the density that present distinctive challenges for simulation and theory alike. For an introduction to some of the main issues, see Chapter 5 of Ref. 6 and also Ref. 12.


16In fact it is possible to preweight with respect to any system observable, although for the problem of fluid phase transitions, the density is usually the most computationally appropriate quantity.


19We can also extrapolate to slightly higher temperatures to pinpoint the critical point. Special finite size scaling techniques are required here, as described in detail in Ref. 12.


**NATURAL SELECTION**

Natural selection, like the interpretation of the Second Law, is one of the most beautiful theories of science, for it is so economical yet powerful. That is, it is simple, yet admits as a consequence complexity. That is the hallmark of high science and the apotheosis of reductionism. From a simple precept (e.g. the unconscious tendency of genes to survive in the face of unconscious competition) the whole gamut of organisms in the creation can be seen to emerge. It is then entirely gratuitous (but science cannot demonstrate that it is incorrect) to suppose that evolution has been guided teleologically or that matter had some in-built tendency to aggregate complexity. Both speculations are entirely possible, just as other profligate speculations are also possible, such as the unsupported musings that this universe is constantly splitting into others that have no communication with us, or that there is a teapot in orbit around Mars.