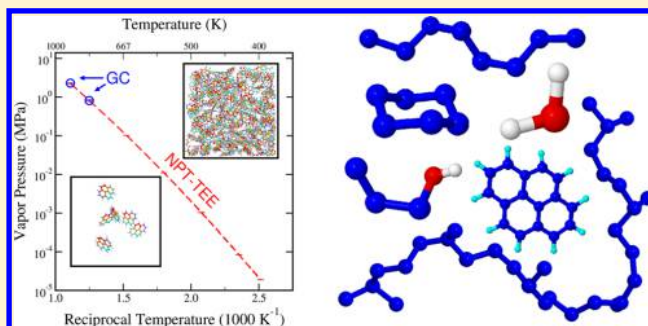


Monte Carlo Simulation Methods for Computing Liquid–Vapor Saturation Properties of Model Systems

Kaustubh S. Rane, Sabharish Murali, and Jeffrey R. Errington*

Department of Chemical and Biological Engineering, University at Buffalo, The State University of New York, Buffalo, New York 14260-4200, United States

ABSTRACT: We discuss molecular simulation methods for computing the phase coexistence properties of complex molecules. The strategies that we pursue are histogram-based approaches in which thermodynamic properties are related to relevant probability distributions. We first outline grand canonical and isothermal–isobaric methods for directly locating a saturation point at a given temperature. In the former case, we show how reservoir and growth expanded ensemble techniques can be used to facilitate the creation and insertion of complex molecules within a grand canonical simulation. We next focus on grand canonical and isothermal–isobaric temperature expanded ensemble techniques that provide a means to trace saturation lines over a wide range of temperatures. To demonstrate the utility of the strategies introduced here, we present phase coexistence data for a series of molecules, including *n*-octane, cyclohexane, water, 1-propanol, squalane, and pyrene. Overall, we find the direct grand canonical approach to be the most effective means to directly locate a coexistence point at a given temperature and the isothermal–isobaric temperature expanded ensemble scheme to provide the most effective means to follow a saturation curve to low temperature.



1. INTRODUCTION

Chemical technology is evolving rapidly, with the frequent discovery of new designer molecules for consumer products and therapeutics. Knowledge of liquid–vapor coexistence properties over a range of conditions is important for many of these technological advances. Molecular simulation can play an important role in the rational design of new compounds and in studying the molecular aspects of poorly understood phenomena. In this paper, we present an efficient, accurate, and precise Monte Carlo simulation-based approach for computing liquid–vapor phase equilibrium of complex molecules over a broad range of conditions.

The goal of this work is to introduce general modeling strategies that enable one to compute the phase coexistence properties of complex fluids. In this context, examples of complex fluids include systems with complex intermolecular interactions (e.g., polar molecules, ionic liquids) and/or complex molecular topologies (e.g., branched and/or cyclic hydrocarbons, asphaltenes). We pursue a strategy in which thermodynamic properties are deduced from histograms of relevant thermodynamic variables.^{1,2} One of the main advantages of this approach stems from the ability to reweight histograms to conditions different from those originally simulated, thereby providing a means to increase the utility of simulation data. Such a strategy has been employed with increasing frequency over the last two decades.^{3–23} Here, we address some of the limitations of this general approach.

Within the histogram-based strategy, one generally deduces liquid–vapor saturation properties from a density probability

distribution collected over a range of densities representative of both the vapor and liquid phases.² grand canonical approaches,^{2,3} in which density fluctuations are facilitated via particle transfers, and isothermal–isobaric approaches,³ in which density fluctuations are facilitated via volume changes, have been introduced. We first address limitations in applying the grand canonical version of this approach to complex flexible molecules. In this case, one often encounters difficulties in transferring molecules to and from the system. Specific issues include (1) generating representative molecular configurations “on-the-fly” and (2) low exchange probabilities encountered within the liquid phase. To address these issues, we work with a strategy that relies upon a “reservoir” grand canonical scheme⁷ and an expanded ensemble growth procedure.²⁴ The former technique provides a collection of molecules from which one can readily draw from during a molecular insertion and the latter technique provides a means to insert/remove molecules from the system in a stagewise manner.

The second general limitation of histogram-based methods that we address in this work is related to the relative difficulty in applying the conventional scheme at moderate to low temperatures. Under such conditions, one often experiences difficulties in sampling the inhomogeneous region between the vapor and liquid phases. Within this region, the system forms structures such as spherical- and cylindrical-shaped droplets and cavities, which promote sampling difficulties related to the low

Received: January 28, 2013

Published: May 17, 2013



transition probabilities between these structures.^{25,26} Moreover, if one wishes to construct a phase diagram over a wide range of temperatures, it simply becomes inefficient to apply the conventional histogram-based approach at many temperatures of interest. In an effort to address these issues, here we describe a means to trace saturation curves over a wide range of temperatures.^{27,28} The approach is similar in spirit to the Gibbs–Duhem approach of Kofke,^{29,30} but this approach allows one to tackle the entire saturation curve in a single simulation.

The determination of liquid–vapor phase coexistence properties is a long-standing problem within the field of molecular modeling and enjoys a rich history. Perhaps the most commonly used strategy for computing the liquid–vapor saturation properties of complex molecules involves combining the Gibbs ensemble Monte Carlo algorithm^{31,32} with configurational-bias techniques.^{33–35} Panagiotopoulos introduced the Gibbs ensemble scheme as a general means for computing phase coexistence properties, and it has proven to be quite robust. Its key advantage is that one can deduce saturation properties without establishing an explicit interface between coexisting phases and without having to traverse the inhomogeneous region between the vapor and liquid phases. In previous work, we compared the efficacy of the Gibbs ensemble strategy relative to a histogram-based approach and found that the Gibbs ensemble provided more-precise estimates of saturation properties over short simulation periods, but that the histogram-based approach proved superior for longer simulation runs.¹⁰ One key advantage of histogram-based approaches, which is particularly relevant for our group, is the ability to apply these methods to study interfacial phenomena.^{12,36,37} Although there are some examples of the application of the Gibbs ensemble method to interfacial systems,^{38–40} it is not always clear how to apply this technique to probe surface phase behavior.

Configurational-bias strategies are commonly used to facilitate the exchange of molecules within grand canonical and Gibbs ensemble simulations. Akin to the Gibbs ensemble, these methods have proven useful in a wide range of problems.^{41–43} The technique considerably increases the probability of inserting a complex molecule into a relatively dense liquid. With this scheme, one is not required to maintain a separate reservoir (as is done within the approach pursued here) and molecules are inserted/removed from a system in a single stage. The drawback of the method is that it becomes progressively more difficult to implement as the complexity of the molecular topology increases. More specifically, as the intramolecular degrees of freedom become more coupled, the ability to grow molecules “on the fly”, using configurational bias, diminishes. We note that the arbitrary trial distribution version of configurational bias⁴⁴ partially addresses this problem.

To demonstrate the utility of the approach presented here, we present liquid–vapor phase equilibria calculations for molecular models for *n*-octane, cyclohexane, 1-propanol, water, squalane, and pyrene. *N*-octane and cyclohexane represent typical nonpolar molecules with different topologies, 1-propanol and water are polar molecules of industrial importance, squalane is a relatively large branched flexible hydrocarbon, and pyrene is a relatively rigid aromatic molecule. Collectively, these molecules present a range of simulation challenges.

The paper is organized as follows. In the next two sections, we describe the simulation methods used to sample systems of interest and the techniques that we use to compute phase coexistence properties. Within section 4, we discuss the molecular models employed and provide details for the simulations performed within this study. In section 5, the results obtained are presented and compared with those in the literature. The paper is concluded in section 6.

2. SIMULATION METHODS

We begin by outlining the molecular simulation tools that are used within this study. The features of many of these techniques are detailed within previous reports. In what follows, we highlight those features that are most relevant to the current study. In the next major section, we describe how these techniques are used within a series of general simulation strategies for computing liquid–vapor phase coexistence properties.

Reservoir grand canonical MC. Within a grand canonical (GC) simulation, molecules are exchanged between a system of interest at a given temperature and activity and a reservoir of molecules at the same temperature.⁴⁵ The configuration space sampled by the system and reservoir are governed by the system and reservoir Hamiltonians, respectively. In practice, GC simulations are often conducted with an implicit reservoir, wherein upon insertion of a molecule a representative configuration is generated “on the fly” in a manner consistent with the reservoir Hamiltonian.⁴⁵ This approach becomes more difficult to implement as the topology of the molecule becomes more complex (e.g., branched, ring topologies). In this work, we employ a scheme in which an explicit reservoir is maintained during a simulation.^{7,46} Within this approach, a reservoir consisting of a finite number of molecules (e.g., 500) is created at the outset of a GC simulation. Upon insertion of a molecule, one of the reservoir configurations is randomly selected and placed within the simulation cell. During the GC simulation, the configuration space of the reservoir is periodically refreshed via a collection of MC moves. This approach provides a relatively efficient means to perform GC simulations with molecules possessing complex topologies. The drawbacks of the technique include the need to invest resources to create and maintain the explicit reservoir and a restriction on the use of configurational-bias strategies^{41–43} to insert molecules. Finally, although not pursued here, we note that fragment-based reservoir techniques⁴⁷ have recently been introduced that allow more flexibility regarding the use of configurational bias.

Two reference states are commonly employed to define the reservoir: the ideal chain and ideal gas states.⁴⁵ Within the ideal chain state the reservoir is governed by a Hamiltonian that includes bonded interactions only (bond stretching, angle bending, torsion). Within the ideal gas state, the reservoir Hamiltonian also includes intramolecular nonbonded interactions (Lennard-Jones, Coulombic). For reasons that we discuss below, we work with an ideal gas reservoir in this study.

Molecules in the reservoir are equilibrated at the outset of the simulation and then periodically updated. In this work, we use a mix of hybrid MC,⁴⁸ pivot,^{49,50} torsion angle rotation, and crank-shaft⁵¹ MC trial moves to sample the reservoir configuration space. The pivot move involves rotation of a sequence of atoms about a fixed point. A torsion angle rotation swings a molecular segment about an axis associated with a torsion angle. The crank-shaft move involves rotation of one or more connected atoms around an axis defined by the extreme

atoms of the selected segment. The Monte Carlo move mix employed depends upon the topology of the molecule. Generally speaking, hybrid MC moves are effective at sampling the local configuration space associated with a free energy basin, but are relatively ineffective at enabling a molecule to traverse a potential energy barrier (e.g., one of the barriers associated with rotation of an alkane chain around a dihedral angle). The pivot, torsion angle rotation, and crank shaft moves are more aggressive in nature and facilitate the sampling of multiple minima within the free energy surface. The probability p_{res} of selecting a reservoir molecule with a configuration r_{res} is given by the Boltzmann distribution,

$$p_{\text{res}}(r_{\text{res}}) = \frac{e^{-\beta E_{\text{res}}(r_{\text{res}})}}{z_{\text{res}}} \quad (1)$$

where E_{res} is the intramolecular interaction energy and z_{res} is the molecular configurational partition function.

Growth Expanded Ensemble. This simulation strategy²⁴ provides a means to insert and remove molecules from a system in a stagewise manner, thereby facilitating the transfer of large molecules into and out of the system. Numerous growth schemes have been proposed.^{24,52–56} Some of these employ a segment-based transfer approach (e.g., atom-by-atom insertion of a molecule), while others utilize a modified Hamiltonian to define the transfer process. We focus on the latter approach here. We transfer molecules to and from the system via an M -stage process wherein the Lennard-Jones and Coulombic interactions between the selected (tagged) molecule and the rest of the system scale with the stage number m . The details of this interaction potential are given in the Model Systems and Simulation Details section.

Within the growth expanded ensemble, a given configuration consists of N complete molecules and at most one partial molecule. The probability π_j of observing a given microstate j at temperature T and activity ξ can be written as

$$\pi_j = \frac{1}{\Xi} \frac{1}{N!} \left(\frac{V\xi}{z_{\text{res}}} \right)^{N+\Delta} e^{-\beta E_j} \quad (2)$$

where Ξ is the partition function of the growth expanded ensemble, V is the system volume, $\beta = 1/k_{\text{B}}T$ is the inverse temperature (where k_{B} is the Boltzmann's constant), E_j is the total configurational energy, and Δ is a parameter that evaluates to 0 when the system consists of full molecules only and 1 when the system contains a tagged molecule. Note that alternative definitions of Δ are also possible. For example, one could employ the definition $\Delta = m/M$. The activity $\xi = qz^{\text{res}}e^{\beta\mu}$, where q represents the component of the molecular partition function stemming from integration over the momenta, provides an alternative representation of the ideal gas reservoir chemical potential μ . As a point of reference, the quantities ξ and z^{res} used in this work are equivalent to $\beta P_{\text{id, gas}}$ and $e^{-\beta\mu_{\text{intra}}^0}$, respectively, described within Appendix G in the work of Frenkel and Smit.⁴⁵

During an MC simulation one attempts to insert (remove) tagged molecules, increase (decrease) the stage number m while the molecule remains partially grown, and complete (start) the growth (destruction) of a molecule. Let us say that one of these moves takes the system from microstate j to microstate k . The acceptance probability of the latter two trial moves is simply $\min\{1, \pi_k/\pi_j\}$. The acceptance probability for inserting a tagged molecule is,

$$p_{\text{acc}} = \min\left\{1, \frac{\pi_k}{\pi_j} \frac{1}{p_{\text{res}}}\right\} = \min\{1, V\xi e^{-\beta(E_k - E_j - E_{\text{res}})}\} \quad (3)$$

The corresponding transition probability for removing a tagged molecule is

$$p_{\text{acc}} = \min\left\{1, \frac{\pi_k}{\pi_j} p_{\text{res}}\right\} = \min\{1, (V\xi)^{-1} e^{-\beta(E_k - E_j + E_{\text{res}})}\} \quad (4)$$

We employ an approach in which attempted transitions are restricted to those between adjacent subensembles. This restriction can be eased, if desired. Finally, we schedule the number of growth stages M to increase with the total molecule number N .

As has been noted by other authors,^{54,57} the probabilities $p_{\text{stg}}(m)$ for observing a system with fixed N in a given growth stage m can vary significantly. As a result, the system tends to get “stuck” at intermediate values of m , leading to infrequent transitions from one value of N to another. We overcome this issue by using flat histogram techniques⁵⁸ to promote uniform sampling of all growth stages. For example, in certain scenarios we are interested in obtaining the density probability distribution of a system over a wide range of densities. In such cases, we weight the simulation such that all $[N, m]$ combinations are uniformly sampled, thereby promoting regular transitions between one total molecule number N and another.

Isothermal–Isobaric Ensemble. In what follows below, we also describe schemes in which the isothermal–isobaric ensemble (NPT)⁴⁵ is used to explore the density and/or temperature dependence of saturation properties. The advantage of this approach stems from the absence of molecule transfers.

Hybrid MC. Within this approach,⁴⁸ short molecular dynamics trajectories are used to propagate a system through configuration space. These moves facilitate rearrangements of the system at multiple scales. Locally, the moves enable molecules to sample their intramolecular degrees of freedom, including bond distances, bending angles, and dihedral angles. HMC moves are also known to be useful in facilitating cooperative many-body rearrangements within a system. The time step used to integrate the equations of motion is adjusted to achieve a target acceptance ratio. Commonly, one finds that this time step is larger than those employed in conventional molecular dynamics (MD) simulations. The number of time steps completed within a single MC move also impacts performance. The study of Mehlig et al.⁵⁹ suggests that 5–10 time steps per MC move is optimal. Here, we use 5 time steps per MC move. Another advantage stems from the fact that one does not need to use the full Hamiltonian to compute the forces required for the equations of motion.⁴⁵ Instead, a “cheap” Hamiltonian can be used to propagate the equations of motion, while the full Hamiltonian is used to compute the MC acceptance probability.

Within this work, we employ three hybrid MC moves: (1) a conventional HMC move, in which all molecules within the system are displaced; (2) an HMC-One move, in which a single molecule is selected and an MD trajectory is completed with that molecule in a field created by the other molecules; and (3) a torsion-free HMC-One move, in which the forces related to the dihedral angles are ignored and the coordinates of the

molecule are propagated in the absence of these forces. The conventional HMC move helps to refresh the intramolecular degrees of freedom and facilitates cooperative rearrangements within the system. The HMC-One move enables us to focus attention on a specific molecule. This preferential targeting is helpful within the growth expanded ensemble scheme, wherein a partially grown molecule is often present. During a grand canonical simulation, moves are regularly attempted to change the growth stage of this molecule. It is therefore desirable for the configuration space within the local environment of this molecule to be updated frequently. To assist in this endeavor, when performing an HMC-One move, we tune the molecule selection such that the partially grown molecule is selected with a much higher probability (e.g., 35%) than the other molecules. Although we did not pursue this opportunity within our work, we suspect that preferentially targeting neighbors of the partially grown molecule would also prove useful.

We now turn our attention to the torsion-free HMC-One move. The motivation for this move stems from the low frequency with which a system traverses free-energy barriers with traditional HMC. As an illustrative example, consider the dihedral angles along a united atom model for an n -alkane. The associated torsion potential contains multiple minima, separated by non-negligible barriers. As noted above, HMC moves sample a local free energy basin rather well, but generally do not facilitate barrier hopping. Within a torsion-free HMC-One move, forces related to the torsion potential are ignored, thereby eliminating the associated free energy barriers and allowing the system to uniformly sample the dihedral angles. In practice, we tune the MD time step based on the standard HMC-One move and subsequently use this value for the torsion-free move. While the resulting acceptance probabilities for the torsion-free moves are lower, the moves facilitate the sampling of important degrees of freedom. Finally, we note that the torsion-free moves are similar in spirit to MD techniques in which select degrees of freedom are coupled to a high-temperature heat bath.⁶⁰ More generally, the torsion-free approach can be viewed as one example of a broader strategy in which carefully selected artificial forces are used to modify select degrees of freedom during an MD trajectory.

grand canonical Temperature-Expanded Ensemble (GC-TEE). This algorithm²⁷ provides a means to explore the temperature and activity dependence of the thermophysical properties of a specific phase. An expanded ensemble⁶¹ is created in which subensembles are differentiated by temperature (or inverse temperature). One also has the freedom to assign unique activity values to each subensemble,²⁷ provided that adjacent subensembles share a common region of configuration space. For example, in what follows below we describe a scheme in which the temperature and activity are scheduled to trace the liquid–vapor saturation curve. During a simulation, attempts are made to transition from one subensemble to another. Let us say that an attempted move takes the system from subensemble j to subensemble k . The acceptance probability for this trial move is

$$p_{\text{acc}} = \min \left\{ 1, \frac{\pi_k}{\pi_j} \right\} \\ = \min \left\{ 1, \left[\left(\frac{\xi_k}{\xi_j} \right) \left(\frac{z_j^{\text{res}}}{z_k^{\text{res}}} \right) \right]^{N+\Delta} e^{-E(\beta_k - \beta_j)} \right\} \quad (5)$$

Note that evaluation of the transition probability requires knowledge of the ratio of the reservoir configurational partition function at temperatures β_j and β_k . This relative free energy curve is obtained by performing a temperature expanded ensemble simulation^{27,61–63} with a collection of isolated molecules that interact via the reservoir Hamiltonian.

Within a GC-TEE simulation, we often use a growth expanded ensemble²⁴ to insert and remove molecules. As noted above, the probability distribution $p_{\text{stg}}(m)$ for observing a system with fixed N and T in a given growth stage m can span one or more orders of magnitude. In this work, GC-TEE simulations are used to sample a particular phase of interest (e.g., the liquid phase). We find that the $p_{\text{stg}}(m)$ distribution does not vary significantly over the relatively narrow range of particle numbers sampled within a given subensemble. In contrast, $p_{\text{stg}}(m)$ typically does vary substantially over the relatively wide range of temperatures sampled within a GC-TEE simulation. To account for this, we introduce stage- and temperature-specific weights that promote uniform sampling⁵⁸ of temperature space and uniform sampling of the intermediate stages related to the molecule growth process. This weighting procedure does not influence the particle number probability distribution collected at a given temperature.

Isothermal–Isobaric Temperature Expanded Ensemble (NPT-TEE). This algorithm is the isothermal–isobaric analogue to the GC-TEE technique discussed above. An expanded ensemble⁶¹ is created in which subensembles are differentiated by temperature (or inverse temperature) and pressure. During a simulation, attempts are made to transition from one subensemble to another. The acceptance probability for a trial move that takes the system from subensemble j to subensemble k is

$$p_{\text{acc}} = \min \{ 1, e^{-E(\beta_k - \beta_j) - V(\beta_k p_k - \beta_j p_j)} \} \quad (6)$$

In contrast to the GC-TEE approach, one does not require information regarding the intramolecular configurational partition function to perform NPT-TEE simulations.

3. PHASE COEXISTENCE

In what follows below, we outline strategies for identifying phase coexistence points.

Direct grand canonical Simulation. Within this approach,^{2,3} we first collect the particle number probability distribution $\Pi_{\text{GC}}(N) = \Pi_{\text{GC}}(N, m = M)$ for a system at a specified inverse temperature β , activity ξ_0 , and volume V over a range of densities that includes both the vapor and liquid phases of interest. This task is accomplished using the reservoir GC⁷ and growth EE²⁴ techniques described above. We then locate the phase coexistence point by identifying the activity at which the grand potentials of the liquid and vapor phases are equal. The grand potential of phase k is given by³

$$\beta\Omega_k(\beta, \xi) = -\beta p_k(\beta, \xi)V = -\ln \left[\sum_{\text{phase } k} \frac{\Pi_{\text{GC}}(N; \xi)}{\Pi_{\text{GC}}(0; \xi)} \right] \quad (7)$$

where the summation extends over all N characteristic of phase k . Histogram reweighting¹ is used to obtain the particle number probability distribution at an activity of interest:

$$\ln \Pi(N; \xi) = \ln \Pi(N; \xi_0) + N \ln(\xi/\xi_0) \quad (8)$$

Saturated densities are taken as the first moments of the vapor and liquid peaks of the density probability distribution shifted to the saturated activity.

Direct Isothermal–Isobaric Simulation. Within this approach,³ we first establish a range of volume to sample that includes states characteristic of the vapor and liquid phases at a given temperature of interest. This volume range is then discretized such that states are separated by a constant shift in $\ln V$. We then determine the volume probability distribution $\Pi_{\text{NPT}}(\ln V)$ for a system at a specified inverse temperature β , pressure p_0 , and molecule number N . The phase coexistence point is identified by determining the pressure at which the Gibbs free energies of the liquid and vapor phases are equal:

$$\sum_{\text{vapor}} \Pi_{\text{NPT}}(\ln V; p) = \sum_{\text{liquid}} \Pi_{\text{NPT}}(\ln V; p) \quad (9)$$

where the summations extend over all $\ln V$ characteristic of a given phase. Histogram reweighting¹ is used to obtain the volume probability distribution at a pressure of interest:

$$\ln \Pi(\ln V; p) = \ln \Pi(\ln V; p_0) - \beta(p - p_0)V \quad (10)$$

Although this NPT-based approach has many intuitively appealing features (e.g., no particle insertions/removals), the method often shows poor performance relative to the GC-based analogue.

grand canonical Temperature-Expanded Ensemble (GC-TEE). This approach²⁷ is used to trace a liquid–vapor saturation curve. It is similar in spirit to the Gibbs–Duhem integration method introduced by Kofke.^{29,30} However, with the current method, one can identify coexistence points over a wide range of temperatures in a single simulation, i.e., one does not move sequentially along a saturation line. In addition, one does not need to invoke a numerical integration procedure to complete the calculation. To initiate the calculation, one must have knowledge of at least one coexistence point, which we label $[\beta_1, \xi_1]$, and a means to estimate the saturated activities at other temperatures of interest. In this work, we use the direct GC approach to locate two coexistence points at relatively high temperature. We then estimate saturated activities over a wide range of temperature by assuming a linear relationship between $\ln \xi$ and β along the saturation line. Next, we perform GC-TEE simulations within the vapor and liquid phases in which subensembles are separated by a constant shift in β and the activity within a subensemble is provided by the aforementioned linear relationship. The subensemble probability distribution $\Pi_{\text{GCTEE}}([\beta, \xi])$ provides the relative grand potential of phase k along the path sampled:

$$\beta \Omega_k(\beta, \xi) = \beta_1 \Omega_k(\beta_1, \xi_1) - \ln \left(\frac{\Pi_{\text{GCTEE}}([\beta, \xi])}{\Pi_{\text{GCTEE}}([\beta_1, \xi_1])} \right) \quad (11)$$

The grand potential of the vapor phase will generally differ from that of the liquid phase for the $[\beta, \xi]$ estimates provided by the linear relationship.

The phase coexistence point is now located by determining the activity ξ_c for which $\Omega_{\text{vap}} = \Omega_{\text{liq}}$. The visited-states probability distributions $P_k(N; [\beta, \xi])$ collected within each subensemble are used to deduce how Ω_k varies with ξ at constant β . The specific procedure employed to compute Ω_k depends upon the difference between ξ_c and the subensemble activity ξ_e . If these values are relatively close, then the particle number probability distribution at ξ_c overlaps significantly with

the one originally collected at ξ_e . In such cases, conventional histogram reweighting¹ is used to evaluate the activity dependence of Ω_k :

$$\beta \Omega_k(\beta, \xi_c) = \beta \Omega_k(\beta, \xi_e) - \ln \left(\frac{\sum_N P_k(N; [\beta, \xi_e]) \left(\frac{\xi_c}{\xi_e} \right)^N}{\sum_N P_k(N; [\beta, \xi_e])} \right) \quad (12)$$

In contrast, if ξ_c is far removed from ξ_e , then there is little overlap between the original and shifted particle number probability distributions and straightforward histogram reweighting proves insufficient. To account for these two scenarios, we first define a maximum activity shift $\Delta \ln \xi = \ln \xi_m / \xi_e$, where ξ_m is the minimum/maximum activity for which we can comfortably reweight $P_k(N; [\beta, \xi_e])$. Here, we establish $\Delta \ln \xi$ by finding the point at which the difference in expectation values for the original and shifted probability distributions is approximately one-half of a standard deviation of $P_k(N; [\beta, \xi_e])$. We typically perform this analysis at a single moderate temperature and subsequently utilize the resulting $\Delta \ln \xi$ to complete the coexistence calculation at all temperatures of interest. When $\ln \xi_c / \xi_e < \Delta \ln \xi$, eq 12 is used to deduce $\Omega_k(\beta, \xi_c)$. When $\ln \xi_c / \xi_e > \Delta \ln \xi$, we first determine $\Omega_k(\beta, \xi_m)$ using eq 12 and subsequently extrapolate the value of $\Omega_k(\beta, \xi_c)$ by assuming a linear relationship between Ω_k and $\ln \xi$.

This analysis provides a revised estimate for the relationship between β and ξ along the saturation curve. The process described above is now repeated with the new $[\beta, \xi]$ values. This iterative procedure is repeated until all saturated activities converge. For the systems discussed below, 2–4 iterations are required to reach this point. We typically perform relatively short runs during the iteration phase and then complete longer runs with the converged activities.

Isothermal–Isobaric Temperature-Expanded Ensemble. This approach is the NPT analog of the GC-based approach described above. The procedure is initiated by using a direct approach to locate two coexistence points at relatively high temperature, of which we label one $[\beta_1, p_1]$, and subsequently estimate the saturated activities at other temperatures of interest by assuming a linear relationship between $\ln p$ and β along the saturation line. Note that the initial high-temperature coexistence points could be generated by a number of techniques, including via direct GC, direct NPT, or Gibbs ensemble simulations. NPT-TEE simulations within the vapor and liquid phases are now conducted in which subensembles are separated by a constant shift in β and the pressure within a subensemble is provided by the aforementioned linear relationship. The subensemble probability distribution $\Pi_{\text{NPTTEE}}([\beta, p])$ provides the relative Gibbs free energy of phase k along the path sampled:

$$\beta G_k(\beta, p) = \beta_1 G_k(\beta_1, p_1) - \ln \left(\frac{\Pi_{\text{NPTTEE}}([\beta, p])}{\Pi_{\text{NPTTEE}}([\beta_1, p_1])} \right) \quad (13)$$

The Gibbs free energy of the vapor phase will generally differ from that of the liquid phase for the $[\beta, p]$ estimates provided by the linear relationship. To facilitate location of the coexistence point, we use visited-states probability distributions $P_k(V; [\beta, p])$ collected within each subensemble to deduce how G_k varies with p at constant β . The procedure employed to

Table 1. Harmonic Bond Stretching Bending Parameters

molecule	atom 1	atom 2	$k_r(\text{K } \text{\AA}^{-2})$	$r_{\text{eq}}(\text{\AA})$
octane, cyclohexane, squalane	C	C	270 000 (ref 71)	1.54 (ref 64)
pyrene	C(aromatic)	C(aromatic)	472 000 (ref 72)	1.3926 (ref 67)
pyrene	C(aromatic)	H	369 384 (ref 72)	1.08 (ref 67)
propanol	C	O	322 000 (ref 71)	1.43 (ref 65)
propanol	O	H	556 000 (ref 71)	0.945 (ref 65)
SPC water	O	H	452 923 (ref 73)	1.000 (ref 66)

locate the coexistence point is analogous to that described for the GC-TEE approach. If the saturated p_c and subensemble p_c pressures are relatively close, then conventional histogram reweighting¹ is used to evaluate G_k at the pressure p_c :

$$\beta G_k(\beta, p_c) = \beta G_k(\beta, p_c) - \ln \left(\frac{\sum_V P_k(V; [\beta, p_c]) e^{-\beta(p_c - p_c)V}}{\sum_V P_k(V; [\beta, p_c])} \right) \quad (14)$$

In contrast, if the saturated pressure is far removed from the subensemble pressure, then there is little overlap between the original and shifted volume probability distributions and straightforward histogram reweighting proves insufficient. For these cases, we use an extrapolation procedure analogous to that described above for the GC-TEE case to obtain $G_k(\beta, p_c)$.

This analysis provides a revised estimate for the relationship between β and p along the saturation curve. The process described above is now repeated with the new $[\beta, p]$ values. For the systems discussed below, two to three iterations are required to converge to a solution. We typically perform relatively short runs during the iteration phase and then complete longer runs with the converged pressures. Finally, although we did not pursue this approach here, we note that the overall computational effort may be further reduced by performing the liquid-phase NPT-TEE simulation just once during the iteration process, since it is observed that the Gibbs free energy of the liquid phase changes by a statistically insignificant amount with successive iterations. In a similar manner, within the GC-TEE approach, computational effort may be reduced by performing the vapor-phase simulation just once during the iteration process.

4. MODEL SYSTEMS AND SIMULATION DETAILS

Molecular Models. The simulation methodology described above is used to determine vapor–liquid equilibria for molecular models of *n*-octane,⁶⁴ cyclohexane,⁴² propanol,⁶⁵ water,⁶⁶ pyrene,⁶⁷ and squalane.⁶⁸ The compounds selected encompass a range of molecular interactions and topologies. For *n*-octane, cyclohexane, propanol, and squalane, the united atom TraPPE model^{64,65,67,69,70} is used. Here, nonbonded interactions between pseudo-atoms are described by Lennard-Jones 12–6 and Coulombic point-charge potentials. The parameters for interactions between unlike pseudo-atoms are calculated using Lorentz–Berthelot combining rules. The TraPPE model employs a harmonic potential to describe angle bending and a triple cosine potential to govern the sampling of dihedral angles. Within the TraPPE model bond distances are taken to be fixed. In this work, we desired to work with flexible molecules, and therefore employed a harmonic potential to describe bond stretching. The potential takes the form $u_{\text{stretch}}(r) = \frac{1}{2}k_r(r - r_{\text{eq}})^2$ and the parameters^{71–73} are provided in Table 1. The decision to work with fully flexible

models was largely driven by our desire to utilize HMC to sample the intramolecular degrees of freedom of a molecule. The HMC algorithm requires the use of a time-reversible and area-preserving integrator.⁴⁵ In this work, we employ the multiple time step algorithm of Tuckerman et al. to propagate the system.⁷⁴ Techniques commonly employed to satisfy geometric constraints (e.g., SHAKE), such as fixed bond lengths, generally do not satisfy the time-reversible and area-preserving criteria. As a result, HMC has traditionally been of limited use when working with models containing constraints. We note that the recent work of Toxvaerd et al.⁷⁵ may provide a solution to this problem, as these authors have introduced a time-reversible algorithm designed to handle bond constraints.

An extension of the TraPPE explicit hydrogen model is used to describe pyrene. Rai and Siepmann introduced parameters for several single-ring aromatic molecules, including benzene.⁶⁷ In this work, we use the Lennard-Jones parameters provided by Rai and Siepmann for all C and H atoms within pyrene. C atoms with a H atom attached carry a charge of $-0.095e$, while fused C atoms do not carry a charge. It follows that all H atoms carry a charge of $+0.095e$. The primary objective of this work is to evaluate the performance of the methods described above, and this approach for modeling pyrene provides a means to accomplish this task. A more-realistic description of pyrene would likely result in the charge distribution being spread across the ring structure.

The rigid SPC/E model⁶⁶ and a flexible version of the SPC model^{54,76} are used to describe water. Within these models, partial charges are placed on O and H atoms, but only the O site is associated with a Lennard-Jones 12–6 potential. Bond lengths and bending angles within the flexible model are governed by harmonic potentials. Parameters are provided in Table 1. The SPC/E model is the only rigid model that we work with. In this case, we simply keep a single copy of the molecule within the reservoir and do not perform MC moves on this structure.

GC Simulations. We employ a cubic simulation cell of $1 \times 10^5 \text{ \AA}^3$ for *n*-octane, cyclohexane, pyrene, and propanol, $1.5 \times 10^5 \text{ \AA}^3$ for squalane, $3.5 \times 10^4 \text{ \AA}^3$ for SPC/E water, and $1 \times 10^4 \text{ \AA}^3$ for SPC water. Additional calculations are performed with other cell sizes for select molecules to examine finite-size effects. The MC move mix includes molecular displacements and rotations, HMC, HMC-One, and torsion-free HMC-One moves, as well as molecular insertions (or growth) and deletions (or reductions). Insertions/deletions are selected with 60% probability and the displacement, rotation, HMC, and HMC-One moves are partitioned such that the computational time invested in each move is roughly equal. This “equal computational cost” approach is adopted primarily for the sake of simplicity. A detailed analysis regarding the efficacy of these moves within the overall scheme pursued here would be valuable. However, this analysis falls outside the scope of this study. Finally, the split between torsion-free and regular HMC-

One move is variable and is based on the topology of the molecule.

The reservoir includes $N_{\text{res}} = 500$ molecules. At the outset of a simulation, we use a single molecular configuration to populate the reservoir and subsequently equilibrate the molecular configurations by completing 2000–5000 MC cycles on the reservoir, with a cycle defined as N_{res} MC steps. During a GC simulation, we generally complete two reservoir MC moves for every one MC move performed with the main simulation cell. The computational cost associated with maintaining the reservoir typically does not exceed 10% of the total computational effort. Simulations employing the ideal-gas reservoir are completed for all molecules studied. Additional simulations are conducted with *n*-octane and squalane employing the ideal-chain reservoir.

Two schemes are employed for describing the manner in which a partial molecule interacts with the rest of the system. The first uses a modified Lennard-Jones potential introduced by Lo and Palmer:⁵³

$$u(r) = \lambda 4\epsilon \left\{ \left[\frac{\sigma}{r + \zeta(1 - \lambda)} \right]^{12} - \left[\frac{\sigma}{r + \zeta(1 - \lambda)} \right]^6 \right\} \quad (15)$$

where r is the site–site separation distance, ϵ and σ are the conventional Lennard-Jones energy and size parameters, respectively, ζ is an adjustable parameter that we set to 0.5σ , and $\lambda = m/M$ is a scaling parameter that controls the growth of a molecule. Coulombic interactions are evaluated with the following potential:

$$u(r) = \frac{1}{4\pi\epsilon_0} \frac{q_i q_j}{[r + \zeta(1 - \lambda)]} \quad (16)$$

where q_i and q_j represent the charges at sites i and j . With this form, the quantity $r + \zeta(1 - \lambda)$ provides an effective separation distance between two sites, which is artificially large in the initial growth process and approaches the true separation distance as the growth process reaches an end. For cases in which a Lennard-Jones length scale is not associated with a site–site interaction, we take $\zeta = 0.5\sigma_{\text{bare}}$, with $\sigma_{\text{bare}} = 3.0$ Å. For example, effective oxygen–hydrogen separation distances within SPC/E water adopt this form.

The second scheme we employ uses the modified Lennard-Jones potential introduced by Shi and Maginn,⁵⁴

$$u(r) = \lambda 4\epsilon \left\{ \frac{1}{\left[\left(\frac{\zeta}{\sigma} \right) (1 - \lambda)^2 + \left(\frac{r}{\sigma} \right)^6 \right]^2} - \frac{1}{\left[\left(\frac{\zeta}{\sigma} \right) (1 - \lambda)^2 + \left(\frac{r}{\sigma} \right)^6 \right]} \right\} \quad (17)$$

and the following Coulomb potential:

$$u(r) = \begin{cases} 2(\lambda - 0.5) \frac{q_i q_j}{4\pi\epsilon_0 r} & \lambda > 0.5 \\ 0 & \lambda < 0.5 \end{cases} \quad (18)$$

Here, introduction of the electrostatic interaction is delayed to prevent overlap of oppositely charged pseudo-atoms at low λ , because of weak repulsive interactions (e.g., bare H atoms in water or propanol). We find that both of the schemes outlined here provide satisfactory results. We did not find that either of the potential forms provided distinct advantages in the limited comparative cases that we examined. Performing a systematic investigation of the merits of these approaches (along with other possibilities) represents a significant undertaking. We reserve this activity for a subsequent study.

A spherical cutoff coupled with a mean-field correction is used to account for long-range Lennard-Jones interactions.⁴⁵ The Ewald summation is used to compute electrostatic interactions.⁴⁵ Lennard-Jones interactions and the real space contribution to the Ewald summation are cutoff at a distance of 14 Å for all molecules except water, for which a 10 Å cutoff distance is used. The algorithm described by Frenkel and Smit⁴⁵ with a precision of 5×10^{-5} is used to determine the damping factor and number of lattice vectors associated with the Ewald summation. To reduce computational expense associated with the growth process, we include the complete reciprocal and self-interaction contributions to the Ewald summation in the first growth stage and subsequently update the real space contribution only with an increase in the stage number.

GC-TEE simulations are performed over a broad range of β with a discretization of $\Delta\beta = 2.0 \times 10^{17} \text{ J}^{-1}$. In addition to the MC moves noted above, we now perform subensemble change moves with a probability of 1%–2%. The reservoir temperature is updated to match the system temperature upon completion of a successful expanded ensemble move. We find that the reservoir remains equilibrated with this scheme. The rationale for this approach stems from the low frequency with which subensemble change moves are attempted and the relatively small $\Delta\beta$ values employed. As noted above, knowledge of the temperature dependence of z^{res} is required to complete GC-TEE simulations. This curve is generated via canonical temperature-expanded ensemble simulations consisting of 100 isolated molecules. While these simulations could be conducted with just a single molecule, we find that utilizing the larger ensemble improves the precision of $z^{\text{res}}(\beta)$ data.

NPT Simulations. We employ a cubic simulation cell consisting of 250 molecules. The MC move mix includes molecular displacements and rotations, HMC, HMC-One, and torsion-free HMC-One moves, as well as volume change attempts. Volume changes are selected with a probability of approximately $1/N$ and the remaining moves are distributed using the “equal computational cost” approach noted above. The spherical cutoff distance for the Lennard-Jones and real-space Ewald contribution is taken to be a constant fraction of the simulation cell length, usually 50%. NPT-TEE simulations are performed in a manner analogous to those described for the GC-TEE simulations.

Computing Probability Distributions. To implement the approach outlined above, one must evaluate probability distributions (e.g., density or temperature probability distributions) that span many orders of magnitude. Directly capturing a probability distribution within a standard simulation is generally not feasible, because of the limited number of macrostates (e.g., particle number, subensemble) that the system naturally samples. As a result, we use a series of advanced sampling schemes^{58,77–83} to compute probability distributions. We detail the approach that we use within the Appendix.

5. RESULTS AND DISCUSSION

Figures 1 and 2 contain saturated density and vapor pressure data, respectively, for *n*-octane and cyclohexane. For both fluids,

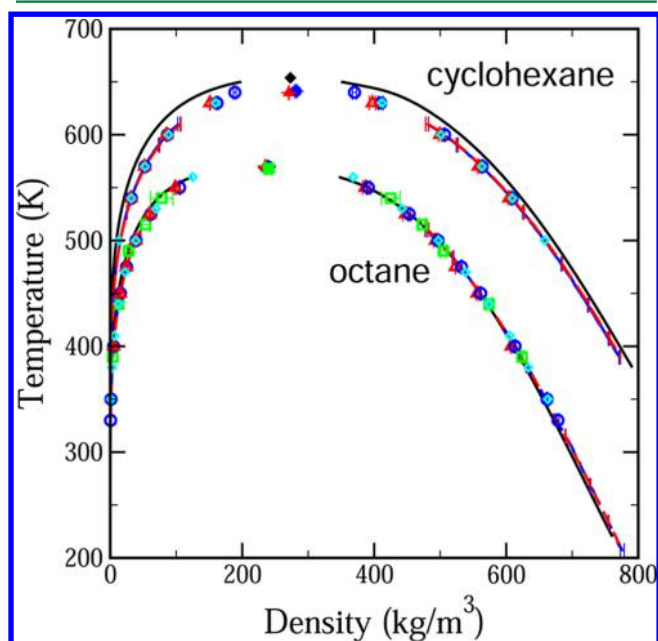


Figure 1. Temperature dependence of the saturated densities for *n*-octane and cyclohexane. The data for cyclohexane are shifted up by 100 K for the sake of clarity. Data are provided from experiment (solid black lines),⁸⁵ direct GCMC (open blue circles), direct NPT (open red triangles), GC-TEE (dashed blue lines), and NPT-TEE (dashed red lines) simulations. Cyan diamonds represent data generated using the Towhee code.^{10,84} Open green squares correspond to the Gibbs ensemble data of Martin and Siepmann.⁶⁴ The continuous dashed lines are drawn by connecting the points associated with each subensemble sampled within TEE simulations (~ 1030 points for *n*-octane and ~ 980 points for cyclohexane). Uncertainties for the TEE approaches are provided at select temperatures.

we provide results from direct GC and NPT simulations at several temperatures, along with results from GC-TEE and NPT-TEE simulations that extend to relatively low temperature. The figures also contain Gibbs ensemble results from the Siepmann group,⁶⁴ transition matrix MC^{3,4} results from our group¹⁰ that were generated with the Towhee code,⁸⁴ and experimental data.⁸⁵ In the Towhee case, a GC approach, coupled with configurational-bias for facilitating insertions and deletions (in lieu of the growth expanded ensemble approach) and transition matrix MC for generating probability distributions, was used to generate coexistence data. The Towhee data were generated with the original TraPPE model, i.e., bond lengths were fixed. Results from the various techniques are in good agreement.

The noted agreement between the results from this study and those from the Siepmann group and the Towhee project suggests that the flexibility of C–C bonds has a negligible impact on the thermophysical properties of an alkane. This observation is consistent with the findings reported previously.^{86,87} Interestingly, Patel et al. recently computed the phase diagram of propane using rigid- and flexible-bond TraPPE united atom models, and found a significant difference between the two phase diagrams.⁸⁸ The modulus of the harmonic bond stretching potential employed in their study (k_r

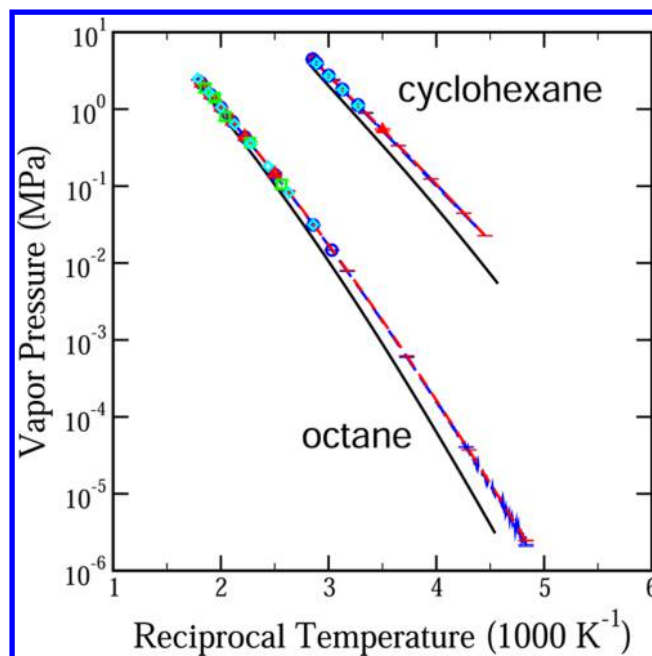


Figure 2. Temperature dependence of the vapor pressures for *n*-octane and cyclohexane. The data for cyclohexane are shifted to the right by 1000 K^{-1} for clarity. Symbols are defined in the same manner as in Figure 1.

$= 1.93 \times 10^5 \text{ K } \text{\AA}^{-2}$) is a factor of 1.4 smaller than the modulus used here. It seems unlikely that this difference could fully explain the relatively large shift in saturation properties. It appears that a more-detailed study is required to clarify this issue.

Figures 3 and 4 provide data related to the convergence of the *n*-octane vapor pressure within the GC-TEE and NPT-TEE approaches. Within each figure, we place the statistical

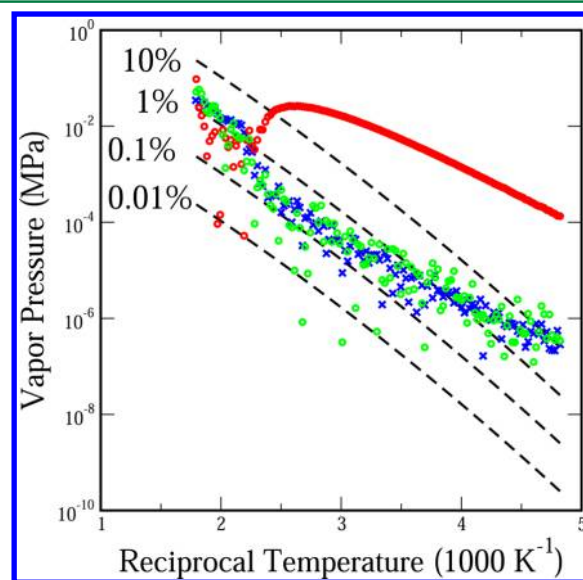


Figure 3. Convergence of the *n*-octane vapor pressures with successive iterations within the GC-TEE approach. Dashed black lines indicate percentages of final vapor pressure estimate ($p_{\text{final}}^{\text{sat}}$). Blue crosses correspond to statistical uncertainty in $p_{\text{final}}^{\text{sat}}$. Open red and green circles represent the absolute difference between $p_{\text{final}}^{\text{sat}}$ and vapor pressures obtained after the first and second iterations, respectively.

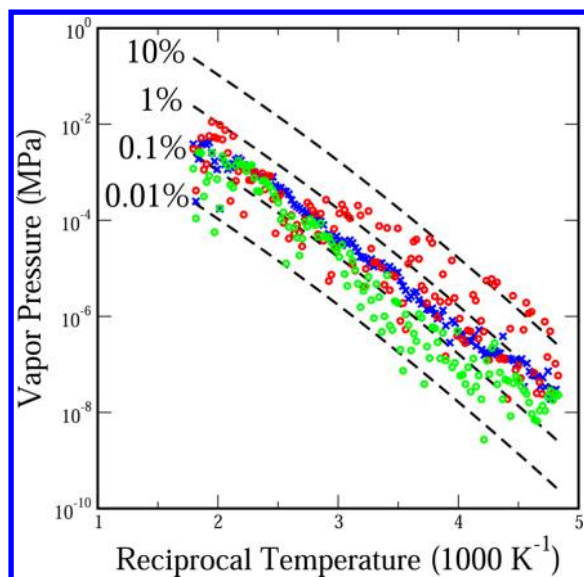


Figure 4. Convergence of *n*-octane vapor pressures with successive iterations within the NPT-TEE approach. Dashed black lines indicate percentages of final vapor pressure estimate ($p_{\text{final}}^{\text{sat}}$). Blue crosses correspond to statistical uncertainty in $p_{\text{final}}^{\text{sat}}$. Open red and green circles represent the absolute difference between $p_{\text{final}}^{\text{sat}}$ and vapor pressures obtained after the first and second iterations, respectively.

uncertainty in the vapor pressure measurement $\sigma_{p_{\text{sat}}}$ at various temperatures along the saturation curve. We also provide guides to indicate the percentage uncertainty in this quantity. Statistical uncertainties are determined by performing four independent sets of simulations. The standard deviation of the results from the four simulation sets is taken as an estimate of the statistical uncertainty. The statistical uncertainty in the vapor pressure provides a natural convergence limit; therefore, we discuss this quantity first. Uncertainties for the vapor pressure computed via the direct GC and NPT approaches are generally less than 0.4% and 1.0%, respectively. Confidence limits for the GC-TEE and NPT-TEE data at moderate to high temperature are similar. At relatively low temperature, the uncertainties associated with the GC-TEE results increase considerably. For example, at the lowest temperatures sampled with *n*-octane, the uncertainty in the vapor pressure is in the vicinity of 10%–20%. In contrast, uncertainties associated with the NPT-TEE approach remain below 1% at temperatures as low as 230 K and do not exceed 2% over the entire temperature range examined. This difference in performance is related to the difficulty of completing particle insertions and deletions within the GC ensemble at low temperatures within the high-density liquid phase. For the *n*-octane simulations, we vary the number of growth stages from four to eight over the temperature range studied here. It is likely that the precision of saturation data would improve with an increase in the number of stages. A more detailed analysis of this issue in a future study would prove useful.

We now return to the issue of convergence for the GC-TEE and NPT-TEE approaches. Within Figures 3 and 4, we place the difference $\Delta p_k^{\text{sat}} = |p_k^{\text{sat}} - p_{\text{final}}^{\text{sat}}|$ between the estimate of the vapor pressure p_k^{sat} obtained after completion of iteration stage k and the final vapor pressure estimate $p_{\text{final}}^{\text{sat}}$. For example, the values p_1^{sat} correspond to those values computed using data generated with the linear relationship between $\ln \xi$ and β (GC-TEE) or $\ln p$ and β (NPT-TEE). With the GC-TEE approach

(Figure 3), we find that the values for Δp_1^{sat} are commensurate with $\sigma_{p_{\text{sat}}}$ at relatively high temperatures, but are significantly larger than $\sigma_{p_{\text{sat}}}$ at moderate to low temperatures. This result is reasonable, given that the iterative process begins using known values for ξ^{sat} generated at relatively high temperature. However, after completion of the subsequent iterative step, we find that the values for Δp_2^{sat} track $\sigma_{p_{\text{sat}}}$ over the entire temperature range, indicating that the procedure converges within two iterations. For the NPT-TEE approach (Figure 4), we find that p_1^{sat} is generally within 10% of $p_{\text{final}}^{\text{sat}}$ over the entire temperature range. It is also clear that the procedure converges after the second iteration step.

One can appreciate the difference in the ability of the two approaches to capture p^{sat} following the first iterative step by considering the linearity of the $\ln \xi$ vs β and $\ln p$ vs β relationships. We find that the latter relation displays less curvature, and, therefore, the linear guess provides a better estimate of the variation of the two field variables along the saturation curve. We note that after we completed the GC-TEE simulations described here, we noticed that the combination $\ln \xi + \ln z^{\text{res}}$ is a more linear function of β than $\ln \xi$. Therefore, we suggest that one consider using this combination to forecast the activity–temperature relation in future studies. In terms of computational resources, we complete the liquid-phase GC-TEE simulations with 60 processors using the scheme described in the Appendix. A total of 10 and 65 wall-clock hours are used to complete the first and second (final) iterative steps, respectively. The analogous vapor-phase GC-TEE simulations require significantly fewer resources ($\sim 10\%$ of the liquid-phase requirements). The NPT-TEE approach requires similar computational resources. Finally, we note that the results obtained with *n*-octane are qualitatively fairly typical of what we observe for the TEE-based iterative procedures. From a quantitative perspective, polar molecules tend to show more curvature in the relevant saturation curve and therefore require one to two more iteration steps.

Figure 5 shows the system-size dependence of estimates for the vapor pressure of cyclohexane at 500 K. The GC-based results display a statistically relevant shift with variation in system size. The vapor pressure computed with the smallest system ($V = 2.2 \times 10^4 \text{ \AA}^3$, $N \approx 82$ within the liquid phase) is 1.3% larger than that in the extrapolated infinite-system limit. Therefore, while the shift is statistically significant, the actual difference between finite-size and thermodynamic-limit values is relatively small. The NPT-based results indicate that increasing the system size beyond 100 molecules has a negligible impact on saturation properties.

Figures 6 and 7 contain saturated density and vapor pressure data, respectively, for pyrene and squalane. We are not familiar with literature data for the pyrene model examined here, so we simply compare our direct GC, direct NPT, and NPT-TEE results. Our direct GC results for squalane are compared to Gibbs ensemble data provided by Zhuravlev et al. In all cases, the results are consistent. The direct GC and NPT data for pyrene highlight the relative precision of the two approaches. Generally speaking, we find that the direct GC approach provides saturation properties with higher precision than the direct NPT method for similar computational effort and comparable system sizes. The precision of both techniques is influenced by the ability of the method to sample the inhomogeneous region that connects the vapor and liquid phases. It is within this region that one often observes inhomogeneous structures like spherically shaped droplets and

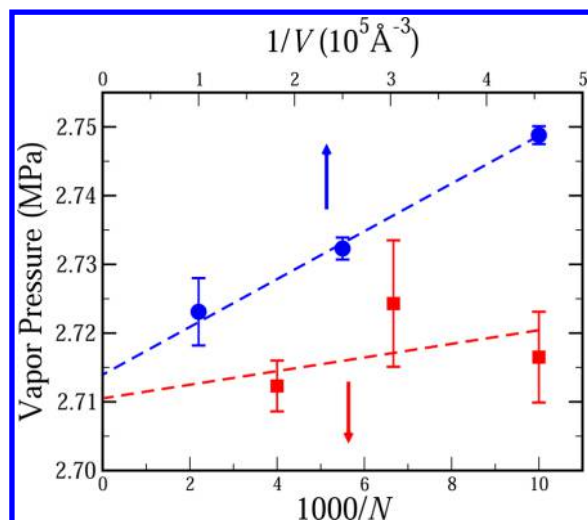


Figure 5. System-size dependence of vapor pressure estimates for cyclohexane at 500 K. Filled blue circles and filled red squares represent the data from direct GCMC and direct NPT simulations, respectively. The results obtained from GCMC and NPT simulations are plotted as a function of reciprocal cell volume (upper axis) and reciprocal number of molecules (lower axis), respectively. The dashed lines provide a linear extrapolation to the infinite system size.

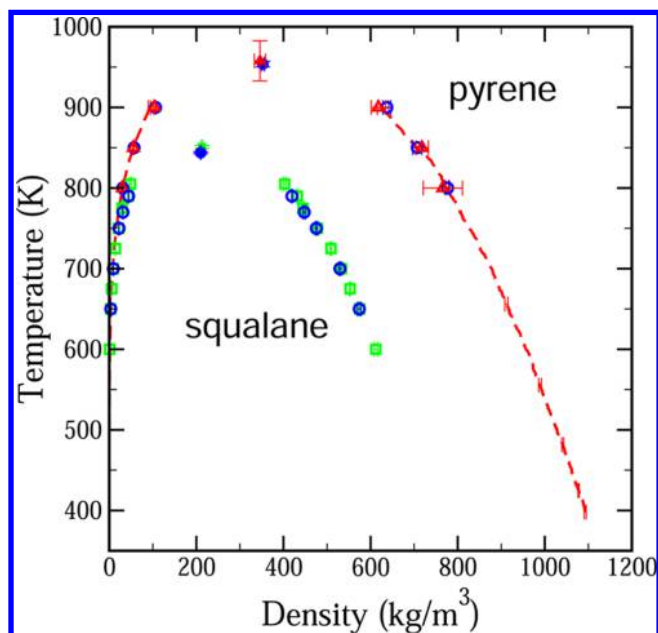


Figure 6. Temperature dependence of the saturated densities for pyrene and squalane. Symbols are defined in the same manner as in Figure 1. In this case, open green squares correspond to the Gibbs ensemble results for squalane of Zhuravlev et al.⁶⁸

cavities.^{25,26} Molecular insertions and deletions with the GC ensemble facilitate the formation and destruction of such structures, which, in turn, leads to better relative precision. As we have noted in earlier studies,⁶³ it is these inhomogeneous structures that render both of the direct techniques discussed here ineffective at sufficiently low temperature. The TEE methods pursued in this work provide a solution to this problem. By coupling the direct and TEE approaches one can compute saturation properties to very low temperature in an efficient manner. For example, the NPT-TEE data for pyrene extend to a reduced temperature T/T_c of 0.40.

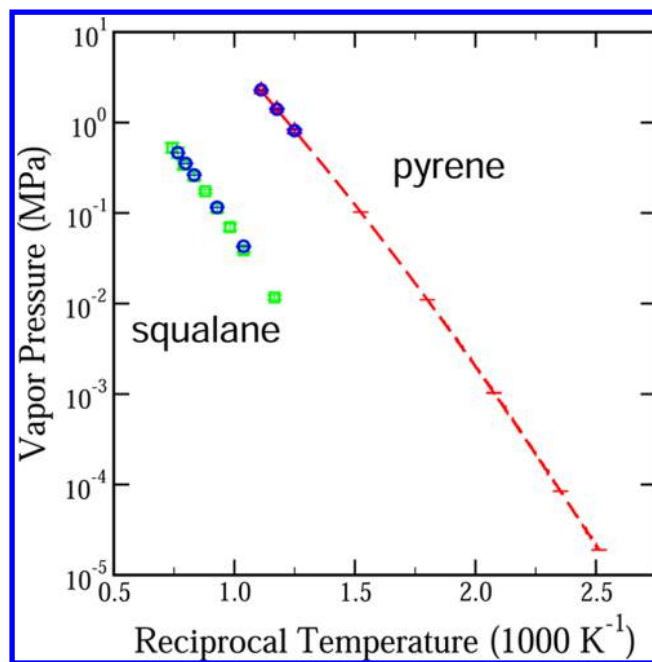


Figure 7. Temperature dependence of the vapor pressures for pyrene and squalane. Symbols are defined in the same manner as in Figure 6.

Squalane and pyrene differ in the extent to which configurational bias (CB) can assist with the insertion of molecules. Squalane is a flexible chain with modest branching. Therefore, there are many intramolecular degrees of freedom that one can exploit while growing a molecule with the CB technique. For molecules with branched topologies, the coupled-decoupled CB algorithm developed by Martin and Siepmann⁶⁹ provides an efficient means to grow molecules. In contrast, pyrene is a relatively rigid, essentially planar, aromatic molecule with limited intramolecular degrees of freedom. When working with perfectly rigid models for aromatic molecules, one can vary the center of mass and the three rotational degrees of freedom only when inserting a molecule. Flexible models for aromatic compounds are more challenging to work with, because special CB algorithms that ensure ring closure are needed.^{44,89–91} Overall, CB is of limited use in facilitating the molecule exchange process for these compounds. An attractive feature of the growth expanded ensemble approach pursued here is that its implementation does not depend upon the architecture of the molecule. The same underlying strategy (and simulation code) is used for very flexible and almost-rigid molecules. Moreover, the number of stages that one employs to grow a molecule is tunable. Finally, we note that an alternative staged exchange process for polycyclic aromatic hydrocarbons was recently introduced by Rafferty et al.⁹²

Working with squalane allows us to better understand how the choice of reservoir Hamiltonian impacts saturation properties. At the outset of this project, we naively opted to work with an ideal chain reservoir, since it is this reference state that is implicitly used in many “on-the-fly” schemes for generating molecular configurations.⁴⁵ GC simulations with the ideal chain reservoir produce squalane vapor pressure values (not shown in Figure 7) that are 7%–15% lower than those reported by Zhuravlev et al.,⁶⁸ with the discrepancy increasing with decreasing temperature. Nonbonded intramolecular Lennard-Jones interactions are ignored within the ideal chain reservoir. As a result, squalane molecules adopt configurations

within the reservoir that are expanded relative to those found within the low-density gas phase. Once a reservoir molecule is inserted into the gas phase, hybrid MC moves molecules to the relatively collapsed configurations promoted by non-bonded interactions (vapor and liquid phase radii of gyration at select temperatures are provided by Zhuravlev et al.⁶⁸). However, the rate at which this transformation takes place is slow, relative to the rate at which molecules are exchanged between the reservoir and the system. This improper sampling of the vapor-phase molecular configurations leads to inaccurate estimates of saturation properties. This problem is overcome by employing an ideal gas reservoir. Within such a reservoir, molecules are influenced by the same intramolecular interactions present within the real system. Therefore, the molecular configuration space sampled within the reservoir is consistent with that sampled within the low-density gas phase. When adopting this approach, we obtain saturation properties that agree with those reported by Zhuravlev et al. Finally, we examine how the nature of the reservoir affects the saturation properties of *n*-octane. In this case, ideal chain estimates for the vapor pressure are 0.5%–2% lower than those obtained with an ideal gas reservoir, which is a small but statistically significant difference.

Figures 8 and 9 contain saturated density and vapor pressure data, respectively, for water and propanol. For the rigid SPC/E model, we include direct GC, direct NPT, GC-TEE, NPT-TEE, Towhee,¹⁰ and Gibbs ensemble data from Errington and

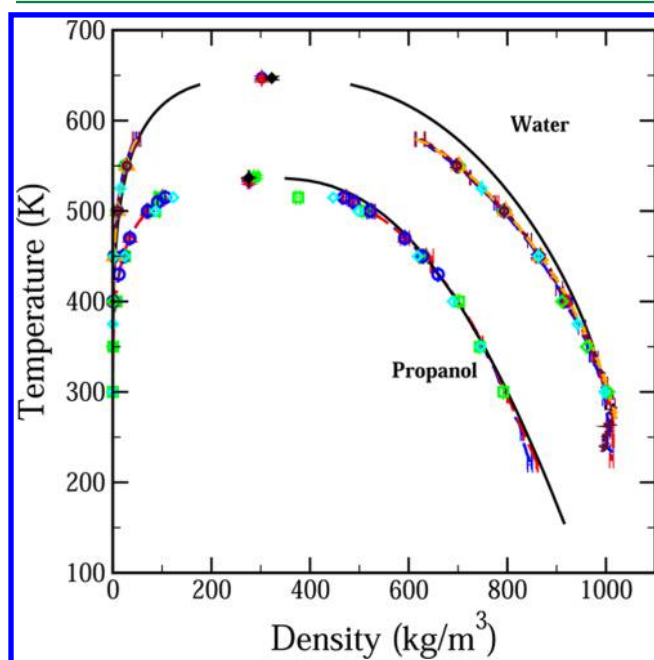


Figure 8. Temperature dependence of the saturated densities for water and propanol. Data are provided for the direct GCMC (open blue circles for rigid SPC/E and propanol, open maroon circles for flexible SPC), direct NPT (open red triangles for rigid SPC/E and propanol, open orange triangles for flexible SPC), GC-TEE (dashed blue lines for rigid SPC/E and propanol, dashed maroon line for flexible SPC), NPT-TEE (dashed red lines for rigid SPC/E and propanol, dashed orange line for flexible SPC) simulations. Open cyan diamonds represent data generated for propanol using Towhee code.^{10,84} Open green diamonds correspond to the results for rigid SPC/E model by Errington and Panagiotopoulos using Gibbs ensemble simulations.⁵ Open green squares correspond to the results for propanol by Chen et al. using Gibbs ensemble simulations.⁶⁵

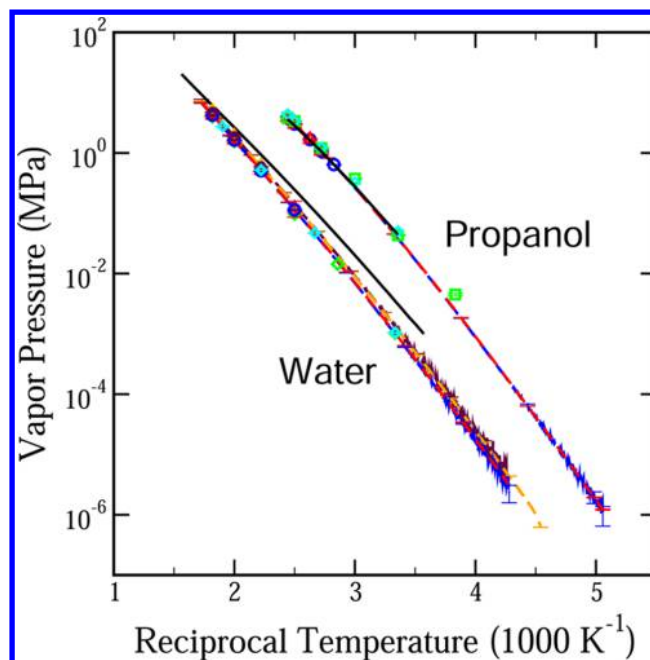


Figure 9. Temperature dependence of the vapor pressures for water and propanol. The data for propanol are shifted to the right by 500 K^{-1} , for the sake of clarity. Symbols are defined in the same manner as in Figure 8.

Panagiotopoulos.⁵ The results generated here compare favorably with those from previous studies. We again find that the precision of saturation properties generated with the direct GC approach is considerably better than with the analogous NPT approach. The saturation properties of the flexible SPC model are quantitatively similar to those for the rigid SPC/E model. Introducing intramolecular flexibility to the SPC model has a significant impact on the phase diagram. The critical temperature of the flexible model is ~ 50 K higher than the rigid model and the vapor pressures of the flexible model are, on average, a factor of 1.6 lower than the rigid version. This observation regarding the flexibility of the O–H bond is not surprising, since many other studies have noted the manner in which thermophysical properties are influenced by this factor.^{93–95} The TEE approaches enable us to compute saturation properties to temperatures as low as 225 K. Both the flexible SPC and rigid SPC/E models show a density maximum along the liquid branch of the coexistence curve, with the maximum for the flexible SPC model more pronounced.

Direct GC, direct NPT, GC-TEE, NPT-TEE, Towhee,¹⁰ and Gibbs ensemble data from Chen et al.⁶⁵ are provided for propanol. The latter two data sets were generated with a rigid bond length version of the model. Similar to what was observed for the SPC case, we find that allowing the O–H bond to be flexible leads to a decrease in the vapor pressure. The shift of $\sim 25\%$ for propanol is considerably smaller than that for SPC water. Figure 10 provides the variation of the propanol vapor pressure at 510 K with system size. Results are qualitatively similar to those for cyclohexane. The quantitative shift in the GC vapor pressure over a similar range of simulation cell volumes is also consistent with the cyclohexane case. It appears that the influence of system size is similar for systems involving both polar and nonpolar molecules.

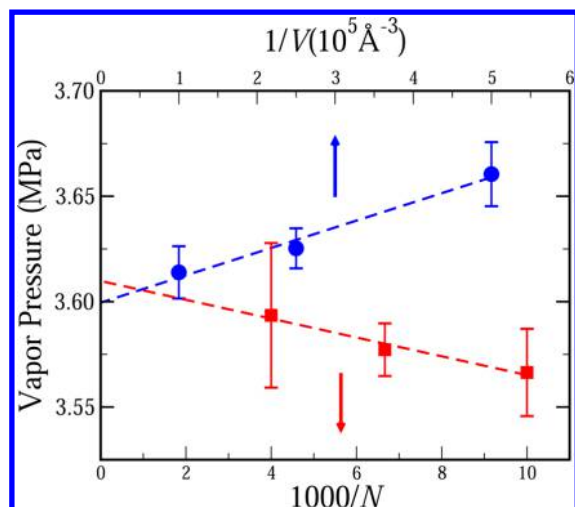


Figure 10. System-size dependence of vapor pressure estimates for propanol at 510 K. The symbols are defined in the same manner as in Figure 5.

6. CONCLUSION

We have discussed various computational strategies that one can use to determine the liquid–vapor phase coexistence properties of complex molecules. We first focused on grand canonical- and isothermal–isobaric-based methods for directly locating a saturation point at a given temperature. Reservoir and growth expanded ensemble techniques were combined to facilitate the molecule exchange process within the grand canonical approach. The grand canonical version of the direct strategy proved to be particularly robust, providing accurate and precise saturation values for the conditions explored here. The isothermal–isobaric version, while often effective, proved less precise, typically rendering statistical uncertainties for saturation properties that exceeded the grand canonical analog.

We also outlined methods for tracing saturation curves over a wide range of temperatures. In this scheme, an expanded ensemble technique is used to determine the free energy of the liquid and vapor phases along a proposed saturation curve and density probability distributions are subsequently used to correct the location of the saturation curve. Here, we found the isothermal–isobaric version of this approach to be superior to the grand canonical analogue. The isothermal–isobaric version proved particularly effective at low temperatures where molecule exchanges become difficult to complete within the grand canonical scheme.

The methods discussed here are general in nature and should prove effective when working with a wide range of complex fluids. We expect that the strategies will also prove useful for studying interfacial phenomena. We are now applying these methods to investigate the phase behavior of realistic models for ionic liquids. The methods are proving robust, providing a means for us to access saturation properties at room temperature.

■ APPENDIX

Computing Probability Distributions

We couple Wang–Landau,^{77,78} transition matrix,^{3,79–83} and multicanonical⁵⁸ Monte Carlo (MC) methods with parallel processing to calculate probability distributions. We typically employ the following recipe. In what follows below, we generically refer to the macrovariable as X , the probability

distribution as $\Pi(X)$, and the relevant dimensionless free energy curve as $\beta C(X) = -\ln \Pi(X)$.

(1) We begin by dividing the macrostate space (e.g., density or inverse temperature range) into a series of nonoverlapping windows (e.g., Window 1 spans from $X = 0$ to 50, Window 2 from $X = 50$ to 100, Window 3 from $X = 100$ to 150, etc.). Within a parallel processing environment, each window is assigned to a specific thread. The main objective of this first step is to obtain a series of equilibrated configurations that sit within the specified macrovariable limits. When working with an intensive macrovariable (e.g., temperature), this criterion is easy to meet. However, when working with an extensive macrovariable (e.g., particle number), one often must “force” the system into the appropriate limits. For example, when performing a grand canonical simulation, we do not allow particle insertions (removals) until the particle number decreases (increases) to a value commensurate with the window limits. This step typically requires relatively little simulation time.

(2) The main objective of the second step is to quickly obtain an approximate version of the probability distribution. We use the Wang–Landau algorithm^{77,78} to accomplish this goal. This approximate probability distribution enables us to quickly achieve near-uniform sampling of the macrovariable within the steps that follow. Moreover, the function provides input that enables us to eventually determine free energy differences using a pseudo-optimized Bennett scheme.⁹⁶ We typically begin the Wang–Landau algorithm with an update factor $\ln f$ between 1.0 and 10^{-2} and terminate the process when $\ln f$ is $<10^{-5}$. This step provides $\beta C_{WL}(X) = -\ln \Pi_{WL}(X)$. We note that some groups^{16,18,19,97} acquire production-quality data by continuing the Wang–Landau convergence process to smaller values of $\ln f$ (e.g., 10^{-8}). While this represents a reasonable approach, we find aspects of the transition matrix (TMMC) scheme to be attractive and, therefore, prefer to use this approach to obtain production-quality data. More specifically, the TMMC approach relies upon equilibrium sampling of configuration space, shows good convergence,⁹⁸ and is directly connected to the Bennett acceptance ratio method⁹⁶ (see below).

(3) We now direct our attention toward obtaining data via the transition matrix approach. Before performing the production runs, we change the windowing scheme and complete a very brief equilibration run. Specifically, we now employ overlapping windows (e.g., Window 1 spans from $X = 0$ to 100, Window 2 from $X = 50$ to 150, Window 3 from $X = 100$ to 200, etc.), with each window assigned to a thread within a parallel processing environment. This setup allows us to periodically swap window limits, such that a given configuration is able to sample the entire macrovariable space. Therefore, within this third step, we simply reorganize the windows and “push” configurations to satisfy the window limits, if needed. This step requires relatively little simulation time.

(4) We now use the transition matrix Monte Carlo algorithm^{58–62} to compute the probability distribution. We conduct a standard Metropolis Monte Carlo simulation^{45,99} and employ a simple bookkeeping scheme to track transition probabilities. At regular intervals during the simulation, we pause to compute the probability distribution using the transition probability data collected to that point. For regions of the macrostate space in which insufficient data are available, we use estimates from the Wang–Landau run (Step 2) to complete the probability distribution. The current estimate of

the probability distribution is used within a multicanonical sampling scheme⁵⁸ to promote uniform sampling of macrostate space, as is outlined in ref 3. It is during these pauses that we also look for opportunities to swap window limits. For completeness, we compute four versions of the probability distribution. Two of these probability distributions, Π_{MU} and Π_{BU} , use the unweighted Metropolis⁹⁹ and Barker¹⁰⁰ trial move acceptance probabilities ($p_{\text{M,acc}}$ and $p_{\text{B,acc}}$, respectively):

$$\Pi_{\text{MU}}(X_2) = \Pi_{\text{MU}}(X_1) \frac{\langle p_{\text{M,acc}}(X_1 \rightarrow X_2) \rangle}{\langle p_{\text{M,acc}}(X_2 \rightarrow X_1) \rangle} \quad (19)$$

with

$$p_{\text{M,acc}}(x_1 \rightarrow x_2) = \min[1, e^{-\beta\Delta\psi}] \quad (20)$$

and

$$\Pi_{\text{BU}}(X_2) = \Pi_{\text{BU}}(X_1) \frac{\langle p_{\text{B,acc}}(X_1 \rightarrow X_2) \rangle}{\langle p_{\text{B,acc}}(X_2 \rightarrow X_1) \rangle} \quad (21)$$

with

$$p_{\text{B,acc}}(x_1 \rightarrow x_2) = \frac{e^{-\beta\Delta\psi}}{1 + e^{-\beta\Delta\psi}} \quad (22)$$

where x_1 and x_2 correspond to microstates representative of the macrostates X_1 and X_2 , respectively, and $\Delta\psi$ gives the change in the sampling distribution (typically $e^{-\beta\Delta\psi} = \pi(x_2)/\pi(x_1)$, where π is the microstate probability) when transitioning from microstate x_1 to x_2 . Fenwick and Escobedo have shown that using the Barker scheme is equivalent to computing the free-energy difference between macrostates using the unoptimized Bennett acceptance ratio method.¹⁰¹ We also compute “optimized” analogs to these probability distributions, Π_{MO} and Π_{BO} , in a manner that leverages the data collected during the Wang–Landau run completed in Step 2. Specifically, we construct the following distributions:

$$\Pi_{\text{MO}}(X_2) = \Pi_{\text{MO}}(X_1) e^{-\beta\Delta C_{\text{WL}}} \frac{\langle g_{\text{M,acc}}(X_1 \rightarrow X_2) \rangle}{\langle g_{\text{M,acc}}(X_2 \rightarrow X_1) \rangle} \quad (23)$$

with

$$g_{\text{M,acc}}(x_1 \rightarrow x_2) = \min[1, e^{-\beta(\Delta\psi - \Delta C_{\text{WL}})}] \quad (24)$$

and

$$\Pi_{\text{BO}}(X_2) = \Pi_{\text{BO}}(X_1) e^{-\beta\Delta C_{\text{WL}}} \frac{\langle g_{\text{B,acc}}(X_1 \rightarrow X_2) \rangle}{\langle g_{\text{B,acc}}(X_2 \rightarrow X_1) \rangle} \quad (25)$$

with

$$g_{\text{B,acc}}(x_1 \rightarrow x_2) = \frac{e^{-\beta(\Delta\psi - \Delta C_{\text{WL}})}}{1 + e^{-\beta(\Delta\psi - \Delta C_{\text{WL}})}} \quad (26)$$

When the free-energy curve provided by the Wang–Landau method closely approximates that obtained from the BO version of the transition matrix algorithm, the BO approach reduces to the optimized version of the Bennett acceptance ratio method.⁹⁶ In other words, the Wang–Landau scheme provides the free-energy estimates necessary to implement the

optimized version of the Bennett method. In practice, we have not observed a case in which Π_{MO} and Π_{BO} differ in any meaningful way. It is one of these distributions that we use to construct the multicanonical weighting function and to complete all phase equilibrium calculations. We did not observe significant differences between the unoptimized and optimized versions of the probability distributions collected in this work. However, more generally, we have observed such cases. For example, we find differences when working with soft fluids (e.g., Gaussian core model).¹⁰²

Finally, we note a couple of ways in which one may wish to modify this scheme. First, several authors^{103–108} have noted that a uniform sampling of macrostate space is not optimal. Instead, one should direct resources to those regions of macrostate space that prove the most challenging. These authors have introduced algorithms for directing the sampling in this manner, which could be incorporated within the general approach presented here. Second, in the discussion above, we note an assignment of one window per thread. In practice, one should consider the tradeoff between a scenario in which one has many windows that proceed through configuration space relatively slowly and one in which one uses relatively few windows that proceed through configuration space quickly by employing a local parallel processing scheme, such as OpenMP.

AUTHOR INFORMATION

Corresponding Author

*E-mail address: jerring@buffalo.edu.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We gratefully acknowledge the financial support of the National Science Foundation (Grant No. CHE-1012356). Computational resources were provided in part by the University at Buffalo Center for Computational Research and the Rensselaer Polytechnic Institute Computational Center for Nanotechnology Innovations.

REFERENCES

- (1) Ferrenberg, A. M.; Swendsen, R. H. *Phys. Rev. Lett.* **1988**, *61*, 2635–2638.
- (2) Panagiotopoulos, A. Z. *J. Phys.—Condens. Matter* **2000**, *12*, R25–R52.
- (3) Errington, J. R. *J. Chem. Phys.* **2003**, *118*, 9915–9925.
- (4) Errington, J. R. *Phys. Rev. E* **2003**, *67*, 012102.
- (5) Errington, J. R.; Panagiotopoulos, A. Z. *J. Phys. Chem. B* **1998**, *102*, 7470–7475.
- (6) Errington, J. R.; Panagiotopoulos, A. Z. *J. Chem. Phys.* **1998**, *109*, 1093–1100.
- (7) Errington, J. R.; Panagiotopoulos, A. Z. *J. Chem. Phys.* **1999**, *111*, 9731–9738.
- (8) Errington, J. R.; Panagiotopoulos, A. Z. *J. Phys. Chem. B* **1999**, *103*, 6314–6322.
- (9) Potoff, J. J.; Errington, J. R.; Panagiotopoulos, A. Z. *Mol. Phys.* **1999**, *97*, 1073–1083.
- (10) Paluch, A. S.; Shen, V. K.; Errington, J. R. *Ind. Eng. Chem. Res.* **2008**, *47*, 4533–4541.
- (11) de Pablo, J. J.; Yan, Q. L.; Escobedo, F. A. *Annu. Rev. Phys. Chem.* **1999**, *50*, 377–411.
- (12) Errington, J. R. *Langmuir* **2004**, *20*, 3798–3804.
- (13) Escobedo, F. A. *Phys. Rev. E* **2006**, *73*, 056701.
- (14) Liu, Y.; Panagiotopoulos, A. Z.; Debenedetti, P. G. *J. Phys. Chem. B* **2011**, *115*, 6629–6635.

- (15) Rosch, T. W.; Errington, J. R. *J. Phys. Chem. B* **2007**, *111*, 12591–12598.
- (16) Ganzenmuller, G.; Camp, P. J. *J. Chem. Phys.* **2007**, *127*, 154504.
- (17) Desgranges, C.; Delhommelle, J. *J. Chem. Phys.* **2009**, *130*, 244109.
- (18) Aleksandrov, T.; Desgranges, C.; Delhommelle, J. *Fluid Phase Equilib.* **2010**, *287*, 79–83.
- (19) Desgranges, C.; Hicks, J. M.; Magness, A.; Delhommelle, J. *Mol. Phys.* **2010**, *108*, 151–158.
- (20) Muller, M.; MacDowell, L. G. *J. Phys.—Condens. Matter* **2003**, *15*, R609–R653.
- (21) Das, S. K.; Binder, K. *Phys. Rev. E* **2011**, *84*, 061607.
- (22) Rathore, N.; Knotts, T. A.; de Pablo, J. J. *J. Chem. Phys.* **2003**, *118*, 4285–4290.
- (23) Orkoulas, G.; Panagiotopoulos, A. Z. *J. Chem. Phys.* **1999**, *110*, 1581–1590.
- (24) Escobedo, F. A.; dePablo, J. J. *J. Chem. Phys.* **1996**, *105*, 4391–4394.
- (25) MacDowell, L. G.; Shen, V. K.; Errington, J. R. *J. Chem. Phys.* **2006**, *125*, 034705.
- (26) Godawat, R.; Jamadagni, S. N.; Errington, J. R.; Garde, S. *Ind. Eng. Chem. Res.* **2008**, *47*, 3582–3590.
- (27) Rane, K. S.; Kumar, V.; Errington, J. R. *J. Chem. Phys.* **2011**, *135*, 234102.
- (28) Orkoulas, G. *J. Chem. Phys.* **2010**, *133*, 111104.
- (29) Kofke, D. A. *Mol. Phys.* **1993**, *78*, 1331–1336.
- (30) Kofke, D. A. *J. Chem. Phys.* **1993**, *98*, 4149–4162.
- (31) Panagiotopoulos, A. Z. *Mol. Phys.* **1987**, *61*, 813–826.
- (32) Panagiotopoulos, A. Z. *Int. J. Thermophys.* **1989**, *10*, 447–457.
- (33) Frenkel, D.; Mooij, G.; Smit, B. *J. Phys.—Condens. Matter* **1992**, *4*, 3053–3076.
- (34) Siepmann, J. I.; Frenkel, D. *Mol. Phys.* **1992**, *75*, 59–70.
- (35) Depablo, J. J.; Laso, M.; Suter, U. W. *J. Chem. Phys.* **1992**, *96*, 2395–2403.
- (36) Errington, J. R.; Wilbert, D. W. *Phys. Rev. Lett.* **2005**, *95*, 226107.
- (37) Sellers, M. S.; Errington, J. R. *J. Phys. Chem. C* **2008**, *112*, 12905–12913.
- (38) Panagiotopoulos, A. Z. *Mol. Phys.* **1987**, *62*, 701–719.
- (39) Chen, B.; Siepmann, J. I.; Klein, M. L. *J. Am. Chem. Soc.* **2002**, *124*, 12232–12237.
- (40) Wick, C. D.; Siepmann, J. I.; Schure, M. R. *Anal. Chem.* **2002**, *74*, 3518–3524.
- (41) Martin, M. G.; Siepmann, J. I.; Schure, M. R. *J. Phys. Chem. B* **1999**, *103*, 11191–11195.
- (42) Lee, J. S.; Wick, C. D.; Stubbs, J. M.; Siepmann, J. I. *Mol. Phys.* **2005**, *103*, 99–104.
- (43) Bai, P.; Siepmann, J. I. *Fluid Phase Equilib.* **2011**, *310*, 11–18.
- (44) Martin, M. G.; Frischknecht, A. L. *Mol. Phys.* **2006**, *104*, 2439–2456.
- (45) Frenkel, D.; Smit, B. *Understanding Molecular Simulation*, Second Edition; Academic Press: London, 2002.
- (46) Macedonia, M. D.; Maginn, E. J. *Mol. Phys.* **1999**, *96*, 1375–1390.
- (47) Shah, J. K.; Maginn, E. J. *J. Chem. Phys.* **2011**, *135*, 134121.
- (48) Duane, S.; Kennedy, A. D.; Pendleton, B. J.; Roweth, D. *Phys. Lett. B* **1987**, *195*, 216–222.
- (49) Lal, M. *Mol. Phys.* **1969**, *17*, 57–64.
- (50) Madras, N.; Sokal, A. D. *J. Stat. Phys.* **1988**, *50*, 109–186.
- (51) Binder, K. *Monte Carlo and Molecular Dynamics Simulations in Polymer Science*; Oxford University Press: New York, 1995.
- (52) Cagin, T.; Pettitt, B. M. *Mol. Phys.* **1991**, *72*, 169–175.
- (53) Lo, C. M.; Palmer, B. J. *J. Chem. Phys.* **1995**, *102*, 925–931.
- (54) Shi, W.; Maginn, E. J. *J. Chem. Theory Comput.* **2007**, *3*, 1451–1463.
- (55) Singh, J. K.; Errington, J. R. *J. Phys. Chem. B* **2006**, *110*, 1369–1376.
- (56) Desgranges, C.; Delhommelle, J. *J. Chem. Phys.* **2012**, *136*, 184107.
- (57) Shroll, R. M.; Smith, D. E. *J. Chem. Phys.* **1999**, *111*, 9025–9033.
- (58) Berg, B. A.; Neuhaus, T. *Phys. Rev. Lett.* **1992**, *68*, 9–12.
- (59) Mehlig, B.; Heermann, D. W.; Forrest, B. M. *Phys. Rev. B* **1992**, *45*, 679–685.
- (60) Rosso, L.; Tuckerman, M. E. *Mol. Simul.* **2002**, *28*, 91–112.
- (61) Lyubartsev, A. P.; Martsinovski, A. A.; Shevkunov, S. V.; Vorontsovvelaminov, P. N. *J. Chem. Phys.* **1992**, *96*, 1776–1783.
- (62) Kumar, V.; Sridhar, S.; Errington, J. R. *J. Chem. Phys.* **2011**, *135*, 184702.
- (63) Grzelak, E. M.; Errington, J. R. *Langmuir* **2010**, *26*, 13297–13304.
- (64) Martin, M. G.; Siepmann, J. I. *J. Phys. Chem. B* **1998**, *102*, 2569–2577.
- (65) Chen, B.; Potoff, J. J.; Siepmann, J. I. *J. Phys. Chem. B* **2001**, *105*, 3093–3104.
- (66) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*, 6269–6271.
- (67) Rai, N.; Siepmann, J. I. *J. Phys. Chem. B* **2007**, *111*, 10790–10799.
- (68) Zhuravlev, N. D.; Martin, M. G.; Siepmann, J. I. *Fluid Phase Equilib.* **2002**, *202*, 307–324.
- (69) Martin, M. G.; Siepmann, J. I. *J. Phys. Chem. B* **1999**, *103*, 4508–4517.
- (70) Wick, C. D.; Martin, M. G.; Siepmann, J. I. *J. Phys. Chem. B* **2000**, *104*, 8008–8016.
- (71) Damm, W.; Frontera, A.; Tirado-Rives, J.; Jorgensen, W. L. *J. Comput. Chem.* **1997**, *18*, 1955–1970.
- (72) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S., Jr.; Weiner, P. *J. Am. Chem. Soc.* **1984**, *106*, 765–784.
- (73) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kucera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.
- (74) Tuckerman, M.; Berne, B. J.; Martyna, G. J. *J. Chem. Phys.* **1992**, *97*, 1990–2001.
- (75) Toxvaerd, S.; Heilmann, O. J.; Ingebrigtsen, T.; Schroder, T. B.; Dyre, J. C. *J. Chem. Phys.* **2009**, *131*, 064102.
- (76) Wu, Y. J.; Tepper, H. L.; Voth, G. A. *J. Chem. Phys.* **2006**, *124*, 024503.
- (77) Wang, F. G.; Landau, D. P. *Phys. Rev. Lett.* **2001**, *86*, 2050–2053.
- (78) Wang, F. G.; Landau, D. P. *Phys. Rev. E* **2001**, *64*, 056101.
- (79) Wang, J. S.; Swendsen, R. H. *J. Stat. Phys.* **2002**, *106*, 245–285.
- (80) Wang, J. S.; Tay, T. K.; Swendsen, R. H. *Phys. Rev. Lett.* **1999**, *82*, 476–479.
- (81) Smith, G. R.; Bruce, A. D. *J. Phys. A: Math. Theor.* **1995**, *28*, 6623–6643.
- (82) Fitzgerald, M.; Picard, R. R.; Silver, R. N. *Europhys. Lett.* **1999**, *46*, 282–287.
- (83) Fitzgerald, M.; Picard, R. R.; Silver, R. N. *J. Stat. Phys.* **2000**, *98*, 321–345.
- (84) Martin, M. G. *The Towhee Code*, <http://towhee.sourceforge.net> (accessed 01/15/2013).
- (85) NIST Chemistry WebBook, <http://webbook.nist.gov/chemistry/> (accessed 01/15/2013).
- (86) Muller, E. A.; Mejia, A. *J. Phys. Chem. B* **2011**, *115*, 12822–12834.
- (87) Mundy, C. J.; Siepmann, J. I.; Klein, M. L. *J. Chem. Phys.* **1995**, *102*, 3376–3380.
- (88) Patel, S.; Wilding, W. V.; Rowley, R. L. *J. Chem. Phys.* **2011**, *134*, 024101.
- (89) Vendruscolo, M. *J. Chem. Phys.* **1997**, *106*, 2970–2976.
- (90) Wick, C. D.; Siepmann, J. I. *Macromolecules* **2000**, *33*, 7207–7218.
- (91) Martin, M. G.; Thompson, A. P. *Fluid Phase Equilib.* **2004**, *217*, 105–110.

- (92) Rafferty, J. L.; Siepmann, J. I.; Schure, M. R. *J. Chromatogr. A* **2011**, *1218*, 9183–9193.
- (93) Raabe, G.; Sadus, R. J. *J. Chem. Phys.* **2007**, *126*, 044701.
- (94) Yuet, P. K.; Blankschtein, D. *J. Phys. Chem. B* **2010**, *114*, 13786–13795.
- (95) Teleman, O.; Jonsson, B.; Engstrom, S. *Mol. Phys.* **1987**, *60*, 193–203.
- (96) Bennett, C. H. *J. Comput. Phys.* **1976**, *22*, 245–268.
- (97) Aleksandrov, T.; Desgranges, C.; Delhommelle, J. *Mol. Simul.* **2012**, *38*, 1265–1270.
- (98) Shell, M. S.; Debenedetti, P. G.; Panagiotopoulos, A. Z. *J. Chem. Phys.* **2003**, *119*, 9406–9411.
- (99) Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. N.; Teller, E. *J. Chem. Phys.* **1953**, *21*, 1087–1092.
- (100) Barker, A. A. *Austr. J. Phys.* **1965**, *18*, 119–133.
- (101) Fenwick, M. K.; Escobedo, F. A. *J. Chem. Phys.* **2004**, *120*, 3066–3074.
- (102) Krekelberg, W. P.; Kumar, T.; Mittal, J.; Errington, J. R.; Truskett, T. M. *Phys. Rev. E* **2009**, *79*, 031203.
- (103) Shell, M. S.; Debenedetti, P. G.; Panagiotopoulos, A. Z. *J. Phys. Chem. B* **2004**, *108*, 19748–19755.
- (104) Trebst, S.; Huse, D. A.; Troyer, M. *Phys. Rev. E* **2004**, *70*, 046701.
- (105) Escobedo, F. A.; Abreu, C. R. A. *J. Chem. Phys.* **2006**, *124*, 104110.
- (106) Escobedo, F. A.; Martinez-Veracoechea, F. J. *J. Chem. Phys.* **2007**, *127*, 174103.
- (107) Escobedo, F. A. *J. Chem. Phys.* **2007**, *127*, 174104.
- (108) Escobedo, F. A.; Martinez-Veracoechea, F. J. *J. Chem. Phys.* **2008**, *129*, 154107.