JCTC Journal of Chemical Theory and Computation

Continuous Fractional Component Monte Carlo: An Adaptive Biasing Method for Open System Atomistic Simulations

Wei Shi and Edward J. Maginn*

Department of Chemical and Biomolecular Engineering, University of Notre Dame, 182 Fitzpatrick Hall, Notre Dame, Indiana 46556-5637

Received January 4, 2007

Abstract: A new open system Monte Carlo procedure designed to overcome difficulties with insertion and deletion of molecules is introduced. The method utilizes gradual insertions and deletions of molecules through the use of a continuous coupling parameter and an adaptive bias potential. The method draws upon concepts from previous open system molecular dynamics and expanded ensemble Monte Carlo techniques and is applied to both the grand canonical and osmotic ensembles. It is shown to yield correct results for the volumetric properties of the Lennard-Jones fluid and water as well as the phase behavior of the CO_2-ethanol binary system.

1. Introduction and Background

The use of atomistic simulations to compute the phase behavior of fluids has reached a high level of sophistication. Accurate vapor−liquid,[1−3] liquid−liquid,[4,5] solid−liquid,[6−8] and solid−vapor[9,10] equilibria calculations have been performed for a wide range of systems. Although a large number of simulation methods have been used for this purpose, they can be split into two main categories. The first category requires that the free energy (chemical potential) be computed for the system at a given state point relative to some reference state. Given this free energy, one can locate coexistence points through the enforcement of the phase equilibria criteria, namely equality of chemical potential, temperature, and pressure. In this category are techniques such as thermodynamic integration, free energy perturbation, expanded ensembles, and histogram methods. Kofke[11,12] has provided excellent reviews of the strengths and weaknesses of these different approaches. The second category of simulation method, and the one we are concerned with here, utilizes an open system in which a chemical potential is imposed and the composition changes in response to this chemical potential. Examples include grand canonical Monte Carlo (GCMC) as well as various open system molecular dynamics techniques. The Gibbs ensemble technique can also be grouped in this category, although the chemical potential is not specified a priori with this technique. Common to both simulation categories is the need to *insert and delete* molecules from the system to either evaluate the chemical potential or change composition. The accuracy and precision of the simulations depend critically on the ability to carry out these insertion and deletion moves. Poor sampling can result when systems are dense and have slowly evolving dynamics, such that suitable "holes" in the system large enough to accommodate the inserted molecule appear infrequently. It also can occur when there are specific interactions between the inserted or deleted species and the rest of the molecules in the system. For example, in strongly hydrogen-bonding systems specific orientations are required for energies to be favorable during an insertion. For deletions of hydrogen-bonded systems, the status quo conformation may be so energetically favorable that it is exceedingly improbable that a deletion can be successful.

A large number of biasing techniques have been developed to overcome these difficulties. While the literature is too extensive to list all efforts, we note that within the Monte Carlo (MC) framework moves have been designed to search for favorable energetic locations in the system[13] as well as cavities where insertions will more likely be accepted.[14] Advanced configurational bias moves,[15−17] multicanonical methods,[18−21] and many other sophisticated sampling techniques have also been developed. Of particular relevance to the present work are the various Monte Carlo "switch"

* Corresponding author phone: (574)631-5687; fax: (574)631-8366; e-mail: ed@nd.edu.

10.1021/ct7000039 CCC: $37.00
Published on Web 05/25/2007

methods in which small "impurity" molecules are inserted and deleted into the system with great efficiency and then transformed into larger "full" molecules.[22] Such an approach has been refined and applied to alkanes,[23] alcohols,[24] and polymers.[25]

One strategy that is particularly appealing is to accompany molecule insertions or deletions with some sort of local relaxation process, such that energetic penalties can be minimized and thus the success probability of the move increased. This is one of the strategies behind various open system molecular dynamics (MD) methods.[5,26−31] The basic idea behind these approaches is similar to that involved in the MC switch methods, in which insertions and deletions are "staged". In the open system MD methods, this is accomplished by making either stochastic or deterministic changes in the coupling strength between a "fractional" molecule and the rest of the system. The composition changes as molecules are continuously created (destroyed) as the fractional molecule becomes fully coupled (decoupled) with the system. In between these changes in coupling strength, dynamics are run to enable the system to accommodate the change in the fractional molecule. The advantage is that the natural forces present can allow for collective or cooperative relaxation of the system in response to the perturbation; there is no need to develop special biasing rules ahead of time. Very recently, a new method has been proposed[32,33] that takes this idea a step further by using local energy minimization to increase the success probability of insertions and deletions.

Here we report the development and application of a method for the efficient insertion and deletion of molecules in the grand canonical (μ, V, T) and osmotic (f_1, N_2, P, T) ensembles. The method, which we call continuous fractional component Monte Carlo (CFC MC), can be generalized to other ensembles such as the Gibbs ensemble. It offers a number of attractive features for simulating the phase behavior of complex systems, including self-adapting biasing capability that requires no foreknowledge of any bias functions. CFC MC draws upon a number of previous methods, including the expanded ensemble GCMC method described by Escobedo and de Pablo,[34] the Monte Carlo "swatch" move developed by Siepmann and co-workers,[23−25] the expanded ensemble molecular dynamics method of Lísal and co-workers,[31] and Boinepalli and Attard's grand canonical molecular dynamics approach.[35] In the next section, a brief summary of related methods is provided, followed by details of the CFC MC technique. Following that, results from CFC MC simulations of three test systems are presented and shown to give equivalent results to accepted methods. We also demonstrated that some of the existing open system methods fail to give the correct results for these test systems.

2. Previous Methods

2.1. Open System Molecular Dynamics. As mentioned above, the advantage of performing insertions and deletions within an MD framework is that the forces allow the system to relax in response to the perturbations brought about by the insertion or deletion move. Cooperative and collective moves, which are difficult to treat with a priori MC biasing rules, occur more or less naturally with MD. One obvious difficulty of open system MD approaches is that standard MD algorithms require a constant number of molecules. To simulate open systems with MD, a modified Hamiltonian is required, as is a method of coupling insertion and deletion moves dynamically to the system.

Çağin and Pettitt[26,27] appear to have been the first to develop an MD method for the grand canonical ensemble. Their grand canonical molecular dynamics (GCMD) method utilizes an extended system Hamiltonian and a dynamical variable λ that links the physical system to a particle bath, thereby enabling compositional changes. The potential for λ is a function of the chemical potential of the system. Pettitt and co-workers successfully applied the GCMD method to a Lennard-Jones fluid[27,36] and water.[28,36] One problem that was observed with the GCMD method is that the dynamics can become unstable due to large accelerations in the coupling parameter when strongly repulsive regions are sampled, as frequently occurs when insertions result in significant molecular overlap. To minimize this, biased insertion and deletion moves were used.

Lo and Palmer[29] proposed an alternative extended Hamiltonian to perform GCMD simulations. This method still makes changes in the strength of the coupling between the fractional molecule and the rest of the system deterministically. The difference is that stochastic moves are made in between composition changes, thereby switching to a new Hamiltonian at each composition. Two other major differences exist. First, Lo and Palmer did not scale the mass of the transforming molecule, while the mass was scaled by λ in the work of Pettitt and co-workers. Second, while Pettitt and co-workers used a simple linear scaling to couple the fractional molecule with the system, in simulating a Lennard-Jones system Lo and Palmer chose the following functional form

$$\phi_f(r_{jf}, \lambda) = \lambda 4\epsilon\left\{\left[\frac{\sigma}{r_{jf} + (1-\lambda)\zeta}\right]^{12} - \left[\frac{\sigma}{r_{jf} + (1-\lambda)\zeta}\right]^{6}\right\} \quad (1)$$

where $\phi_f(r_{jf}, \lambda)$ is the interaction potential between the transforming "fractional" molecule and the other "integer" molecules in the system, ϵ and σ are the well depth and collision diameter for the LJ potential, j denotes integer molecules, and f denotes the fractional molecule. The parameter ζ can be adjusted to optimize the performance of the algorithm. Note that the normal Lennard-Jones potential is recovered when $\lambda = 1$. This type of scaling significantly reduces the repulsive forces present when atoms are close to one another at small values of λ and thus improves stability. Biased insertions are typically not required with this type of scaling.

Shroll and Smith[30] extended the GCMD method in two ways. First, they introduced a bias potential to improve the probability of transitions among adjacent values of λ. Without this bias potential, the system can become "stuck" at values of λ between 0 and 1 whenever the free energy barrier for λ transitions is large. This appears to not be an issue for the simple Lennard-Jones system[35] but becomes a factor for more complex systems such as water. In the spirit of expanded ensemble methods,[37] this bias potential should be close to the free energy difference between adjacent λ states to

maintain a uniform distribution of λ values. Shroll and Smith determined their bias potential for water by first computing the free energy difference between adjacent states via thermodynamic integration. The obvious drawback of this approach is that a separate thermodynamic integration must be carried out for each state point, which mitigates the advantages of using a bias potential. However, the authors note that if the free energy is insensitive to the state points, the same bias potential can be used for multiple state points. The second modification introduced by Shroll and Smith was the use of a continuous λ that "rolls over" when it reaches the boundaries of $\lambda = 0$ or $\lambda = 1$. That is, the velocity of the coupling parameter $\dot{\lambda}$ is reassigned to the new fractional molecule whenever the bounds on λ are exceeded. In the original GCMD methods, dynamic λ transitions abruptly end whenever λ exceeds it bounds. We also note that these authors found that the use of a nonlinear scaling of the potential avoided instabilities when the strongly repulsive regions of the potential were sampled.

2.2. Hybrid Methods. All of the above techniques employ fully deterministic methods for making changes in the value of the coupling parameter λ (and thus the composition) by employing an extended Hamiltonian. As with any extended system, they require the use of "time constants" or virtual masses for the extended variables, which control the rate of change of the extended variable. It has been pointed out[38] that to properly compute free energies using techniques such as this there must be a dynamical adiabatic separation between the evolution of the fractional molecule coupling and the remaining degrees of freedom. In other words, the transformation must be sufficiently slow to ensure a large separation in time scales between the dynamics of the system and that of λ. This is potentially a problem with fully dynamic methods, since there is no a priori way of knowing what the relevant "time constants" should be to ensure adiabatic separation.

It is also possible to exploit the favorable properties of dynamic insertions via *hybrid approaches* in which stochastic "open system" MC moves are used to change composition while deterministic MD trajectories are also used to relax the system. As with the former methods, particle number is a continuous variable so that insertions and deletions are not done with full molecules. Because composition changes are done stochastically using Metropolis-like acceptance rules, however, these techniques can be constructed to satisfy detailed balance and thus are guaranteed to satisfy the limiting probability distribution of the ensemble. Compared to extended system approaches, these methods are relatively simple to implement.

Boinepalli and Attard[35] have proposed such a hybrid method for the grand canonical ensemble. For a given state of the fractional molecule, a regular microcanonical MD step is taken, followed by a stochastic attempt to increase or decrease the value of the coupling parameter for a fractional particle. This method was implemented for the Lennard-Jones system and found to give correct volumetric properties.

Lísal and co-workers[31] have recently proposed a hybrid MD−MC method called expanded-ensemble osmotic molecular dynamics. As with the Boinepalli and Attard method, composition changes are treated stochastically by making gradual changes in the coupling strength of a fractional molecule with the other molecules in a system. As did Shroll and Smith, these authors used a bias potential to overcome problems with the coupling parameter becoming "stuck" at values between 0 and 1. The bias must be removed in the acceptance rules to obtain proper averages. Following Shroll and Smith, thermodynamic integration was used to determine the values for the bias potential.

3. Continuous Fractional Component Monte Carlo Method

The CFC MC method descibed here draws on a number of features of the previous methods outlined above. It is a hybrid method in which continuous composition changes are handled stochastically, while MD and, if desired, MC moves are used to relax the system. Extended variables (or bias potentials) are used to overcome large free energy barriers for insertion or deletion moves. Unlike previous methods, these bias factors are determined in a self-adaptive manner via an interative Wang−Landau approach.[39] A nonlinear potential is used to avoid problems with repulsive overlap. We implement the method in both the expanded grand canonical and expanded osmotic ensembles and then use the technique to compute volumetric properties of the Lennard-Jones fluid and the simple point charge (SPC) model of water as well as phase equilibria for CO_2/ethanol mixtures. We demonstrate the accuracy and efficiency of the method by comparing CFC MC results with previous simulation studies and our own calculations using existing methods.

3.1. Grand Canonical Ensemble. In this section, we show how CFC MC is implemented in the grand canonical ensemble. To prove the validity of any stochastic move, it is sufficient to show that it satisfies detailed balance, which we write as

$$\Pi_{mn}\alpha_{mn}p_m = \Pi_{nm}\alpha_{nm}p_n \tag{2}$$

where Π_{mn} is the one-step transition or acceptance probability of going from state m to state n, α_{mn} is the probability of attempting such a move, and p_m and p_n are the ensemble probability distribution functions for states m and n, respectively. Now consider a system at fixed temperature T, chemical potential μ, and volume V that contains two types of fully flexible molecules containing n_a atoms: N fully interacting "integer" molecules and a single "fractional" molecule whose potential couples to the integer molecules via a pseudocontinuous coupling parameter λ of the type discussed above that ranges from 0 to 1. The fractional molecule can exist in a large number of states M, and at each state j there is an expanded system variable η_j associated with it. The partition function for this system is that of an expanded grand canonical ensemble having the following form[34,40]

$$\Psi(\mu, V, T, \eta) = \sum_{N=0}^{\infty}\sum_{j=1}^{M}\frac{q_t^{N+1}\, exp[\beta\mu N]\, \exp[\eta_j]}{N!}Z(N, j, V, T) \tag{3}$$

where $\beta = 1/k_BT$. Z is the configurational integral given by

$$Z(N, j, V, T) = \int \exp[-\beta\phi(\mathbf{q})] \prod_{i=1}^{N+1} d\mathbf{q}_i \tag{4}$$

where $\mathbf{q}_i$ is the vector of generalized coordinates for molecule i, which includes both integer molecules and the fractional molecule. The variable q_t contains kinetic energy contributions from all atoms in the system (integer and fractional) as well as the Jacobian of the transformation from atomic Cartesian coordinates to internal cooordinates.[40,41]

It is often more convenient to work with fugacities f rather than chemical potentials. It is straightforward to show that the probability distribution function for some state m having N_m integer molecules and one fractional molecule is[40]

$$p_m = \frac{1}{\Psi} \frac{(f\beta V)^{N_m} q_t \Omega}{N_m!(Z^{\mathrm{IG}}/\Omega)^{N_m}} \exp[\eta_m] \exp[-\beta\phi_m] \tag{5}$$

where η_m is the bias factor associated with the particular value of λ at state m, ϕ_m is the total potential energy of state m, Z^{IG} is the ideal gas configurational integral of a single molecule, and Ω is the volume associated with the generalized coordinates of a single molecule. Formally,

$$Z^{\mathrm{IG}} = \int \exp[-\beta\phi_{\mathrm{intra}}(\mathbf{q})] d^{3n_a}q \tag{6}$$

$$\Omega = \int d^{3n_a}q \tag{7}$$

where $\phi_{\mathrm{intra}}(\mathbf{q})$ is the intramolecular contribution to the potential energy of a single molecule with generalized coordinates given by the vector $\mathbf{q}$. For the special case of an atomic species, $\phi_{\mathrm{intra}} = 0$ and $n_a = 1$, so that $Z^{\mathrm{IG}} = \Omega = V$.

Given the probability distribution function in eq 5, a set of Metropolis-like acceptance rules can be derived for making changes in the value of λ. Assume that state m consists of N_m integer molecules and one fractional molecule. If a change in the coupling parameter occurs such that $\lambda > 1$, then the existing fractional molecule is converted to an integer molecule, and a new fractional molecule is added to the system by randomly selecting a position and assigning the new fraction molecule a coupling parameter of $\lambda - 1$. The new state n contains $N_n = N_m + 1$ integer molecules and a single fractional molecule. The move is accepted with a probability reminiscent of a grand canonical ensemble insertion move

$$\Pi_{mn}^{\mathrm{ins}} = \min\left(1, \frac{\alpha_{nm}}{\alpha_{mn}} \frac{\Omega}{Z^{\mathrm{IG}}} \frac{f\beta V}{N_m + 1} \exp[\eta_n - \eta_m] \exp[-\beta(\phi_n - \phi_m)]\right) \tag{8}$$

Likewise, when a transition occurs for a system of N_n integer molecules and one fractional molecule such that $\lambda < 0$, the fractional molecule is removed from the system, and one of the remaining integer molecules is randomly assigned a coupling parameter value of $1 + \lambda$. This molecule thus becomes the new fractional molecule. The transition probability for such a move from state n to state m is similar to the grand canonical deletion move probability and is given by

$$\Pi_{nm}^{\mathrm{del}} = \min\left(1, \frac{\alpha_{mn}}{\alpha_{nm}} \frac{Z^{\mathrm{IG}}}{\Omega} \frac{N_n}{f\beta V} \exp[\eta_m - \eta_n] \exp[-\beta(\phi_m - \phi_n)]\right) \tag{9}$$

For moves in which λ changes but is still between (0,1), the acceptance probability is

$$\Pi_{mn}^{\lambda} = \min(1, \exp[\eta_n - \eta_m] \exp[-\beta(\phi_{\mathrm{inter},n} - \phi_{\mathrm{inter},m})]) \tag{10}$$

where $\phi_{\mathrm{inter},m}$ and $\phi_{\mathrm{inter},n}$ are the intermolecular contributions to the potential energy of states m and n, respectively. Note that no intramolecular energies are altered by this type of move.

Finally, thermal equilibration moves at constant values of λ occur. Any MC sampling method desired can be used for these moves.[42,43] In the present work, a hybrid Monte Carlo procedure was used[44] in which velocities are selected from a Maxwell−Boltzmann distribution, followed by a number of microcanonical MD steps. The new positions are accepted with probability

$$\Pi_{mn}^{\mathrm{HMC}} = \min(1, \exp[-\beta(H_n - H_m)]) \tag{11}$$

where H_m and H_n are the Hamiltonians of the system in states m and n.

Note that in eqs 10 and 11 the attempt probabilities α_{mn} and α_{nm} are not included, because for these moves they are equal and thus cancel out. That is, there is no bias as to whether an attempt to increase or decrease λ is made. However, the weighting factors η do influence the likelihood of a successful change in λ and are optimized for performance, as discussed below. On the other hand, the attempt probabilities have been explicitly included in eqs 8 and 9, because moves in which the number of integer molecules changes are compound moves. The probabilities associated with these moves consist of the product of the symmetric probability of increasing or decreasing the value of λ and the asymmetric probability of adding or removing a molecule from the system. There is a great deal of flexibility in the choice of these latter probabilities. In this work, when λ exceeds unity and a new fractional molecule is added to the system, a rigid conformation is chosen from a "reservoir" of ideal gas molecules generated "on the fly" during the simulation.[40,45] These conformations occur with probability

$$p^{\mathrm{IG}} = \frac{\exp[-\beta\phi_{\mathrm{intra}}(\mathbf{q})]\Omega}{Z^{\mathrm{IG}}} \tag{12}$$

Given such a conformation, a random position and orientation of this molecule is chosen, and the new fractional molecule is inserted into the system. The attempt probability for this move is $\alpha_{mn} = p^{\mathrm{IG}}$. For the reverse move when λ falls below zero, an existing integer molecule is chosen at random to become the new fractional molecule. The attempt probability for this move is $\alpha_{nm} = 1$. Thus the ratio of attempt probabilities in eqs 8 and 9 is

$$\frac{\alpha_{nm}}{\alpha_{mn}} = \frac{Z^{\mathrm{IG}}}{\exp[-\beta\phi_{\mathrm{intra}}]\Omega} \tag{13}$$

Using this method for generating conformations simplifies the acceptance ratios in eqs 8 and 9 to

$$\Pi_{mn}^{\mathrm{ins}} = \min\left(1, \frac{f\beta V}{N_m + 1} \exp[\eta_n - \eta_m] \exp[-\beta(\phi_{\mathrm{inter},n} - \phi_{\mathrm{inter},m})]\right) \quad (14)$$

and

$$\Pi_{nm}^{\mathrm{del}} = \min\left(1, \frac{N_n}{f\beta V} \exp[\eta_m - \eta_n] \exp[-\beta(\phi_{\mathrm{inter},m} - \phi_{\mathrm{inter},n})]\right) \quad (15)$$

where $\phi_{\mathrm{inter},m}$ and $\phi_{\mathrm{inter},n}$ are only the intermolecular contributions to the potential energy for states m and n, respectively. Note that the intramolecular potential energies cancel out, given that the fractional molecule's intramolecular degrees of freedom are sampled from eq 12 and the existing molecules are unperturbed by the move. Note that additional biasing moves are not used, given the gradual coupling of the fractional molecule with the system and the use of the extended system biasing factors η.

3.2. Osmotic Ensemble. The CFC MC method can also be applied to the osmotic ensemble[46−48] in which for a c-component system with k solute molecules and $c - k$ solvent molecules, the following variables are fixed: f_1, f_2, ..., f_k, N_{k+1}, N_{k+2}, ..., N_c, P, T. The number of molecules of species $1 - k$ will fluctuate in this ensemble. Such an ensemble is particularly effective in simulating solubilities, especially for light species in low-volatility solvents. Here, we develop osmotic CFC MC acceptance rules for the binary case where species 1 is the solute and species 2 is the solvent. Generalizing to multicomponent systems is relatively straightforward.

The probability density of a system of volume V containing N_1 integer solute molecules having generalized coordinates $\mathbf{q}_1$, a fractional molecule in state j at $\mathbf{q}_j$ with coupling parameter λ and associated bias factor η_j, and N_2 solvent molecules at $\mathbf{q}_2$ is[46]

$$\rho(N_1, V, \eta_j, \mathbf{q}_1, \mathbf{q}_j, \mathbf{q}_2; f_1, N_2, P, T) \propto \frac{V^{N_2}}{N_1!} \times \frac{(\beta f_1 V)^{N_1}(V \exp[\eta_j])}{(Z^{\mathrm{IG}}/\Omega)^{N_1+1}} \exp[-\beta PV] \exp[-\beta(\phi(\mathbf{q}_1, \mathbf{q}_j, \mathbf{q}_2))] \quad (16)$$

In eq 16, the dependence of the potential energy on the coordinates of all species is explicitly noted.

Changes in λ that result in an increase or decrease in the number of integer molecules are accepted with the same probability as that in eqs 14 and 15, assuming the ratio of attempt probabilities is given by eq 13. Likewise, changes in λ that do not alter the number of integer molecules are accepted according to eq 10, and thermal equilibration moves are performed with hybrid Monte Carlo (HMC) and accepted according to eq 11.

Random changes in volume are performed to maintain a constant pressure. These moves occur at constant λ and $N_{\mathrm{tot}} = N_1 + N_2 + 1$, with acceptance probability given by

$$\Pi_{mn}^{\mathrm{vol}} = \min\left(1, \exp\left[-\beta\left\{P(V_n - V_m) + (\phi_n - \phi_m) - N_{\mathrm{tot}} k_{\mathrm{B}} T \ln\frac{V_n}{V_m}\right\}\right]\right) \quad (17)$$

4. Simulation Details

4.1. Lennard-Jones. The CFC MC method was first implemented and tested for grand canonical simulations of the Lennard-Jones fluid, in which the following cut and shifted potential was used

$$\phi_{\mathrm{LJ,cut-shift}}(r) = \begin{cases} \phi_{\mathrm{LJ}}(r) - \phi_{\mathrm{LJ}}(r_c) & r \le r_c \\ 0 & r > r_c \end{cases} \quad (18)$$

where the cutoff distance was taken as $r_c = 2.5\sigma$ and ϕ_{LJ} is given by

$$\phi_{\mathrm{LJ}}(r) = 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] \quad (19)$$

To minimize problems with repulsive overlap, the following scaled potential[49] was used to describe the interaction between a fractional species and the integer species

$$\phi_f(r_{if}, \lambda) = \lambda 4\epsilon\left\{\frac{1}{\left[\zeta(1-\lambda)^2 + \left(\frac{r_{if}}{\sigma}\right)^6\right]^2} - \frac{1}{\left[\zeta(1-\lambda)^2 + \left(\frac{r_{if}}{\sigma}\right)^6\right]}\right\} \quad (20)$$

As with eq 1, r_{if} is the distance between an integer species i and fractional species f, while ζ is an adjustable parameter that was set to 0.5. This scaled potential has the correct behavior at the limits of $\lambda = 0$ and $\lambda = 1$ but is well-behaved at very close overlap, as often occurs when inserting a species.

Due to the relative ease with which λ transitions occur in this system, it was found that bias factors were not necessary. Simulations were carried out at three state points with and without bias factors, and the results were identical. Thus for most of the Lennard-Jones results reported here, bias factors were not used. Sixty percent of moves were HMC moves used to thermally equilibrate the system, while 40% of the moves were used to change the value of the coupling strength between the fractional atom and the rest of the system. Each HMC move involved selection of velocities from a Maxwell−Boltzmann distribution followed by 5 microcanonical MD steps. Changes in λ were made uniformly up to a maximum change of $\Delta\lambda_{\max}$. The HMC time step and $\Delta\lambda_{\max}$ were adjusted during equilibration to achieve 50% move acceptance rates. Equilibration consisted of 2 million MC steps, after which averages were taken over 10 million MC steps. Cubic boxes with edge lengths of 5.87, 9, and 13.2 σ were used. Reduced units[42] were used to enable easy comparison with previous work and an equation of state. The reduced chemical potential was

$$\mu^* = \mu/\epsilon - k_{\mathrm{B}}T/\epsilon \ln(\Lambda^3/\sigma^3) \quad (21)$$

where Λ is the de Broglie wavelength. When computing the average composition, only the integer number of molecules was considered.

To enable comparison with the CFC-GCMC simulations, the GCMD method of Lo and Palmer was implemented and run for the Lennard-Jones system, as was the hybrid method of Boinepalli and Attard. In addition, traditional GCMC simulations were carried out.

4.2. Water. Traditional GCMC and CFC-GCMC simulations were conducted for a fully flexible model of water in which harmonic potentials were used for bond lengths and bond angles. A flexible model was required, since hybrid Monte Carlo requires the use of reversible integrators; traditional constraint dynamics methods are not reversible. The nominal bond lengths and angles as well as partial charges and Lennard-Jones parameters were taken from the SPC water model.[51] Because SPC is a rigid model, bond length and bond angle force constants were taken from the CHARMM22 force field[50] and set to 450 kcal/(mol Å ^{2}) and 55 kcal/(mol rad ^{2}), respectively.

As described below, simulations of this water model were also attempted with the method of Boinepalli and Attard, but convergence was difficult. To enable comparison between this method and the CFC approach, a "soft" water model was also simulated, in which all potential parameters were the same except the partial charges were removed from the oxygen and hydrogen atoms. By removing the partial charges, the soft water model does not hydrogen bond, and thus it is much easier to insert or delete molecules.

For the traditional GCMC simulations, 60% of the moves were HMC moves, 20% were insertion moves, and 20% were deletion moves. For CFC-GCMC, 60% of the moves were HMC moves, while 40% were random changes in λ. A typical GCMC simulation consisted of 1 million equilibration moves followed by 2 million moves for the production run.

The simulations for the "real" water model were conducted at 298 K and a fugacity of 3746.8 Pa in a cubic box with an edge length of 13.9623 Å. This corresponds to the conditions used previously by Mezei[52,53] in cavity bias GCMC simulations of a rigid water model. "Soft" water simulations were carried out at 300 K and fugacities ranging from 100 to 300 Pa. Electrostatics were treated using an Ewald sum with a real space cutoff of 6.5 Å and tin foil boundary conditions. A switching function of the following form was used for the Lennard-Jones interactions between all atoms in the system

$$\phi_{\text{switch}}(r) = \begin{cases} \phi_{ij}(r) & r < r_{\text{on}} \\ \phi_{ij}(r) \times \dfrac{(r_{\text{off}}^2 - r^2)^2(r_{\text{off}}^2 + 2r^2 - 3r_{\text{on}}^2)}{(r_{\text{off}}^2 - r_{\text{on}}^2)^3} & r_{\text{on}} \le r \le r_{\text{off}} \\ 0 & r > r_{\text{off}} \end{cases} \tag{22}$$

where $\phi_{ij}(r)$ denotes the full interaction between two atoms i and j with a distance of r, and r_{on} and r_{off} were set to be 6.0 Å and 6.5 Å, respectively. The switching function guarantees that the potential energy and force are continuous over the entire range of r. The scaled potential given in eq 20 was used for interactions between the fractional molecule and integer molecules. In addition, the partial charges on the fractional molecule were scaled as $q_f = \lambda^5 q_i$, where q_f and q_i are the partial charges on a fractional and integer atom, respectively. This nonlinear scaling was found to moderate strong electrostatic interactions that result when molecules were inserted close to an existing species. To avoid singularities, a hard core cutoff of 1.0 Å was used between fractional and integer atoms. If an insertion of a fractional molecule resulted in two atoms coming closer than 1.0 Å, then the insertion was rejected. This hard core cutoff was required for the SPC model because the hydrogen atoms are modeled with only a partial point charge. The lack of a van der Waals radius can cause singularities during insertions. If all atoms in the system have van der Waals radii, then the hard core cutoff will be unnecessary.

A reservoir of water molecules used for insertions was generated by simulating 500 ideal gas waters using HMC in parallel with the CFC-GCMC simulations. When an insertion is required, a configuration is randomly chosen from the reservoir, given a random orientation, and then randomly inserted into the system. After 100 MC steps, additional HMC moves are performed on the reservoir to ensure new conformations are generated such that the distribution in eq 12 is properly sampled. It should be noted that for the simple water model used here, the distribution in eq 12 could be sampled analytically with no need to generate a reservoir of configurations. This is not the case, however, for more complex solutes having intramolecular nonbonded interactions. It was verified that the reservoir method used here did indeed satisfy the analytic distribution for this model consistent with eq 12.

As with the Lennard-Jones case, 5 MD time steps were used for each HMC move, with the time step adjusted during equilibration to obtain roughly 50% acceptance rates. Random changes in λ were made, with the maximum change $\Delta\lambda_{\text{max}}$ adjusted during equilibration to achieve 50% acceptance of these moves. It was found that $\Delta\lambda_{\text{max}}$ was typically about 0.2, and the HMC time step was about 1.0 fs. During the production phase, success rates for λ changes averaged about 50%, while HMC moves were successful about 80% of the time.

As described below, it was found that bias factors were necessary to adequately simulate the real water system. Ideally, the bias factors should yield a flat distribution in λ values, and this criterion can be used to optimize the bias factors during the equilibration stage of the simulation. The simple updating method of Smith and Bruce[54] and the method of Wang and Landau[39] were both investigated. It was found that the Wang−Landau approach was superior. In this approach, the (0,1) range of λ is divided into ten adjacent bins of [0,0.1),[0.1,0.2),...,[0.9,1.0], and each bin j is assigned a bias factor η_j. Initially, all bias factors are set to $\eta_j = 0$. During the equilibration phase, each time a value of λ falls within the range of a bin j, the corresponding value of η_j is modified according to

$$\eta_j = \eta_j - \ln(\upsilon) \tag{23}$$

where υ acts as a scaling factor. Initially, $\ln(\upsilon)$ was set to unity. One hundred iterations were performed, each consisting of 10 000 steps. During this phase, histograms were collected for the number of times a given λ bin was visited. After each iteration, the distribution of λ was checked, and if the minimum probability of each λ bin was greater than 1%, then υ was modified according to the following expression

$$\upsilon = \sqrt{\upsilon} \tag{24}$$

The histograms were zeroed after each iteration.

4.3. CO_2 and Ethanol. To test the osmotic CFC MC method, the solubility of CO_2 in ethanol was computed. This is an ideal test system to study, since it was the basis of the 2004 Industrial Fluid Properties Simulation Challenge.[55] As a result, there are two sets of previous simulations that used the same force field against which comparison can be made. The first set of simulations was conducted by Errington and co-workers and utilized a transition matrix MC method.[56] The second simulation study was performed by Zhang and Siepmann and utilized Gibbs ensemble MC.[57] These two studies had consistent results, so agreement with these calculations is a strong indication of the validity of the CFC method.

Force field parameters were the same as that used in the two previous studies[56,57] with the exception that the bond lengths and angles for both CO_2 and ethanol were flexible in the CFC calculations but were held rigid in the other two works. Force constants were taken from the CHARMM22 force field.[50] The use of flexible models is expected to have a minimal impact of the phase equilibria properties. A flexible model was used in the CFC simulations because HMC was used for sampling, which requires the use of reversible and symplectic integration schemes. This is nontrivial to accomplish with constrained degrees of freedom, and so fully flexible models were used. A switching function, eq 22, was used for the Lennard-Jones interactions, with $r_{on} = 14.0$ Å and $r_{off} = 14.19$ Å. The Ewald sum method was used with a real space cutoff of 14.19 Å and $\kappa^{-1} = 0.2857$ Å^{-1}.

To begin the calculations, NPT simulations were run for 216 ethanol molecules at 323 K. Ninety-eight percent of the moves were HMC moves, each consisting of 5 MD time steps. The remaining 2% of the moves were volume change moves. One million equilibration moves followed by 2 million production moves were made. An HMC time step of 1.1 fs resulted in about 31% of the HMC moves being accepted, while a maximum volume change of 260 Å^3 resulted in 56% of the volume moves being accepted.

Once it was confirmed that the computed densities matched those in the literature, isotherms were calculated using osmotic CFC MC simulations at varying CO_2 fugacity and a fixed number of ethanol molecules, temperature, and pressure. In a typical application of the osmotic ensemble, the solvent is usually much less volatile than the solute, such that the total pressure of the system is equal to the partial pressure of the solute. The solute fugacity can be directly related to the pressure via $f_i = \phi_i P$ where ϕ_i is the solute fugacity coefficient. For the calculations here, CO_2 is certainly much more volatile than ethanol. Instead of performing calculations at varying pressure and fugacity, isotherms were computed at constant pressure and varying fugacity. In all cases, $P \geq f_i/\phi_i$, where ϕ_i was calculated for pure CO_2 at a given T and P from the Peng−Robinson equation of state. Since the simulations were of a binary single phase system, the phase rule was not violated. However, in a two phase vapor−liquid equilibrium experiment, a third sparingly soluble component (such as helium) would need to be added to the vapor phase to independently vary pressure and CO_2 fugacity. Not surprisingly, the calculations show that the solubility of CO_2 is a very weak function of total pressure, and so isotherms computed in this manner can be compared with experiment with little error.

The attempt probabilities for translation, volume change, and λ change moves were 60%, 2%, and 38%, respectively. The Wang−Landau scheme was used to optimize the bias factors, with anywhere from 20 to 100 iterations used to obtain an acceptably flat distribution of λ values. As an example, for a simulation at 323 K and a total pressure of 20 bar with a fugacity of 10 bar, the weighting factors for each of the ten λ bins were 0.0, 0.511, 1.47, 1.784, 1.613, 1.289, 0.845, 0.397, −0.232, and −1.194, respectively. The resulting λ distribution was quite flat; the ratio of the most probable to least probable λ bin was roughly 1.6.

5. Results

5.1. Lennard-Jones. A series of CFC GCMC simulations was conducted for the Lennard-Jones fluid at T^* values ranging from 0.769 to 1.25 and ρ^* values ranging from 0.005 up to 0.886. The results were compared against the accurate equation of state of Johnson et al.,[58] standard GCMC simulations carried out as part of this work, and the simulation results reported by Lo and Palmer.[29] We also implemented the methods reported by Lo and Palmer[29] and Boinepalli and Attard[35] and conducted simulations with these algorithms. This was done to compare the efficiency and accuracy of these methods against the CFC method. The only difference between our implementation of the Boinepalli and Attard method and theirs is that they used stochastic temperature control, while we used HMC. This should not alter the results.

Figure 1 shows the isotherm calculated at $T^* = 1.0$ and is typical of the type of agreement obtained between the different methods. A complete listing of all the results is provided in Table 1. With the exception of our implementation of Lo and Palmer's method, all the results agree with each other and the equation of state within numerical accuracy. The results reported by Lo and Palmer agree with our calculations using their method at low densities, but at higher densities our implementation of their method systematically overpredicts the densities. It is possible there is a mistake in our implementation of the Lo and Palmer method, although it is unclear why our implementation agrees at low density but not high density. We are unaware of any other publications in which the Lo and Palmer method has been implemented and tested. It would be interesting to see if other groups can reproduce the original results.

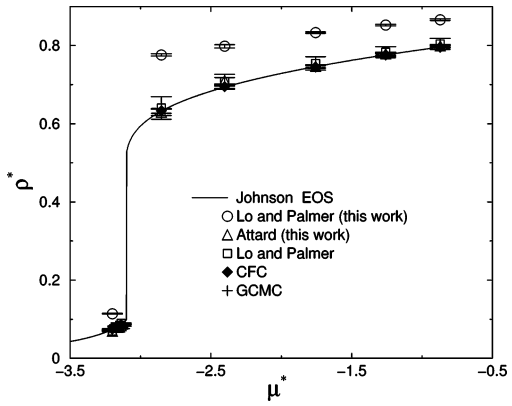

Figure 1. Computed isotherms for the Lennard-Jones fluid at $T^* = 1.0$ using various methods. The line is from the equation of state of Johnson et al.[58] The circles are from this work using the method of Lo and Palmer.[29] The triangles are from this work using the method of Boinepalli and Attard.[35] The squares are the results reported by Lo and Palmer.[29] The filled diamonds are from CFC GCMC simulations, while pluses are from standard GCMC simulations. All the numerical values are also shown in Table 1. With the exception of our implementation of Lo and Palmer's method, all results agree with each other and the equation of state.

In any case, these results indicate that the CFC algorithm is correct and has been properly implemented in our locally developed software package. For the Lennard-Jones system at these state points, there is no advantage to using the CFC GCMC method (or the other slow growth open system methods). Standard GCMC is sufficient for computing volumetric properties of the Lennard-Jones fluid under these conditions. At lower temperatures and higher densities, however, there may be a need to use more sophisticated techniques. These state points were not explored, because the goal was to test the method under conditions where standard techniques give reliable results to compare against.

It is interesting to investigate the role the bias potentials η_j play on the results and the performance of the method. As mentioned earlier, some test simulations were run with and without the use of the extended system bias potentials. As shown in Table 2, the final result for the Lennard-Jones fluid under these conditions is independent of whether or not the bias potentials are used. As shown below, however, the use of bias potentials for systems more complex than Lennard-Jones leads to more efficient sampling. The CFC results agree with the GCMC results regardless of whether or not a bias potential is used. We note one exception. At the lowest temperature, the GCMC and the CFC results are very close but fall just outside of the error bars. It may be that the uncertainty of these calculations has been underestimated. Figure 2 shows the distribution of λ with and without bias potentials for the Lennard-Jones system at $T^* = 0.769$, $L/\sigma = 5.87$, and $\mu^* = -0.816$. Although ultimately both calculations yield the same density, the simulations without the bias potentials have very few intermediate values of λ. The fractional molecule is almost always at states where $\lambda < 0.3$ and $\lambda > 0.6$. With the use of bias potentials, however, the fractional molecule spends roughly the same amount of time at all values of λ. The latter situation leads to much more reliable sampling.

As an aside, it turns out that the pseudopotential in the Boinepalli and Attard method can be reduced to a specific value of the CFC bias factor for the Lennard-Jones system via the following relation

$$\exp(\eta) = \frac{\exp[\beta\mu\lambda]V^{\lambda-1}N!}{\Lambda^{3(\lambda-1)}(N+\lambda)!} \tag{25}$$

where N is the number of integer molecules present in the system. This means that the Boinepalli and Attard method and the CFC method are equivalent for the Lennard-Jones system if the bias factors are given by eq 25. The advantage of the CFC approach is that the bias factors can be adjusted in a self-adaptive way to optimize the performance of the algorithm.

5.2. Water. Water provides a more stringent test of the CFC method. For this system, hydrogen bonding and orientational order make random insertions and deletions of molecules more difficult. To begin these simulations, the "soft" water model was initially simulated at low densities. Under these conditions, the fluid should exhibit ideal gas behavior, and so algorithms can easily be checked to ensure they capture this limiting behavior. In addition to CFC GCMC calculations, standard GCMC simulations were also carried out on this system, as were simulations using our implementation of the Boinepalli and Attard method. The simulations were conducted at 300 K in a cubic box with edge lengths of 1000 Å and fugacities ranging from 100 to 300 Pa. The results are shown in Figure 3. As can be seen, all results agree. As with the case of the Lennard-Jones simulations, the use of bias factors improves sampling somewhat but is unnecessary to get accurate results. That is, simulations were performed with and without bias factors, and the results were identical. This is not surprising, given the low density and lack of hydrogen bonding in this system.

Next, simulations were carried out for the flexible SPC model of water using our implementation of Boinepalli and Attard's method and CFC GCMC. Mezei[52,53] has conducted cavity-bias GCMC simulations on a rigid SPC water model using a cubic box with edge lengths of 13.9623 Å at $T =$ 298 K and $B = -6.0$, where

$$B = \ln(\beta f\ V) \tag{26}$$

This corresponds to a fugacity of 3746.9 Pa. The number of water molecules at this condition was found to be 90.30 $\pm$ 0.9,[52,53] which corresponds to a water density of 0.992 $\pm$ 0.0099 g/cm^3.

Using our implementation of Boinepalli and Attard's method, calculations were started at the same temperature and box volume used by Mezei, but with only 60 water molecules initially in the box. This corresponds to an initial density of 0.659 g/cm^3. After 3 million steps, it was observed that the number of water molecules in the system had not changed; the fractional molecule became "stuck" at an intermediate value. Figure 4 shows the composition and probability distribution of λ. It can be seen that λ lies mostly in the range of 0.3−0.5 and never exceeds 0.65 or goes lower

Table 1. Results from Simulations of the Lennard-Jones Fluid[a]

T^*	μ^*	ρ^* EOS	ρ^* GCMC	ρ^* LP	ρ^*LP (pw)	ρ^* BA (pw)	ρ^* CFC
1.0	−3.2	0.0744	0.0729 (2)	0.0699 (6)	0.114 (1)	0.0698 (9)	0.075 (1)
1.0	−3.162	0.0817	0.0807 (4)	0.0808 (8)	–	0.079 (2)	0.0807 (8)
1.0	−3.150	0.0844	0.0823 (5)	0.0844 (18)	–	0.080 (2)	0.083 (1)
1.0	−3.139	0.0871	0.0857 (5)	0.0896 (10)	–	0.084 (1)	0.086 (1)
1.0	−2.852	0.632	0.638 (1)	0.640 (29)	0.776 (3)	0.629 (4)	0.633 (6)
1.0	−2.403	0.693	0.697 (1)	0.707 (19)	0.798 (4)	0.710 (4)	0.695 (5)
1.0	−1.757	0.746	0.750 (1)	0.754 (17)	0.833 (2)	0.747 (2)	0.745 (3)
1.0	−1.258	0.776	0.780 (1)	0.783 (14)	0.852 (2)	0.779 (2)	0.776 (4)
1.0	−0.87	0.796	0.799 (1)	0.804 (14)	0.866 (2)	0.799 (2)	0.796 (2)
0.769	−4.127	0.00501	0.00510 (1)	0.0050 (0)	0.00489 (9)	0.0050 (1)	0.0048 (2)
0.769	−3.797	0.00802	0.00800 (1)	0.0081 (0)	0.0078 (2)	0.0082 (2)	0.00801 (6)
0.769	−3.646	0.01004	0.01003 (2)	0.0099 (0)	0.0097 (3)	0.0100 (3)	0.00996 (8)
0.769	−3.583	0.01105	0.01101 (2)	0.0108 (0)	0.0108 (3)	0.0116 (3)	0.01100 (8)
0.769	−2.999	0.779	0.784 (1)	0.785 (20)	0.864 (3)	0.789 (3)	0.787 (9)
0.769	−2.868	0.788	0.794 (1)	0.794 (11)	0.871 (3)	0.799 (3)	0.793 (8)
0.769	−2.727	0.798	0.803 (1)	0.804 (11)	0.879 (3)	0.801 (3)	0.801 (4)
0.769	−0.816	0.886	0.890 (1)	0.895 (21)	0.934 (2)	0.889 (2)	0.886 (4)
1.1	−3.35	0.0815	0.0810 (3)	–	–	0.077 (2)	0.0803 (7)
1.15	−3.3	0.106	0.104 (1)	–	–	0.109 (4)	0.104 (1)
1.2	−3.2	0.151	0.146 (1)	–	–	0.148 (4)	0.142 (4)
1.25	−3.14	0.189	0.184 (1)	–	–	0.177 (6)	0.185 (2)
1.25	−3.03	0.244	0.241 (2)	–	–	0.249 (2)	0.242 (4)
1.25	−2.95	0.291	0.301 (3)	–	–	0.293 (4)	0.302 (5)
1.25	−2.9	0.321	0.343 (3)	–	–	–	0.332 (8)
1.25	−2.8	0.381	0.406 (4)	–	–	–	0.410 (5)
1.25	−2.76	0.403	0.432 (2)	–	–	–	0.431 (8)
1.25	−2.73	0.419	0.443 (1)	–	–	–	0.446 (3)
1.175	−2.68	0.511	0.529 (1)	–	–	0.526 (2)	0.523 (6)
1.15	−2.63	0.552	0.562 (2)	–	–	0.552 (6)	0.560 (6)
1.1	−2.55	0.610	0.616 (2)	–	–	0.611 (3)	0.619 (3)

[a] Abbreviations are as follows. EOS: Equation of State of Johnson et al.;[58] GCMC: Grand Canonical Monte Carlo Simulation, present study; LP: results of Lo and Palmer;[29] LP (pw): results of the present work using the Lo and Palmer algorithm; BA (pw): results of the present work using the Boinepalli and Attard[35] algorithm; CFC: Continuous Fractional Component GCMC Calculations. A "–" indicates no results available.

Table 2. Simulation Results for the Lennard-Jones Fluids Using Standard GCMC Simulations as Well as CFC GCMC with and without Extended System Bias Potentials[a]

T^*	L/σ	μ^*	ρ^* (GCMC)	ρ^* (CFC, no η)	ρ^* (CFC, with η)	η
0.769	5.87	−0.816	0.890 (1)	0.883 (3)	0.886 (4)	1.8, 3.5, 5.6, 5.9, 5.5, 4.8, 3.8, 2.7, 1.4, 0.0
1.0	9	−3.52	0.0416 (1)	0.0417 (2)	0.0418 (2)	0.0, 0.1, 0, 0, −0.3, −0.5, −0.8, −1.1, −1.5, −1.7
1.25	9	−3.03	0.241 (2)	0.242 (4)	0.242 (7)	0.7, 0.9, 0.8, 0.7, 0.5, 0.4, 0.2, 0.0, 0.8, 0.6

[a] Numbers in parentheses are the standard deviations in the last digit.

than 0.1. Apparently, under the conditions investigated here, the pseudopotential used for insertions and deletions in this approach is incapable of "pushing" the fractional molecule out of the intermediate λ range. This does not indicate that the Boinepalli and Attard method is flawed; as they point out one can add an additional bias potential to the method to improve sampling. What it does indicate is that the addition or deletion of water molecules is difficult with these types of methods.

Standard GCMC as well as CFC GCMC calculations were then run at this state point, with 60 and 90 water molecules initially placed in the simulation box. This corresponds to initial densities of 0.659 g/cm^3 and 0.989 g/cm^3. The convergence of these simulations is shown in Figure 5. Regardless of the initial density, both methods converge to the same density of 1.03 ± 0.01 g/cm^3, again indicating that the CFC method gives equivalent results to GCMC. These densities are slightly higher than that reported by Mezei. The difference is probably due to the fact that an Ewald sum was used for the present calculations, while a charge cutoff was used by Mezei. In addition, Mezei used a rigid model for water, while the water model used here had flexible bond lengths and angles.

Although the CFC and standard GCMC simulations yield identical densities, the sampling efficiency of the CFC method is much better than standard GCMC. At the final water density, less than 0.01% of water insertion or deletion moves were accepted for the standard GCMC simulations. This means the composition changes once on average every 10 000 insertion or deletion attempts. With the CFC method, however, 2.4% of the moves involving a change in λ resulted in the creation or removal of an integer molecule, and 52%

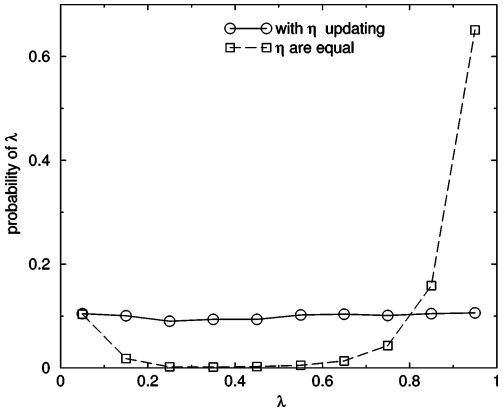

Figure 2. The probability distribution of λ during CFC GCMC simulations of the Lennard-Jones system with and without biasing potentials. The simulations are performed on a system at $T = 0.769$, $L/\sigma = 5.87$, and $\mu^* = -0.816$. The lines are to guide the eyes. Circles are for calculations using bias potentials, and squares are for unbiased simulations. The use of bias potentials promotes a more uniform λ distribution, while the simulations without bias potentials tend to become "stuck", especially near $\lambda = 1$.

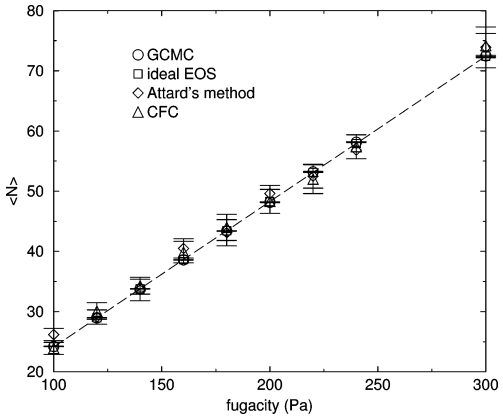

Figure 3. The average number of integer molecules as a function of fugacity for a "soft" water model at 300 K and very low density. The circles are from standard GCMC simulations, the squares are for the ideal gas, the diamonds are results from our implementation of Boinepalli and Attard's method[35] extended to molecules, and the triangles are from the CFC GCMC calculations. The line is a guide for the eye.

of all attempted changes in λ were accepted. This means the composition changes once on average every 42 attempts to change λ. Figure 6 shows the instantaneous values of λ over a portion of the water simulation. Clearly, λ traverses the entire range from fully coupled to ideal gas many times during the simulation. Figure 7 shows the initial convergence characteristics of the CFC simulation starting from 90 water molecules (left) and the distribution of λ values during the final 2 million moves of the production run (right). While the λ distribution is peaked near the ends, all values are visited frequently enough to maintain good transition frequencies. Better updating schemes for the bias potentials would likely improve the distribution of λ even more.

Mezei also used cavity biased GCMC to calculate the density of the rigid SPC water model at $T = 298$ K and $B = -5.7$, which corresponds to $f = 5057.7$ Pa. At these conditions, the computed density was 1.006 ± 0.010 g/cm^3, which is 1.4% higher than the density computed at $f = 3746.9$ Pa. Using CFC GCMC, the density was found to be 1.046 ± 0.007 g/cm^3 or 1.5% higher than that calculated at the lower fugacity. Again, there is a slight difference in the absolute values, but the relative increase in density is nearly the same. Interestingly, standard unbiased GCMC becomes extremely inefficient at this state point; we were unable to obtain reliable statistics due to the vanishingly low probabilities of successful insertions and deletions.

Obviously, one can develop powerful biasing schemes that will improve the insertion and deletion success rates of GCMC, and Mezei's cavity biasing approach is one such method. The CFC method can be thought of as another biased MC move, but one that is self-adapting in terms of the use of gradual insertions and deletions as well as bias factors that promote these gradual transitions. It is thus

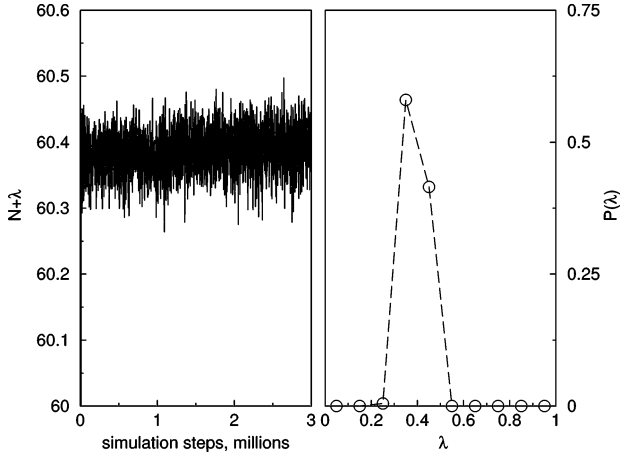

Figure 4. The real number of $N + \lambda$ fluctuation and the probability of λ for the flexible SPC water model from the method by Boinepalli and Attard.[35] The dashed line is to guide the eye.

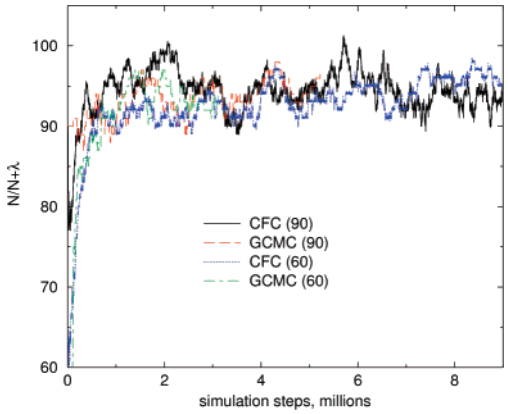

Figure 5. The real number of molecules $N + \lambda$ from the CFC method and the integer number of molecules N from GCMC simulation for the flexible SPC water model. The solid black and long-dash red lines correspond to the CFC and standard GCMC simulations starting from 90 water molecules, respectively. The dotted blue and dot-dash green lines are for the CFC and GCMC simulations starting from 60 water molecules, respectively.

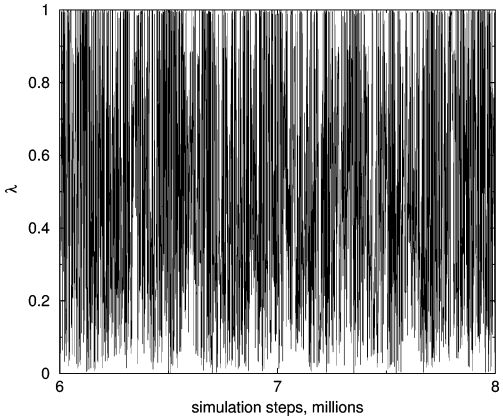

Figure 6. Instantaneous values of λ during the simulation of SPC water. For clarity, only the last 2 million steps of the simulation are shown. λ clearly traverses the entire range of values during the simulation.

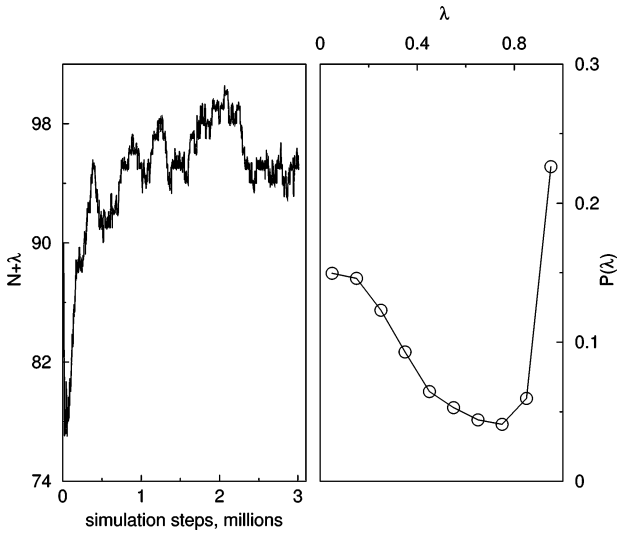

Figure 7. Left: The number of water molecules in the simulation box (integer plus fractional) as a function of simulation duration. Right: The average probability of a given value of λ, using the updating scheme of Wang and Landau[39] to modify the importance weighting factors η. During the last 2 million steps, a total of 798 335 attempts are made to change λ, out of which just over 51% are accepted. Out of these changes in λ, a molecule was successfully removed from the system 10 727 times, and a molecule was added to the system a total of 10 728 times.

relatively straightforward to implement and requires little a priori tuning. We also note that, as expected, the equilibrium properties do not depend on the exact values of the weighting factors used. Two different sets of weighting factors obtained from different updating schemes yielded the same density for the SPC water system. Obviously, very poor choices for weighting factors will have a negative effect on the calculations, but it appears that the method is robust enough such that a fairly wide range of weighting factors will achieve favorable results.

5.3. CO_2 and Ethanol. CFC osmotic simulations were performed to compute the solubility of CO_2 in ethanol. Isothermal−isobaric simulations were first performed on the pure ethanol system at 323 K and 0.294 bar to compare

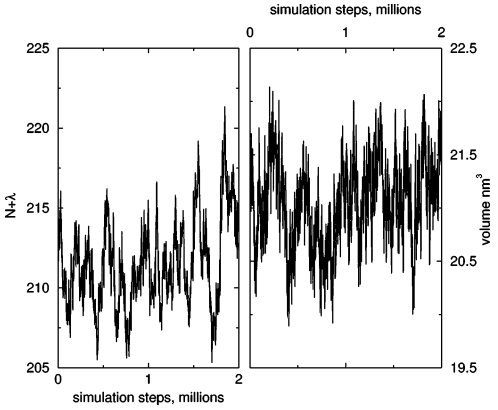

Figure 8. Left: The total number of CO_2 molecules (integer plus fractional) in ethanol at $T = 323$ K, $P = 40$ bar, and $f =$ 10 bar. Right: The volume fluctuation during the simulation.

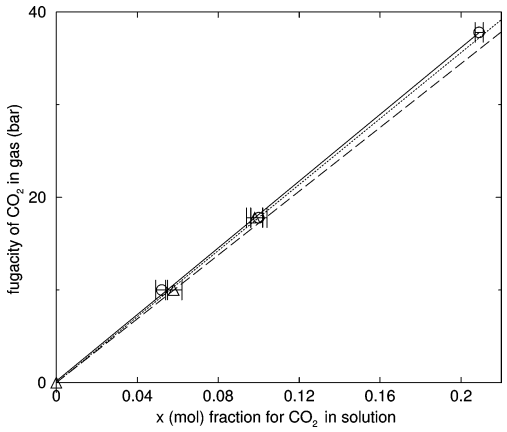

Figure 9. The fugacity of CO_2 in the gas phase versus the mole fraction of CO_2 in the solution phase at 323 K. The three circles are from CFC osmotic simulations at a total pressure of 40 bar. The triangles are from simulations at a total pressure of 20 bar. The solid line is a linear fit to the 40 bar simulation results and yields an estimated Henry's constant of 18.0 ± 0.3 MPa. A linear fit to the 20 bar results yields a Henry's constant of 18.1 ± 0.6 MPa. The dotted line represents a Henry's constant of 17.8 ± 0.6 MPa calculated by Errington et al.,[56] while the long dashed line represents a Henry's constant of 17.2 ± 0.4 MPa obtained by Zhang and Siepmann.[57]

against the previous results of Errington and co-workers.[56] The pure ethanol density at these conditions was calculated as 0.751 ± 0.002 g/cm^3. This is consistent with the value reported by Errington and co-workers of 0.7579 ± 0.0005 g/cm^3. The small difference may be attributed to the fact that a flexible model for ethanol was used here, while Errington and co-workers used a rigid model. Also, the way in which the Lennard-Jones potential was truncated differs slightly between the two methods.

CFC osmotic simulations were then run at 323 K and varying CO_2 fugacities. The total pressure of the system was set to either 20 bar or 40 bar. An example of the convergence of the number of CO_2 solutes and system volume for a fugacity of 10 bar is shown in Figure 8. The isotherms calculated from these simulations are shown in Figure 9. Note that the solubilities calculated at total pressures of 20 and 40 bar agree within the uncertainty of the calculations,

confirming the insensitivity of the solubility to total pressure. By fitting a straight line through the data, a Henry's law constant may be obtained. For the CFC calculations the Henry's law constant was found to be 18.0 ± 0.3 MPa at 40 bar and 18.1 ± 0.6 MPa at 20 bar, respectively. This agrees very well with the results obtained from transition matrix MC calculations (17.8 ± 0.6 MPa) and Gibbs ensemble calculations (17.2 ± 0.4 MPa). The Henry's constants for these two previous calculations are shown as dotted and dashed lines in Figure 9. We note that the two previous studies[56,57] have used low pressure and concentration of CO_2. The fact that the osmotic CFC calculations agree with these previous calculations confirms that the method yields correct results.

6. Conclusions

A new open system method called continuous fractional component Monte Carlo (CFC MC) has been presented. The method is designed to overcome difficulties with the insertion and deletion of molecules and relies upon gradual changes in a coupling parameter to increase or decrease the interaction strength between a fractional molecule and the rest of the molecules in the system. An adaptive bias potential is used to enable these transitions to occur efficiently. The bias potential is adjusted using the Wang−Landau updating scheme in such a way as to maintain a uniform distribution of fractional molecule states. In between these coupling strength moves, hybrid Monte Carlo steps are used to relax the system.

Acceptance rules were developed for the grand canonical and osmotic ensembles. The grand canonical CFC method was used to calculate the volumetric properties of the Lennard-Jones fluid and SPC water. The results agreed with accepted values and a standard implementation of grand canonical Monte Carlo. An implementation of another hybrid method[35] was also found to give the correct results for the Lennard-Jones fluid but failed for the SPC water case, due to an inability to make successful changes in the coupling parameter. Without the use of a bias potential, the system became "stuck" at intermediate states. It was not possible to achieve correct results for the Lennard-Jones fluid at high density with an implementation of a previously developed[29] grand canonical molecular dynamics procedure. The osmotic CFC method was used to compute the solubility of CO_2 in ethanol. Results from the calculations agree with two previously published simulation studies.

The CFC method has been verified by comparing calculated volumetric properties and solubilities against previous accepted results. In the future we wish to extend the CFC method to other ensembles and apply it to more challenging systems where other simulation approaches fail or become very inefficient.

Acknowledgment. Support for this work was provided by the Air Force Office of Scientific Research under contract number F49620-03-1-0212.

References

(1) Potoff, J. J.; Siepmann, J. I. *AIChE J.* **2001**, *47*, 1676.

(2) Gao, G. T.; Wang, W. C. *Fluid Phase Equil.* **1997**, *130*, 157.

(3) Martin, M. G. *Fluid Phase Equil.* **2006**, *248*, 50.

(4) Crozier, P. S.; Rowley, R. L. *Fluid Phase Equil.* **2002**, *193*, 53.

(5) Morrow, T. I.; Maginn, E. J. *J. Chem. Phys.* **2005**, *122*, 54504.

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

(7) Agrawal, P. M.; Rice, B. M.; Thompson, D. L. *J. Chem. Phys.* **2003**, *119*, 9617.

(8) Eike, D. M.; Maginn, E. J. *J. Chem. Phys.* **2006**, *124*, 164503.

(9) Chen, B.; Siepmann, J. I.; Klein, M. K. *J. Phys. Chem. B* **2001**, *105*, 9840.

(10) Zhao, X. S.; Chen, B.; Karaborni, S.; Siepmann, J. I. *J. Phys. Chem. B* **2005**, *109*, 5368.

(11) Kofke, D. A.; Cummings, P. T. *Fluid Phase Equil.* **1998**, *150*, 41.

(12) Kofke, D. A. *Fluid Phase Equil.* **2005**, *228*, 41.

(13) Snurr, R. Q.; Bell, A. T.; Theodorou, D. N. *J. Phys. Chem.* **1993**, *97*, 13742.

(14) Mezei, M. *Mol. Phys.* **1980**, *40*, 901.

(15) Siepmann, J. I.; Frenkel, D. *Mol. Phys.* **1992**, *75*, 59−70.

(16) Mooij, G. C. A. M.; Frenkel, D.; Smit, B. *J. Phys.: Condens. Matter* **1992**, *4*, L255−L259.

(17) Laso, M.; de Pablo, J. J.; Suter, U. W. *J. Chem. Phys.* **1992**, *97*, 2817−2819.

(18) Berg, B. A.; Neuhaus, T. *Phys. Rev. Lett.* **1992**, *68*, 9.

(19) Orkoulas, G.; Panagiotopoulos, A. Z. *J. Chem. Phys.* **1999**, *110*, 1581.

(20) Yan, Q. L.; de Pablo, J. J. *J. Chem. Phys.* **1999**, *111*, 9509.

(21) Fenwick, M. K.; Escobedo, F. A. *J. Chem. Phys.* **2003**, *119*, 11998.

(22) De Pablo, J. J.; Prausnitz, J. M. *Fluid Phase Equil.* **1989**, *53*, 177.

(23) Martin, M. G.; Siepmann, J. I. *J. Am. Chem. Soc.* **1997**, *119*, 8921.

(24) Chen, B.; Siepmann, J. I.; Klein, M. L. *J. Am. Chem. Soc.* **2002**, *124*, 12232.

(25) Wick, C. D.; Siepmann, J. I.; Theodorou, D. N. *J. Am. Chem. Soc.* **2005**, *127*, 12338.

(26) Çağin, T.; Pettitt, B. M. *Mol. Phys.* **1991**, *72*, 169−175.

(27) Çağin, T.; Pettitt, B. M. *Mol. Simul.* **1991**, *6*, 5.

(28) Ji, J.; Çağin, T.; Pettitt, B. M. *J. Chem. Phys.* **1992**, *96*, 1333−1342.

(29) Lo, C. M.; Palmer, B. *J. Chem. Phys.* **1995**, *102*, 925−931.

(30) Shroll, R. M.; Smith, D. E. *J. Chem. Phys.* **1999**, *111*, 9025−9033.

(31) Lísal, M.; Smith, W. R.; Kolafa, J. *J. Phys. Chem. B* **2005**, *109*, 12956−12965.

(32) Theodorou, D. N. *J. Chem. Phys.* **2006**, *124*, 34109.

(33) Uhlherr, A.; Theodorou, D. N. *J. Chem. Phys.* **2006**, *125*, 84107.

(34) Escobedo, F. A.; de Pablo, J. J. *J. Chem. Phys.* **1996**, *105*, 4391−4394.

(35) Boinepalli, S.; Attard, P. *J. Chem. Phys.* **2003**, *119*, 12769−12775.

(36) Lynch, G. C.; Pettitt, B. M. *J. Chem. Phys.* **1997**, *107*, 8594−8610.

(37) Lyubartsev, A. P.; Martsinovski, A. A.; Shevkunov, S. V.; Vorontsov-Velyaminov, P. N. *J. Chem. Phys.* **1992**, *96*, 1776−17837.

(38) Rosso, L.; Minary, P.; Zhu, Z.; Tuckerman, M. E. *J. Chem. Phys.* **2002**, *116*, 4389.

(39) Wang, F.; Landau, D. P. *Phys. Rev. Lett.* **2001**, *86*, 2050−2053.

(40) Macedonia, M. D.; Maginn, E. J. *Mol. Phys.* **1999**, *96*, 1375−1390.

(41) June, R. L.; Bell, A. T.; Theodorou, D. N. *J. Phys. Chem.* **1990**, *94*, 8232.

(42) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids;* Clarendon: Oxford, 1987.

(43) Frenkel, D.; Smit, B. *Understanding Molecular Simulation;* Academic Press: San Diego, 2002.

(44) Duane, S.; Kennedy, A. D.; Pendleton, B. J.; Roweth, D. *Phys. Lett. B* **1987**, *195*, 216−222.

(45) Errington, J. R.; Panagiotopoulos, A. Z. *J. Chem. Phys.* **1999**, *111*, 9731.

(46) Spyriouni, T.; Economou, I. G.; Theodorou, D. N. *Phys. Rev. Lett.* **1998**, *80*, 4466−4469.

(47) Banaszak, B. J.; Faller, R.; de Pablo, J. J. *J. Chem. Phys.* **2004**, *120*, 11304.

(48) Zervopoulou, E.; Mavrantzas, V. G.; Theodorou, D. N. *J. Chem. Phys.* **2001**, *115*, 2860.

(49) Beutler, T. C.; Mark, A. E.; van Schaik, R. C.; Gerber, P. R.; van Gunsteren, W. F. *Chem. Phys. Lett.* **1994**, *222*, 529−539.

(50) MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L., Jr.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E., III; 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.

(51) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. In *Intermolecular Forces*; Dordrecht: The Netherlands, 1981.

(52) Mezei, M. *Mol. Phys.* **1987**, *61*, 565−582.

(53) Mezei, M. *Mol. Phys.* **1989**, *67*, 1207−1208.

(54) Smith, G. R.; Bruce, A. D. *J. Phys. A* **1995**, *28*, 6623−6643.

(55) Industrial Fluid Properties Simulation Challenge. http://fluidproperties.org (accessed January 4, 2007).

(56) Cichowski, E. C.; Schmidt, T. R.; Errington, J. R. *Fluid Phase Equil.* **2005**, *236*, 58.

(57) Zhang, L.; Siepmann, J. I. *Theor. Chem. Acc.* **2006**, *115*, 391.

(58) Johnson, J. K.; Zollweg, J. A.; Gubbins, K. E. *Mol. Phys.* **1993**, *78*, 591−618.

CT7000039