Molecular dynamics simulations in the grand canonical ensemble: Formulation of a bias potential for umbrella sampling

Robert M. Shroll and David E. Smith
Department of Chemistry and Biochemistry, MSC 3C, New Mexico State University, P.O. Box 30001, Las Cruces, New Mexico 88003-8001

(Received 18 December 1998; accepted 3 February 1999)

An extended Hamiltonian technique for performing grand canonical ensemble molecular dynamics simulations has been reformulated to include umbrella sampling, thus improving the efficiency of particle creation and annihilation processes. This was accomplished through incorporation of a bias potential in the Hamiltonian that modifies the free energy contour between integer particle number states. The extended Hamiltonian includes a continuous particle number variable that is the sum of the integer particle number and a coupling parameter, the latter being used in a scaling function for a fractional molecule’s interactions with the rest of the system. Equations of motion for the coupling parameter, derived from the Hamiltonian, integrate to yield density fluctuations at constant chemical potential. This new method may be adapted to a wide range of potential scaling functions. The technique was applied to calculations of extended simple point charge water density versus chemical potential, using both linear and nonlinear scaling. For each scaling function, a bias potential was constructed using a thermodynamic integration technique. Grand canonical ensemble simulations then yielded results in agreement with independent calculations. © 1999 American Institute of Physics. [S0021-9606(99)50517-5]

I. INTRODUCTION

In open chemical systems, matter exchange between the system and its surroundings is governed by the chemical potential. It is often desirable to know how the composition of such a system changes as a function of chemical potential and other constrained thermodynamic parameters. In an adsorption isotherm, for example, the amount of a substance adsorbed at constant temperature is determined for various values of its chemical potential, measured in terms of vapor pressure. Analysis of the isotherms reveals structural and surface binding properties of the substrate. Computer simulations may also be used to investigate particle density and density fluctuations at constant chemical potential. This may be accomplished through simulation of a system in direct contact with a particle bath. However, this approach is inefficient, with a great deal of computational effort wasted on the particle bath, the system–bath interface, and lengthy equilibration times. These deficiencies may be addressed by sampling from the grand canonical ensemble where chemical potential is one of the constrained thermodynamic variables.

Several alternative open-system simulation techniques have been developed recently. The first grand canonical ensemble simulation methods were based upon a Monte Carlo (MC) approach.1–3 MC techniques are well suited theoretically for discrete number variations, with particle insertion and deletion criteria based upon the Widom test particle relationship.4–5 Nevertheless, simple MC methods can be very inefficient when applied to condensed phases. They rely upon local density fluctuations to produce configurations for which particle insertion or deletion is energetically favorable. For molecular liquids, such as water, these fluctuations are rare. Methods development research, mostly based upon cavity-bias schemes,6 has recently provided significantly improved sampling efficiency.7–13

As with all MC-based approaches, grand canonical MC methods do not directly yield dynamical information about the system. This provides motivation for the development of molecular dynamics-based, open-system methods. Typically, the equations of motion used in molecular dynamics (MD) techniques are poorly suited to simulate discrete processes such as particle number variations.1 This has led to the development of MD methods that sample at constant chemical potential through the use of MC-based stochastic procedures for particle insertions and deletions.14–19 While successful at generating dynamical information, no proof yet exists that these techniques produce grand canonical ensemble averages. In addition, they suffer from the same particle insertion and deletion difficulties as the MC methods.

An alternative, MD approach to open-system simulations, first developed by Çağin and Pettitt,20 involves representing the particle number as a continuous “extension” variable that is treated dynamically. The approach is based upon extended-system methods1,2 in which a Hamiltonian, containing potential and kinetic energy terms for the extension variables, is postulated and shown to yield correct averages for the desired ensemble. Equations of motion for the extension variables are derived in the usual fashion. This technique was initially applied to Lennard-Jones particles,20 and subsequently shown to be applicable to more complicated systems.21–26 The extended-system approach has the advantage of being theoretically tractable, yielding correct ensemble averages. In addition, it is possible that this dynamical approach may circumvent some of the computational difficulties associated with discrete insertion and dele-
tion methods. In particular, the system is given the opportunity to equilibrate in the presence of fractional particles, responding to either accept or reject them in a gradual fashion.

Here, the grand canonical molecular dynamics (GMD) methodology is reformulated to include an umbrella sampling procedure. This is accomplished through the addition of a bias potential that modifies the free energy along the coupling-parameter path. For the extended simple point charge (SPC/E) water model, the free energy barrier between integer particle-number states was calculated to be much higher than is typically accessible via thermal excitation. The bias potential was used to reduce the magnitude of this barrier, in a fashion analogous to that of a catalyst, improving the efficiency of particle-number fluctuations without changing the probability distribution of the integer particle-number states. Grand canonical ensemble averages were thereby maintained without needing to subtract effects of the bias potential.

In this paper, an extended-system Hamiltonian and its corresponding equations of motion are presented. The Hamiltonian contains a general coupling-parameter path, connecting integer particle-number states, and a term for the bias potential. The extended-system, microcanonical partition function is related to the grand canonical partition function, and necessary restrictions on the bias potential are thereby defined. Equations for the free energy along the coupling parameter path are derived, and a thermodynamic integration (TI) procedure for constructing an optimal bias potential is presented. A particle insertion and deletion procedure is also defined and discussed.

The bias potential GMD method was applied to a calculation of the average water density as a function of chemical potential for the SPC/E water model. Both linear and nonlinear single-particle scaling functions were investigated. Linear scaling is problematic near the uncoupled particle state due to overlapping of particles on the highly repulsive energy surface. The resulting numerical instabilities associated with unbounded coupling-parameter accelerations were avoided through a modification of the interaction potential at very short separations. The need for biased particle insertions was therefore eliminated, in a manner similar to that used by Lo and Palmer. Interaction potential modifications were not required for the nonlinear path. An optimal bias potential was calculated for each path using the TI method. Results of the GMD simulations are compared with excess chemical potentials measured in the TI calculations and with values reported in the literature.

II. METHODOLOGY

A. Grand canonical ensemble Hamiltonian

The theoretical basis for sampling the grand canonical ensemble using MD techniques has been given in detail elsewhere. The following provides a necessary justification of the bias potential method and is included for clarity.

The intent of grand canonical ensemble MD simulations is to use classical equations of motion to investigate systems in which the temperature, volume, and chemical potential are fixed and the number of particles fluctuates. This is accomplished by adding extension variables to the Hamiltonian that scale particle velocity and couple a fractional particle to the system via a continuous function of a coupling parameter, \( \lambda \). A continuous particle number variable \( \nu \) may be defined as

\[
\nu = N + \lambda. \tag{1}
\]

At any given time, \( N \) is the number of “whole” particles and \( \lambda \), which varies between zero and one, represents the extent to which a single, “fractional” particle is coupled to the rest of the system. An extension variable, \( s \), is used to scale the particle momenta. A Hamiltonian for the grand canonical ensemble may be written as

\[
\mathcal{H} = \sum_{i=1}^{N} \frac{p_i^2}{2m_i} + \frac{p_f^2}{2m_f} + \frac{p_s^2}{2Q} + \frac{p_{\lambda}^2}{2W} + \sum_{i=1}^{N-1} \sum_{j>i}^{N} U(q_{ij}) + \sum_{i=1}^{N-1} \sum_{j>i}^{N} U(q_{ij}, \lambda) + (g_N + 1)kT \ln(s) + U(\lambda), \tag{2}
\]

where \( T \) is the equilibrium temperature, \( k \) is the Boltzmann constant, \( g_N \) represents the degrees of freedom for the physical system, and primed summations exclude intramolecular interactions. The functions \( U(\lambda) \) and \( (g_N + 1)kT \ln(s) \) are the potential energy terms for the extension variables. The third and fourth terms in the Hamiltonian are the kinetic energies of the extension variables and allow for the derivation of their equations of motion. The “mass” parameters for the extension variables, \( Q \) and \( W \), have units of energy \( \times \) time\(^2\). They may also be viewed as governing the extent of coupling between the system and a virtual, constant temperature, particle bath. The first and fifth terms in the equation are the Hamiltonian for an \( N \) particle system with momenta, \( \{p_1, p_2, ..., p_N\} \), and positions, \( \{q_1, q_2, ..., q_N\} \). The second term is the kinetic energy of the fractional particle whose momentum is given by \( p_f \) and whose position is given by \( q_f \). The sixth term is the fractional particle potential energy.

The fractional particle interactions may be scaled as

\[
\sum_{i=1}^{N-1} \sum_{j>i}^{N} U(q_{ij}, \lambda) = f(\lambda) \sum_{i=1}^{N} U(q_{ii}), \tag{3}
\]

where the function \( f(\lambda) \) must obey the boundary conditions

\[
f(0) = 0, \quad f(1) = 1. \tag{4}
\]

The potential energy for the particle number extension variable is given by

\[
U(\lambda) = -(N + h(\lambda))\mu - U_{\text{bias}}(\lambda) + h(\lambda)\mu_{\text{id}}, \tag{5}
\]

where \( \mu \) is the chemical potential, \( U_{\text{bias}}(\lambda) \) is the bias potential, and \( \mu_{\text{id}} \) is the ideal chemical potential for an \( N + 1 \) particle system. This form is chosen to assure sampling from a grand canonical ensemble distribution, as discussed below. The function \( h(\lambda) \) must obey the same boundary conditions as \( f(\lambda) \), but need not be identical to it. The ideal chemical potential may be defined either with or without a rotational
contribution, which in turn alters the value of the total chemical potential by an additive constant. This corresponds to a change in thermodynamic reference state, and has no other effect on simulation trajectories or measured thermodynamic properties of the system.

Equations of motion may be derived from Eq. (2) in the standard way. These equations may be transformed from virtual space to real space by making the substitution $dt = s \, dt'$, where the prime indicates real space. The second-order equations of motion in real space are then given by

$$
m_i \ddot{q}_i = - \sum_{j \neq i}^N \frac{\partial U(q_j)}{\partial q_i} - \frac{\partial U(q_i)}{\partial q_i} - \frac{m_i \dot{q}_i^2}{s},
$$

$$
m_j \ddot{q}_j = - \frac{f(\lambda)}{\lambda} \sum_{i=1}^N \frac{\partial U(q_i)}{\partial q_j} - \frac{m_j \dot{q}_j^2}{s},
$$

$$
Q \ddot{s} = s \sum_{i=1}^N m_i q_i^2 + m_j \dot{q}_j^2 - g_N kT + \frac{Q s^2}{s},
$$

$$
W \ddot{s} = s \left[ - \frac{\partial f(\lambda)}{\partial \lambda} \sum_{i=1}^N U(q_i) - \frac{\partial U(\lambda)}{\partial \lambda} + \frac{W \dot{s}}{s} \right],
$$

where the primes have been omitted.

Substituting the equation for $U(\lambda)$ into Eq. (9) gives

$$
W \ddot{s} = s \left[ - \frac{\partial f(\lambda)}{\partial \lambda} \sum_{i=1}^N U(q_i) + \frac{\partial h(\lambda)}{\partial \lambda} \mu + \frac{\partial U_{bias}}{\partial \lambda} - \frac{\partial h(\lambda)}{\partial \lambda} \mu_{id} \right] + \frac{W \dot{s}}{s}.
$$

This corresponds to Eq. (2.10) of Lo and Palmer with an extra term for the bias potential. The excess chemical potential is defined as $\mu_{ex} = \mu - \mu_{id}$. Substituting this definition into Eq. (10) yields

$$
W \ddot{s} = s \left[ - \frac{\partial f(\lambda)}{\partial \lambda} \sum_{i=1}^N U(q_i) + \frac{\partial h(\lambda)}{\partial \lambda} \mu_{ex} + \frac{\partial U_{bias}}{\partial \lambda} \right] + \frac{W \dot{s}}{s}.
$$

It is therefore possible to use either the excess chemical potential or the total chemical potential as an input parameter for influencing fluctuations in the particle-number variable. While either approach seems reasonable for systems with small density fluctuations, constraining the total chemical potential may be preferred since this both mimics experimental conditions and explicitly includes the $N$ dependence of the chemical potential.

### B. Partition function

The grand canonical partition function, $Q_{\mu VT}$, may be written in terms of canonical partition functions,

$$
Q_{\mu VT} = \sum_{N=0}^\infty Q_{NVT} \exp(\beta \mu N),
$$

where $\beta = 1/kT$. The probability ratio for states with particle number $N+1$ and $N$ is given by

$$
\frac{Q_{(N+1)VT}}{Q_{NVT}} = \exp(\beta \mu).
$$

It will now be shown that the extended-system Hamiltonian generates a partition function that when evaluated at its $\lambda$ end-points gives an equivalent ratio. The extended-system provides a smooth free-energy interpolation between discrete particle numbers that allows for `continuous' sampling of the grand canonical ensemble.

Simulations of the extended system sample from a microcanonical ensemble whose partition function is

$$
\Xi = \sum_{N=0}^\infty \frac{\hbar^{-3(N+1)}}{N!} \int dp_x \int ds \int dp_x \int \limits_0^1 d\lambda \int dp^N \int dq^N \int dp_f \int dq_f \exp(-\beta H_{\lambda}),
$$

where

$$
H_0 = \sum_{i=1}^N \frac{p_i^2}{2m_i} + \sum_{j}\left( \sum_{i=1}^N U(q_j) + \sum_{i=1}^N U(q_{ij}) \right),
$$

$$
H_{\lambda} = \frac{p_{\lambda}^2}{2W} + U(\lambda).
$$

Integration with respect to $p_{\lambda}$ yields
\[
\Xi = \sum_{N=0}^{\infty} \frac{h^{-3(N+1)}}{N!} \sqrt{2\pi QkT} \sqrt{2\pi WkT} \exp(\beta E) \\
\times \int_0^1 d\lambda \int dp^N \int dq^N \int dp'_f \int dq'_f \exp(-\beta U(\lambda)) \exp(-\beta U(\lambda)).
\]
(19)
The reduced partition function, \(\Xi_{N+\lambda}\), is defined by its relationship to the full partition function,
\[
\Xi_{N+\lambda} = \sum_{N=0}^{\infty} \int_0^1 \Xi_{N+\lambda} d\lambda.
\]
(20)
The ratio of the reduced partition function for the 0 and 1 configurations is
\[
\frac{\Xi_{N+1}}{\Xi_N} = \frac{\exp(-\beta U(1)) Q_{N+1}}{\exp(-\beta U(0)) Q_N} \frac{Q_{N+1}^{id}}{Q_N^{ex}}.
\]
(21)
\(Q_{N+1}^{id}\) and \(Q_N^{ex}\) are the kinetic and potential parts of the partition function. The \(N+1\) term in the denominator is a direct result of not scaling the kinetic energy of the fractional particle. Substitution of Eq. (5) into Eq. (21) gives
\[
\frac{\Xi_{N+1}}{\Xi_N} = \exp(-\beta U_{bias}(1) - \beta U_{bias}(0)) \times \frac{\exp(\beta(N+1)\mu - \beta(N)\mu) Q_{N+1}}{Q_N},
\]
(22)
which is equivalent to the grand canonical ratio in Eq. (13) if \(U_{bias}(0) = U_{bias}(1)\).
(23)
Effects of the bias potential on calculated ensemble averages do not need to be subtracted out as long as this last restriction is obeyed.

C. Bias potential

The equilibrium particle number distribution is governed by a free energy, \((-pV)_{N+\lambda}\), that is related to the reduced partition function by
\[
(-pV)_{N+\lambda} = -kT \ln(\Xi_{N+\lambda}).
\]
(24)
Using this relationship, the \(\lambda\)-coordinate free energy may be expressed in terms of an integral of the mean force along that coordinate,
\[
\Delta(-pV)_\lambda = \int_0^\alpha \left[ \left( \frac{\partial H^\lambda}{\partial \lambda} \right)_{N+\lambda} + \frac{dU(\lambda)}{d\lambda} \right] d\lambda
\]
\[
= \int_0^\alpha \left[ \sum_{i=1}^{N} \frac{\partial U(q_i^f, \lambda)}{\partial \lambda} \right]_{N+\lambda} + \frac{dU(\lambda)}{d\lambda} d\lambda
\]
\[
= \mu^{calc}_ex(a) - \mu^{ex}h(a) - U_{bias}(a) + U_{bias}(0).
\]
(25)

The fractional-particle excess chemical potential, \(\mu^{calc}_ex(a)\), may be determined from a canonical ensemble simulation using a single-particle TI technique.\(^{30}\) It should not be confused with \(\mu^{ex}\), the input excess chemical potential used in the GMD simulations. The excess Helmholtz free energy, calculated through a TI-scaling of all solvent particles simultaneously,\(^{30}\) differs from the excess chemical potential by \(kT\),
\[
\Delta A_{ex} = \mu^{calc}_ex + kT = \int_0^\alpha \left( \sum_{i=1}^{N} \frac{\partial U(q_i^f, \lambda)}{\partial \lambda} \right) d\lambda + kT.
\]
(26)
The bias potential term in the extended-system Hamiltonian allows for modification of the free-energy path between \(\lambda\) end points. In the fashion of “umbrella sampling,” a logical choice for the bias potential would effectively flatten the free-energy profile in order to achieve uniform sampling and increase computational efficiency.\(^{2}\) A linear path between integer-\(\lambda\) configurations may be achieved through setting the second derivative of the free energy to zero,
\[
\frac{\partial^2(-pV)_{N+\lambda}}{\partial \lambda^2} = \frac{\partial}{\partial \lambda} \left( \sum_{i=1}^{N} \frac{\partial U(q_i^f, \lambda)}{\partial \lambda} \right)_{N+\lambda}
\]
\[
- \mu^{ex} \frac{d^2h(\lambda)}{d\lambda^2} - \frac{d^2U_{bias}(\lambda)}{d\lambda^2} = 0.
\]
(27)
An equation for the bias potential may then be constructed by two successive integrations with respect to \(\lambda\), producing
\[
\int_0^\alpha \frac{dU_{bias}(\lambda)}{d\lambda} d\lambda = \int_0^\alpha \left( \sum_{i=1}^{N} \frac{\partial U(q_i^f, \lambda)}{\partial \lambda} \right) d\lambda
\]
\[
- \mu^{ex} \int_0^\alpha \frac{d^2h(\lambda)}{d\lambda^2} d\lambda + C = \int_0^\alpha d\lambda,
\]
(28)
where \(C\) is an integration constant. The final expression for the bias potential is obtained by carrying out the integrations and replacing \(a\) with \(\lambda\),
\[
U_{bias}(\lambda) = U_{bias}(0) + \mu^{calc}_ex(\lambda) - \mu^{ex}h(\lambda) + C.
\]
(29)
The bias potential required for a linear progression between \(\lambda\) configurations may be obtained from Eq. (29) for any given \(\mu^{ex}\) and \(h(\lambda)\) following a TI determination of \(\mu^{calc}_ex(\lambda)\). The constant \(C\) is chosen to assure the equivalence of \(U_{bias}(1)\) and \(U_{bias}(0)\). This approach implies that a separate thermodynamic integration must be carried out for each configuration characterized by a different value of \(N\). However, if the free energy surface changes little relative to \(kT\) for the values of \(N\) which are sampled during a simulation, then only one thermodynamic integration is necessary.

III. SIMULATION DETAILS

The SPC/E model for liquid water of Berendsen et al.\(^{28}\) was employed in all simulations. Each SPC/E water molecule consists of three atomic sites and has a rigid geometry with bond lengths of 1 Å and a tetrahedral bond angle of 109.47°. A Lennard-Jones (LJ) interaction between oxygen atoms provides a weak van der Waals attraction and a short-range intermolecular repulsion, while partial charges on the atomic sites account for the remaining electrostatic interactions. Each molecule has a permanent dipole moment of 2.35
D that is enhanced significantly over the experimental gas-phase value (\(\sim 1.85\) D) at ambient temperature.\textsuperscript{30} Thus, SPC/E water, while not explicitly including polarizability, accounts for water polarization in the liquid environment in an average fashion. The model differs from its SPC parent\textsuperscript{30} as it is parametrized to include a self-polarization energy correction that is necessary, but often neglected, in static polarization models.

The total potential energy for a system of SPC/E water molecules is given by

\[
U = \sum_i^{N} \sum_j^{N} \left[ 4\epsilon \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + q_i q_j \frac{e^2}{r_{ij}},
\]

\[
= U_{LJ} + U_C,
\]

where the self-polarization energy correction has been included implicitly. The values of the potential parameters are given in Table I. The Coulomb part of the potential energy, \(U_C\), was evaluated for the periodically replicated system using the Ewald summation method,\textsuperscript{1,41} with a total of 690 wave vectors and a value of \(\kappa L = 5.5\). All short-range interactions were truncated using the minimum image convention.

For most calculations, the simulation system was composed of a periodically replicated cube of side length 14.74 Å. For 107 water molecules, this corresponds to a density of \(1.0\) g/cm\(^3\). Thermodynamic integration calculations performed with more or less water molecules were maintained at the same water density. The target temperature was 300 K. Simulations were performed using a modified version of the SOLVENT MD code described elsewhere.\textsuperscript{32} The equations of motion were integrated using a 2.0 fs time step and a modified velocity-Verlet algorithm.\textsuperscript{36,43} The constrained dynamics were achieved through the use of the RATTLE\textsuperscript{44} variant of the SHAKE\textsuperscript{45} algorithm. Initial configurations were equilibrated in the canonical ensemble for a minimum 20 ps.

### A. Thermodynamic Integration

Calculations of \(\mu_{\text{calc}}(\lambda)\) were performed using a “slow growth” version of the TI method,\textsuperscript{48} in which the average in the integrand of Eq. (26) is calculated from a single time step and \(\lambda\) is varied continuously and slowly. All calculations were performed in both forward and reverse \(\lambda\) directions. The rate of \(\lambda\) variation was chosen such that there was no systematic hysteresis observed.

Special care is required in calculating free energies of molecule creation or annihilation processes.\textsuperscript{29} For example, if the function \(\sum_i^{N} U(q_i,\lambda)\) is taken to be a simple linear scaling of the full interaction potential,

\[
\sum_i^{N} U(q_i,\lambda) = \lambda \sum_i^{N} U(q_i),
\]

then evaluation of the TI integral in Eq. (26) is problematic. The integrand in the equation will be proportional to \(\lambda^{-34}\) as \(\lambda\) approaches zero since the short-range repulsions are dominated by the \(r^{-12}\) LJ term in the interaction potential.\textsuperscript{34,35}

This and similar problems have been dealt with in a variety of ways, two of which have been incorporated here. The first maintains linear scaling but with a modification of the short-range part of the interaction potential, and is similar to the method used by Mon and Griffiths.\textsuperscript{33} The second approach uses nonlinear \(\lambda\) scaling.\textsuperscript{29,34,35}

The high-energy portions (\(\gg kT\)) of the water interaction potential may be modified arbitrarily without impacting water behavior or calculated averages. Here, the interaction potential has been modified at short distances to make the interaction energy at zero separation finite. This assures that the integrand in Eq. (26) is finite and well-behaved as \(\lambda\) approaches zero. Modifications were restricted to distances that are never (over the course of a typical simulation) sampled by the fractional particle in the limit of \(\lambda = 1\). Thus, while the configuration spaces for \(\lambda = 0\) (noninteracting particle) and \(\lambda = 1\) (full particle) remain unaltered, the path by which they are connected has been changed. Since the free energy is a state function, its value, calculated as an integral over the new path, will be unchanged.

Nonlinear \(\lambda\) scaling provides an alternative approach that does not require modification of the interaction potentials. Here, the function \(\sum_i^{N} U(q_i,\lambda)\) was taken to be

\[
\sum_i^{N} U(q_i,\lambda) = \lambda^2 \sum_i^{N} U(q_i).
\]

If the fractional particle energy is a function of \(\lambda\) to the fifth power or higher, the integrand in Eq. (26) is well-behaved.\textsuperscript{29}

The free energies calculated from the TI method were used to construct bias potentials for the GMD simulations. In addition, since \(\mu_{\text{calc}}\) is equal to the excess chemical potential of bulk water at the simulation density,\textsuperscript{30} the TI calculations provided an independent measure of the chemical potential for comparison with the GMD results.

### B. GMD

GMD simulations were performed using both linear and nonlinear scaling, as in the TI calculations, by setting \(f(\lambda)\) in Eq. (3) to \(\lambda\) and \(\lambda^2\), respectively. The short-range part of the interaction potential was again modified for the linear-scaling case. Bias potentials were constructed by fitting the expression on the right-hand side of Eq. (29) with a tenth-order polynomial, where the value of \(h(\lambda)\) was set to \(\lambda\) and \(C\) was chosen such that \(U_{\text{bias}}(1) = U_{\text{bias}}(0)\). Separate bias potentials were constructed for the linear and nonlinear scaling cases from the corresponding \(\mu_{\text{calc}}(\lambda)\) functions.

When a fractional particle becomes full, at \(\lambda = 1\), that particle is reassigned to the full-particle list. The coordinates of a new fractional particle are determined randomly and its velocity is selected from a Maxwell–Boltzmann distribution. Likewise, when \(\lambda\) decreases to zero, the fractional particle is eliminated and a new fractional particle is chosen at random from among the remaining particles. For both cases, the velocity of the coupling parameter coordinate, \(\lambda\) is unaltered.
TABLE II. The excess chemical potential for the SPC/E model of liquid water. Estimated uncertainty in \( \mu_{\text{ex}} \), for all calculations is \( \pm 0.1 \) kcal/mol.

<table>
<thead>
<tr>
<th>Method</th>
<th>System size/molecules</th>
<th>( \mu_{\text{ex}} ) (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thermodynamic integration</td>
<td>50</td>
<td>-5.7</td>
</tr>
<tr>
<td></td>
<td>107</td>
<td>-5.7</td>
</tr>
<tr>
<td></td>
<td>207</td>
<td>-5.7</td>
</tr>
<tr>
<td>GMD (( \lambda ) scaling)</td>
<td>107</td>
<td>-5.6</td>
</tr>
<tr>
<td>GMD (( \lambda^5 ) scaling)</td>
<td>107</td>
<td>-5.6</td>
</tr>
</tbody>
</table>

This procedure differs significantly from that of Pettitt and co-workers,20,22,25 who employed biased insertions and deletions in an attempt to match particle number coordinate accelerations. We find, as did Lo and Palmer,24 that biased insertions become unnecessary when the singularity at the uncoupled particle state is removed. The insertion and deletion approaches used here sample randomly from the equilibrium distributions for the \( \lambda = 0 \) and \( \lambda = 1 \) states, respectively.

IV. RESULTS AND DISCUSSION

A. Thermodynamic Integration

TI calculations with linear \( \lambda \) scaling were performed in both forward and reverse directions for systems containing 50, 107, and 207 full particles and one fractional particle. A \( \lambda^5 \)-scaling calculation was performed in both directions for a single, 107 particle system. The total simulation time for each calculation was at least 1.0 ns. Results are presented in Table II. The reported excess chemical potentials have been adjusted by the self-polarization free energy of 1.25 kcal/mol. The value of this correction is equivalent to the self-polarization energy for SPC/E water, as justified in the Appendix. A ‘‘tail correction’’ for the LJ contributions to the energy from molecules outside the simulation box was also added to the calculated chemical potentials. This has a negligible influence on all but the smallest system results.

The measured excess chemical potentials are -5.7 kcal/mol, with no significant size effects detected. The linear and nonlinear scaling results agree to within the estimated uncertainty (\( \pm 0.1 \) kcal/mol) of the calculations. Shelley \textit{et al.} measured an equivalent value using a grand canonical MC approach.10 Hermans \textit{et al.} used a TI method to calculate the excess Helmholtz free energy of SPC/E water at 300 K.39 After applying the self-polarization free energy correction and adding \( kT \), their measured excess chemical potential is also equal to -5.7 kcal/mol. None of these results is in particularly good agreement with the experimental excess chemical potential of -6.34 kcal/mol.39

B. GMD

The advantage of including a bias potential in the GMD simulations may be illustrated by plotting free-energy profiles for the system obtained through TI calculations as shown in Fig. 1. Profiles for both \( \lambda \) and \( \lambda^5 \) scaling were calculated from Eq. (25) using an input excess chemical potential of -5.7 kcal/mol and excluding the bias potential. The barrier heights for the two paths differ due to modifications in the short-range interaction potential for the linear case. In both cases, it is clear that thermal energies will generally be inadequate to provide sampling across the large free-energy barrier. As noted above, the bias potential is constructed to essentially eliminate this free energy peak.

Bias potential GMD simulations were performed using both linear scaling with fixed excess chemical potential and nonlinear scaling with fixed total chemical potential. The excess chemical potential ranged between approximately -5.0 and -6.4 kcal/mol. These values correspond to ‘‘corrected’’ chemical potentials for the SPC/E model. The system was equilibrated in the grand canonical ensemble for at least 200 ps at each chemical potential. A typical particle number trajectory is given in Fig. 2. Averages were subsequently collected for a minimum of 2 ns. The calculated particle number averages are plotted in Fig. 3 as a function of the excess chemical potential. The data were fit via linear regression, with the corresponding lines used to extract values for \( \mu_{\text{ex}} \), at a water density of 1.0 g/cm\(^3\). The results, presented in Table II, are both in good agreement with the TI values.
The extension variable masses $Q$ and $W$ were chosen to be 0.38 and 190 ps$^2$ kcal/mol, respectively. While the choice for $Q$ is standard and should be reasonable, the value of $W$ was not optimized and may be larger than necessary. The reported simulation times ($\geq 2$ ns) may therefore be longer than required. Work currently in progress will further address this issue.

V. CONCLUSIONS

A bias potential was incorporated into an extended-system GMD simulation algorithm. The bias potential allows for systematic and theoretically sound modification of the free energy of noninteger regions of the particle number coordinate. This enhances computational efficiency in sampling the full range of the particle-number variable in the grand canonical simulations. The method may be applied with arbitrary coupling parameter paths as demonstrated through the uses of both linear and nonlinear potential scaling functions.

Though an added computational expense, the separate TI calculation of $\mu_{\text{ex}}^\text{calc}(\lambda)$ provides a significant benefit to the bias potential GMD method; since the free-energy profile is determined explicitly, it may then be adjusted systematically. The resulting gain in sampling efficiency may more than counterbalance the expense of the TI computation. In general, the TI simulation need not be done with high precision or accuracy, as any reasonable estimate (within $\sim kT$) of the shape of $\mu_{\text{ex}}^\text{calc}(\lambda)$ should yield an effective bias potential. In addition, relatively few TI calculations should be required since a bias potential determined for one system (bulk water) may be adequate for use in others (ionic solutions).

The bias potential GMD method is currently being applied in our group for modeling of the ion exchange properties of clay minerals. Ion exchange on clays is often accompanied by changes in interlayer hydration that are driven by chemical potential gradients. The GMD method will allow for modeling of ion exchange under conditions of constant water chemical potential. The highly constrained interlayer environment of low water content clays should provide an interesting challenge for the method.

ACKNOWLEDGMENT

This work was supported by the US Department of Energy Environmental Management Science Program.

APPENDIX: SELF-POLARIZATION CORRECTION

The chemical potential of SPC/E water, calculated directly from the TI simulations and used as an input parameter in the GMD simulations, must be corrected for the self-polarization energy of the water molecules. The proper self-polarization free-energy correction is therefore derived in this Appendix.

The theory for the interaction potentials of polarizable molecules has been presented previously and will not be reproduced here. One contribution to the potential energy of a polarizable system is a positive self-polarization energy that depends upon both the molecular polarizability and the extent of polarization. For effective pair potentials with static polarization, such as SPC/E water, Berendsen et al. showed that the self-polarization energy, $U_{\text{sp}}$, is given by

$$U_{\text{sp}} = \sum_{i} (p - p^{0})^{2}/2\alpha,$$

where $p$ is the total molecular dipole moment, $p^{0}$ is the isolated-molecule dipole moment, and $\alpha$ is the molecular polarizability. The SPC/E model utilizes atom-centered point charges to represent the molecular dipole, with the charges enhanced over their gas-phase values to provide for the static polarization. The total potential energy of a system of SPC/E water molecules is therefore given by

$$U = U_{\text{LJ}} + U_{C} + U_{\text{sp}},$$

where $U_{\text{LJ}} + U_{C}$ is the energy measured directly in the simulation. The potential energy of the system, $U$, is obtained through addition of the self-polarization energy. For the SPC/E model, the self-polarization energy correction, per water molecule, is taken to be 1.25 kcal/mol. Note that this energy correction, while not explicitly included in the Hamiltonian in Eq. (2), is applied implicitly whenever determining the potential energy of a system containing SPC/E water molecules.

As described in the text, the excess chemical potential is calculated using the TI method from

$$\mu_{\text{ex}}^\text{calc} = \int_{0}^{1} \left( \frac{\partial U}{\partial \lambda} \right)_{\lambda} d\lambda,$$

where $U$ is the potential energy of the fractional particle and is given by

$$U = \sum_{i=1}^{N} U(q_{i\lambda},\lambda) + U_{\text{sp}}^{\text{f}}(\lambda).$$

The last term represents the self-polarization energy of the fractional particle. Substituting this expression into Eq. (A3) gives

$$\mu_{\text{ex}}^\text{calc} = \int_{0}^{1} \left( \sum_{i=1}^{N} \frac{\partial U(q_{i\lambda},\lambda)}{\partial \lambda} \right) d\lambda + \int_{0}^{1} \frac{\partial U_{\text{sp}}^{\text{f}}(\lambda)}{\partial \lambda} d\lambda.$$
where the last term defines the self-polarization free-energy for the fractional particle, denoted $U_{\text{corr}}$.

In order to calculate the free-energy correction, a potential scaling function for the fractional particle charges must be chosen. The scaling function is arbitrary as long as it varies smoothly and obeys the restriction of Eq. (4). It is convenient to scale the charges in both the polarized and unpolarized molecules by the same function, $f(\lambda)$. Equation (A1) may then be expressed as

$$U_{\text{sp}}^f(\lambda) = \frac{(f(\lambda)p - f(\lambda)p^0)^2}{2\alpha}. \quad (A6)$$

Substituting back into the last term of Eq. (A5) results in the final expression for the self-polarization free-energy,

$$U_{\text{corr}} = \int_0^1 \frac{\alpha}{\partial \lambda} \frac{(f(\lambda)p - f(\lambda)p^0)^2}{2\alpha} d\lambda \quad (A7)$$

$$= \frac{(p - p^0)^2}{2\alpha}. \quad (A8)$$

The self-polarization free energy is therefore equivalent to the self-polarization energy for a single particle, and a correction of 1.25 kcal/mol should be applied to both calculated ($\mu_{\text{ex}}$) and input ($\mu_{\text{ex}}$) values of the excess chemical potential. Although the fractional particle self-polarization energy depends on the coupling parameter, the GMD simulations may also follow the dynamics of a system of uncorrected water molecules, with a self-polarization correction added after the fact. Including the self-polarization energy dynamics alters the free-energy surface between integer-$\lambda$ states and adjusts the full particle end points by exactly the self-polarization free energy. Thus, the corrected chemical potential, calculated in either fashion, is equivalent. Finally, it should also be noted that the self-polarization free energy is not equal to the difference in free energy between SPC and SPC/E water, reported recently to be 0.83 kcal/mol.50