

Direct calculation of solid-liquid equilibria from density-of-states Monte Carlo simulations

Ethan A. Mastny and Juan J. de Pablo

Department of Chemical and Biological Engineering, University of Wisconsin-Madison, Madison, Wisconsin 53706-1691

(Received 19 November 2004; accepted 26 January 2005; published online 31 March 2005)

A density-of-states Monte Carlo method is proposed for simulations of solid-liquid phase equilibria. A modified Wang–Landau density-of-states sampling approach is used to perform a random walk in regions of potential energy and volume relevant to solid-liquid equilibrium. The method provides a direct estimate of the relative density of states $[\Omega(U, V)]$ and thus the relative free energy within these regions, which is subsequently used to determine portions of the melting curve over wide ranges of pressure and temperature. The validity and usefulness of the method are demonstrated by performing crystallization simulations for the Lennard-Jones fluid and for NaCl. © 2005 American Institute of Physics. [DOI: 10.1063/1.1874792]

I. INTRODUCTION

The melting/freezing phase transition is probably the most familiar of first-order transitions, as it is ubiquitous in nature. Solid-liquid equilibria are of fundamental importance to a variety of technologies in separation of chemicals, pharmaceutical development, and semiconductor processing. Solid state properties and melting transitions are also of fundamental interest in the earth, atmospheric, and space sciences.

Theoretical methods for prediction of solid-fluid equilibria have lagged behind their counterparts for the study of fluid phase equilibria. Molecular simulations have provided valuable insights into the structure and thermophysical properties of solids, but most simulations to date have not been able to provide direct predictions for the freezing transition of a liquid into a solid. For pure systems, prediction of freezing points requires finding a point where the temperature, pressure, and chemical potential (T , P , and μ) of the solid and liquid phases are identical. The chemical potential (or free energy) cannot be measured directly or instantaneously in simulations; the relative free energy of the two phases must therefore be calculated by devising iterative, multiple-step thermodynamic constructs.

Previous simulation reports of solid-liquid equilibria have relied on laborious and computationally demanding thermodynamic integration techniques. Such methods rely on precise knowledge of the free energy at carefully chosen reference states for both the solid (e.g., an Einstein crystal) and the liquid (e.g., an ideal gas). Numerous, lengthy simulations are then required to calculate the free energy difference between the reference state and the system of interest, until a point is identified where both the solid and the liquid phases have identical temperature, pressure, and chemical potential.¹ The computational demands of performing thermodynamic integration as well as the more challenging task of identifying and applying a valid integration path (free of singularities) are such that only a handful of systems have been studied to date.

Having identified a particular melting point, a Gibbs–Duhem integration² can subsequently be used to trace a melting curve. Gibbs–Duhem integrations as well as thermodynamic integrations are inherently serial and can suffer from an incorrect initial coexistence point (which would lead to an incorrect coexistence line). Errors can be compounded by inadequate numerical integration of the Gibbs–Duhem equation.³

Melting point data have also been generated through dynamic simulations, where two phases are put in contact and allowed to equilibrate to the melting temperature.¹ While this method can be applied to simple systems (such as Lennard-Jones) with large numbers of particles, it is not well suited for complex molecules or systems where interfacial effects can be significant.

Most direct methods for calculating melting curves to date require an interface or an initial coexistence point; an exception is provided by the so-called phase-switch Monte Carlo (MC) technique.^{4,5} This method can be used to sample both crystalline and fluid phases in a single simulation through biased sampling techniques. It provides individual melting points directly, but knowledge of the solid phase configuration is required and its implementation can be demanding.

In this work, a novel density-of-states (DOS) method is proposed for direct simulation of solid-liquid phase transitions. The method relies on a two-dimensional, multiple-replica variant of the density-of-states sampling technique originally proposed by Wang and Landau.^{6–8} This density-of-states method finds a system's fundamental equation $[S(U, V)]$, spanning a region of phase space that encompasses solid and liquid equilibrium states, over a wide range of temperatures and pressures. In addition, this method is conceptually simple, can be readily implemented for a variety of systems, and, as shown in this work, is computationally efficient. Its usefulness is demonstrated in the context of the Lennard-Jones fluid and NaCl; these are systems for

which melting results from thermodynamic integration are available, thereby permitting a direct comparison to more established techniques.

This density-of-states method relies on sampling regions of energy and volume space where the solid and liquid phases are unstable, thereby providing a measure of the relative free energy of the two phases. To connect the free energy of the solid and liquid phases, there must exist a portion of energy and volume space that can be simultaneously sampled by both solid and liquid phases.

II. METHODS

A. Model

In the Lennard-Jones fluid pairs of particles interact through a potential energy function of the form

$$\phi_{LJ}(r) = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right], \quad (1)$$

where σ and ϵ represent characteristic size and energy parameters, respectively. A cutoff radius of $r_c/\sigma=2.5$ was applied for all simulations presented in this work. Standard long-range corrections were applied to determine the total potential energy of the system. Each simulation consisted of either 256 or 500 particles arranged in a cubic box with periodic boundary conditions.

As an example of a more complex system, NaCl, an ionic liquid is also considered in this work. The ionic pair-pair interaction potential is of the Born–Mayer–Huggins–Fumi–Tosi form

$$\phi_{ij}(r) = A_{ij} \exp[B(\sigma_{ij} - r)] - \frac{C_{ij}}{r^6} - \frac{D_{ij}}{r^8} + \frac{q_i q_j}{\tau}. \quad (2)$$

The parameter values and a review of their validity for predicting NaCl solid state properties have been discussed recently by Anwar *et al.*⁹ All NaCl simulations are performed with 512 particles (256 Na–Cl pairs) in a cubic box with periodic boundary conditions. Ewald summations are used to evaluate the long-range component of Coulombic interactions. All non-Coulombic terms were cutoff at 9.0 Å and corrected using standard tail-correction expressions.

B. General simulation method

In the density-of-states sampling technique,⁷ configurations are generated with probability

$$p(U) \propto \frac{1}{\Omega(U)}, \quad (3)$$

where $\Omega(U)$ is the density-of-states for potential energy U . A running estimate of $\Omega(U)$ is gradually improved as the simulation proceeds.

For simulations of solid-liquid equilibria, Ω must be determined as a function of both energy and volume; the resulting $\Omega(U, V)$ function can then be used to determine the precise location of phase boundaries (and any other thermodynamic property of interest). The extension of density-of-states sampling to multiple dimensions (e.g., en-

ergy and number of particles) has been discussed in the literature and need not be described in detail here.¹⁰

A random walk in energy and volume space is implemented whose ultimate goal is to sample all sets of (U, V) uniformly or with probability $1/\Omega(U, V)$. Simple coordinate rescaling trial volume moves can be used in addition to particle displacement moves. In this formulation, a trial move is accepted or rejected according to

$$p_{acc}(U_1, V_1 \rightarrow U_2, V_2) = \min \left[1, \left(\frac{V_2}{V_1} \right)^N \frac{\Omega(U_1, V_1)}{\Omega(U_2, V_2)} \right]. \quad (4)$$

A running estimate of Ω is updated on the fly; every time a distinct point in (U, V) space is visited, the corresponding $\Omega(U, V)$ is updated according to $\Omega(U, V) = f \times \Omega(U, V)$, where f is an arbitrary convergence factor which tends to unity as the calculation proceeds. A histogram $h(U, V)$ records visits to distinct energy and volume states. When that histogram becomes sufficiently flat, it is erased, the value of f is modified according to $f = \sqrt{f}$, and the calculation proceeds starting with a new (i.e., empty) histogram. This sampling method satisfies detailed balance with an accuracy proportional to $\ln(f)$.⁷ Therefore, as $f \rightarrow 1$ detailed balance is essentially satisfied.

To sample the regions of phase space that are relevant to melting, relatively wide ranges of energy and volume must be explored. To facilitate convergence and to minimize the time spent by the random walker on “easy” regions of phase space, the thermodynamic space of interest is subdivided into overlapping windows; the density-of-states for each window can then be evaluated and scaled to find the density-of-states over the desired, wider region of thermodynamic space. Following previous reports,^{6,8} the DOS algorithm can be merged with a parallel sampling scheme. As the density-of-states simulations are performed in independent windows, a swap move between the systems in overlapping windows is implemented. This move is accepted if both the energy and volume of both systems lie within the overlapping portion of the windows when the move is proposed.

Once the density of states is known, thermodynamic properties such as the entropy S , potential energy $U(T, P)$, and heat capacity $C_v(T, P)$ can be calculated.⁷ The probability of being in a given (U, V) state for a given temperature and pressure is found by Eq. (5),

$$p(U, V) = \frac{\Omega(U, V) e^{-\beta U} e^{-\beta P V}}{\sum \sum \Omega(U, V) e^{-\beta U} e^{-\beta P V}}. \quad (5)$$

C. Simulation details

1. Initialization

At each stage of the algorithm described above, the goal is to sample a predefined (U, V) space uniformly. Selection of appropriate bounds of the (U, V) space to be sampled is important for the discernment of melting data. The following procedure was used in this work (see Fig. 1 for an example).

(i) Guess a set of melting temperatures that correspond to a set of pressures where equilibrium data are desired.

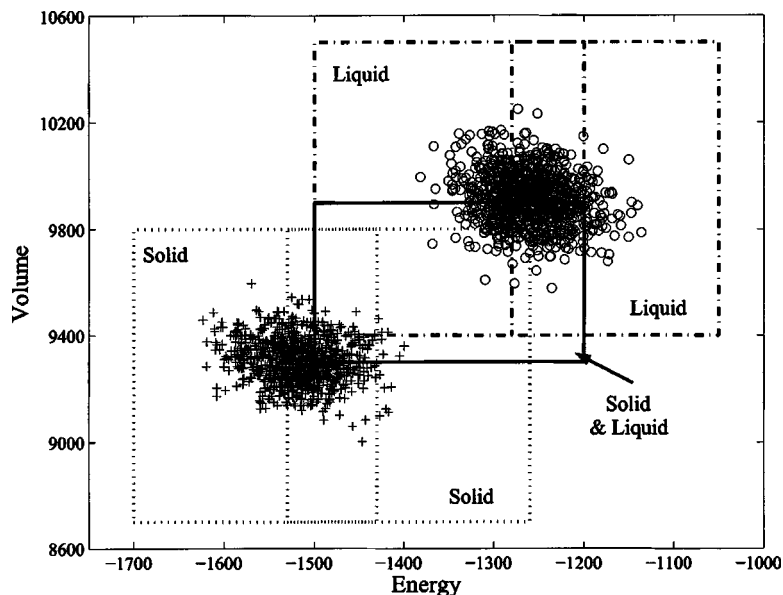


FIG. 1. The crosses in this figure denote locations in (U, V) space sampled by a short NPT simulation of a solid phase. Likewise the circles denote locations sampled by a NPT simulation of a liquid phase at the same temperature and pressure. The traced boxes denote the boundaries of overlapping simulation windows. Each window is labeled with the particular phase the window was initialized with. The box in the middle labeled, "Solid & Liquid" comprises two completely overlapping simulation windows, one initialized with a solid phase and the other with a liquid phase. This initialization was used for melting point predictions of a 256 particle Lennard-Jones system in the vicinity of $T^* = 2$.

(ii) Run short simulations, in the constant temperature and pressure ensemble (NPT), for both the solid and liquid phases, for each T, P pair. Examine the (U, V) space explored by each equilibrated phase.

(iii) Create a set of overlapping windows that encompass the (U, V) space observed in the NPT simulations sampled by each phase.

(iv) Create initial configurations from short NPT simulations of both solid and liquid phases. Solidlike initial conditions are placed in windows with lower energy and volume, and fluidlike initial conditions are placed in windows with high energy and volume.

It is advantageous for simulation efficiency to initialize each window with a structure that corresponds to the phase (solid or liquid) that is anticipated for that energy (and volume) range. For the systems studied in this work (with simple crystal structures), we have found that this DOS algorithm can transform the initial phase to the appropriate phase that corresponds to a given window, even if the window was initialized improperly.

2. Merging

In this simulation method no attempt is made to try to relate the density of states to a reference state of known value. Therefore Ω is only known within a multiplicative constant, or $\ln \Omega$ is only known to within an additive constant. Since density-of-state estimates are accumulated independently in each window, an additive constant must be added to each $\ln \Omega$ estimate to obtain a continuous $\ln \Omega$ surface over the entire (U, V) space. This constant is chosen by attempting to best fit the shape of the $\ln \Omega$ surface in adjacent windows in the overlapping regions.

3. Piecewise linear construction

Ω and $\ln \Omega$ are a continuous functions of energy and volume. This algorithm requires that the (U, V) space be divided into discrete bins. The values of $\ln \Omega$ accumulated in each bin during the simulation are assumed to correspond to

a single (U, V) pair centered in the bin, and not all (U, V) pairs within the bin. If $\ln \Omega$ is essentially linear within a given bin, a piecewise-linear construction of $\ln \Omega$ is a good approximation of the continuous $\ln \Omega$ function. Therefore, when applying the acceptance criteria in Eq. (4), $\ln \Omega(U_1, V_1)$ is determined by performing a two-dimensional linear interpolation of the $\ln \Omega$ values surrounding (U_1, V_1) , and likewise for (U_2, V_2) .

Using this linear interpolation technique greatly improves the simulation efficiency. This ensures that (U, V) space is also sampled uniformly within a particular bin. Linear interpolation also makes it easier to select the bin size: the bins need to be small enough to ensure that the local linear approximation of $\ln \Omega$ is valid.

III. RESULTS AND DISCUSSION

Equation (5) gives the probability of finding the system in a state (U, V) . At coexistence, two distinct points in phase space have different values of U and V but are equally likely to occur, i.e., they exhibit the same probability $p(U, V)$. Melting points can therefore be determined at any given pressure (or temperature) by reweighting $\Omega(U, V)$ until the temperature (or pressure) is found where the probability of each phase is equal.

Figure 2 shows the probability distribution as a function of potential energy and volume, for 256 Lennard-Jones particles, at one estimated melting point ($T^* = 2.03$ and $P^* = 21.5$), calculated by Eq. (5). Each peak shown in Fig. 2 has equal probability; this bimodal distribution therefore represents a melting point.

Figures 3(a) and 3(b) show the results of DOS melting curve predictions for the Lennard-Jones system (500 particles) for two different portions of the melting curve. Each solid line shows the results of an independent DOS simulation, initiated with a different random number seed. These data are compared with those of Agrawal and Kofke, who calculated the melting curve using Gibbs–Duhem integration from a previously available melting point.¹¹ The data in Fig.

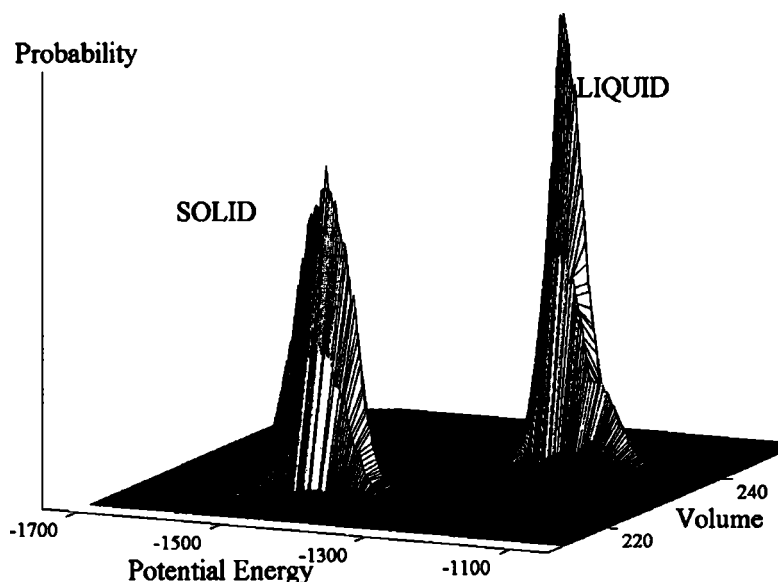


FIG. 2. The surface shown here is the probability distribution of the Lennard-Jones system at a melting point ($T^* = 2.03$ and $P^* = 21.5$). Each peak has equal probability, therefore the solid and liquid phases are in equilibrium.

3(a) are also compared to a point generated from a recent study by Errington using phase-switch Monte Carlo.⁴

The DOS melting curve predictions for both portions of the melting line are in good agreement with previous studies of the Lennard-Jones system. The slight disagreement between our results and those of Agrawal could be due to a slightly inaccurate reference point employed by those authors for their calculations.⁴ Each DOS curve in Fig. 3 represents results from an independent simulation; the spread of these curves therefore provides a measure of the uncertainty of our calculations. Simulations of 256-particle systems yield melting curves with values similar to those for the 500-particle systems. The statistical uncertainty of the 500-particle systems, however, is approximately two to three times smaller. While most studies of melting have not quantified statistical uncertainty (partly as a consequence of the computational demands of existing simulation techniques), we believe that the accuracy of the proposed simulation scheme is highly satisfactory.

Figure 4 shows the melting curves calculated from five independent DOS simulations for the NaCl system. At ambient pressure (0.1 MPa) the coexistence point is calculated to be 1050.2 ± 2.7 K. This is approximately at 2% below the experimental melting point (1074 K).¹² Also at ambient pressure we find that $dP/dT = 3.1$ MPa/K as calculated from DOS simulations. These results are consistent with a recent calculation of the melting point of NaCl by thermodynamic integration, which at -30 ± 40 MPa places it at 1074 K, with $dP/dT \approx 3$ MPa at that point.⁹ Despite a small offset from experimental value of the melting point at ambient pressure, we find the accuracy and precision of the DOS method to be satisfactory.

The source of uncertainty in our calculations arises from errors in the estimate of the density of states. These errors occur because we assume that a histogram is “sufficiently flat” when the minimum entry in the histogram is no less than 80% of the average value.⁸ Density-of-state estimates

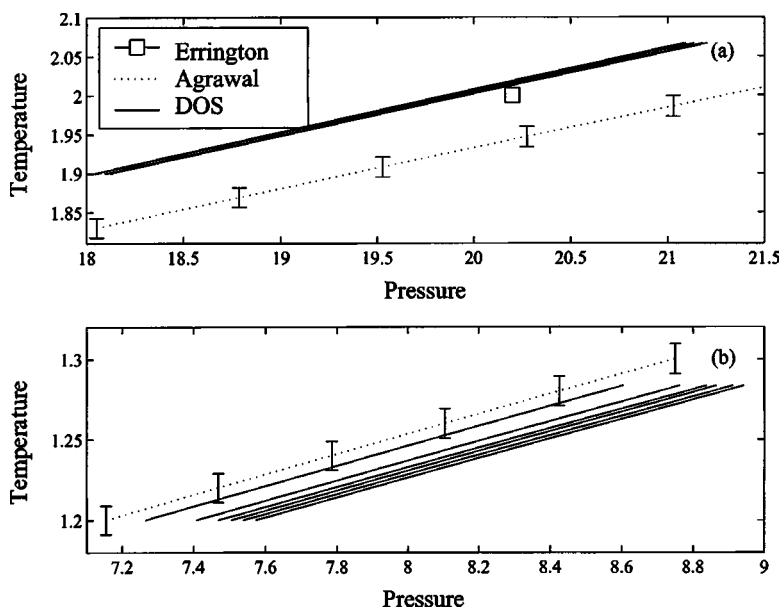


FIG. 3. The solid lines represent melting curves of the Lennard-Jones system predicted from six independent density-of-states simulations. The dotted line was traced from Gibbs–Duhem integration, and the box represents a melting point generated from phase-switch simulations. Box (a) represent a portion of the Lennard-Jones melting curve near $T = 2$ and box (b) is another portion of the same melting curve near $T = 1.25$.

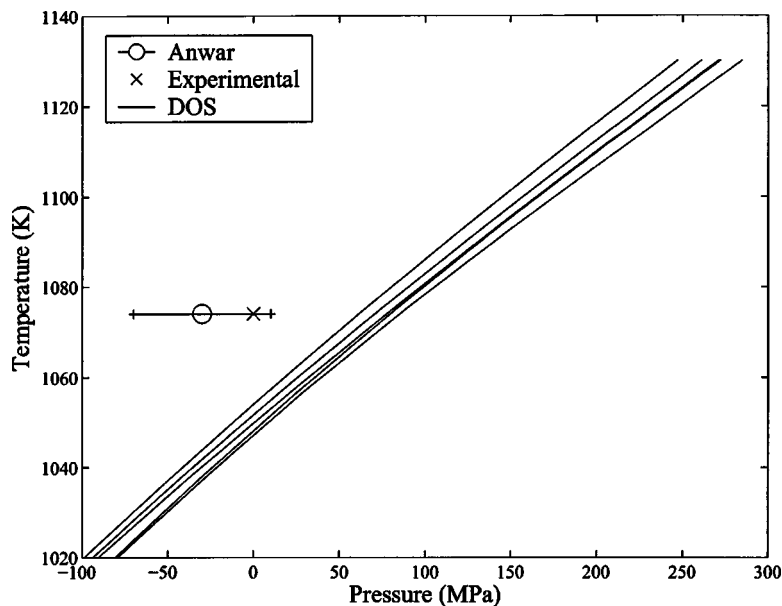


FIG. 4. The solid lines are melting curves of the NaCl system generated from five independent density-of-states simulations. The point marked with an x is the experimental melting point, and the point marked with the circle (and error bounds) is the melting point predicted by thermodynamic integration.

with minor errors can result in a “sufficiently flat” histogram. More stringent flatness criteria would help alleviate this problem, but at the expense of higher computational demands. We find that a 80% flatness criteria provides a good balance between efficiency and accuracy. To calculate statistical errors, we estimated the density of states with several identical simulations, but with different random number seeds. The statistical error in these calculations may be significantly reduced by implementing the configurational temperature method of Yan *et al.*⁸ but in the context of a configurational pressure.

According to our calculations for the Lennard-Jones system, Gibbs–Duhem integration from an already available reference point would require approximately twice the simulation time for results of much lower accuracy. More importantly, in cases where a reference point is not available, considerable effort must be devoted to the calculation of that point (the quality of the ensuing coexistence results hinge critically on precise knowledge of that reference state). An additional limitation of integration methods is that they are inherently serial; the calculation of a new state cannot be pursued until the prior state point calculation is complete. In contrast, the DOS method pursued in this work is inherently parallel, thereby facilitating its implementation with parallel processors.

Also according to our estimates, the phase-switch MC method proposed by Errington produces a single phase coexistence point (of comparable accuracy to that generated by DOS) in approximately three times the computational requirements. In addition to computational efficiency, the main advantage of DOS over phase switch is the relative simplicity of the method and the fact that an underlying lattice structure need not be specified.

The results for NaCl by Anwar *et al.* were generated by thermal integration; one of the main challenges of those calculations was to identify a suitable thermal path for that integration. Those authors report that three different paths were explored for the liquid phase before adopting the procedure described in their work. This challenge does not arise in

DOS calculations. Additionally Anwar’s simulations were more than an order of magnitude more demanding than DOS simulations.

The Lennard-Jones fluid and NaCl systems are the only systems we have applied this methodology to; however, we do not anticipate any significant problems in extending this technique to other systems. Note that, as described here, DOS Monte Carlo is applicable to systems where the solid and liquid phases have the same composition; extensions to multicomponent systems would require working with more computationally demanding multidimensional histograms.

IV. CONCLUSIONS

A many-replica, multidimensional density-of-states method has been proposed for direct simulation of solid-liquid phase transitions. This method uniformly explores the energy and volume space relevant to both phases. Compared with currently available methods, this density-of-states algorithm is simple in concept and application and can be used to generate solid-fluid phase equilibria data relatively rapidly. For the Lennard-Jones system, estimates of the density of states as a function of energy and volume appear to yield entire portions of the Lennard-Jones melting curve with significantly less effort than that required by other techniques. The results of the simulations presented here are in good agreement with literature reports. For NaCl, the solid-liquid coexistence curve determined from DOS simulations is consistent with the melting point recently predicted by thermodynamic integration.⁹ Together, the results presented in this work for a simple and an ionic fluid suggest that the DOS method provides a viable alternative for simulation of solid-fluid phase behavior; we are currently pursuing applications to anisotropic crystals, and the results will be presented in a future publication.

ACKNOWLEDGMENTS

This material is based upon work supported under a National Science Foundation Graduate Research Fellowship.

Financial support from the Department of Energy (DE-FG02-99ER14961) is gratefully acknowledged.

¹J. R. Morris and X. Song, J. Chem. Phys. **116**, 9352 (2002).

²R. Agrawal and D. A. Kofke, Mol. Phys. **78**, 1331 (1993).

³D. Frenkel and B. Smit, *Understanding Molecular Simulation From Algorithms to Applications* (Academic, New York, 2002).

⁴J. R. Errington, J. Chem. Phys. **120**, 3130 (2004).

⁵N. B. Wilding and A. D. Bruce, Phys. Rev. Lett. **85**, 5138 (2000).

⁶N. Rathore, T. A. KnottsIV, and J. J. de Pablo, J. Chem. Phys. **118**, 4285 (2003).

⁷F. Wang and D. P. Landau, Phys. Rev. Lett. **86**, 2050 (2001).

⁸Q. Yan and J. J. de Pablo, Phys. Rev. Lett. **90**, 035701 (2003).

⁹J. Anwar, D. Frenkel, and M. G. Noro, J. Chem. Phys. **118**, 728 (2003).

¹⁰Q. Yan, R. Faller, and J. J. de Pablo, J. Chem. Phys. **116**, 8745 (2002).

¹¹R. Agrawal and D. A. Kofke, Mol. Phys. **85**, 43 (1995).

¹²J. Akella, S. N. Vaidya, and G. C. Kennedy, Phys. Rev. **185**, 1135 (1969).