JCTC Journal of Chemical Theory and Computation

Accelerated Superposition State Molecular Dynamics for Condensed Phase Systems

Michele Ceotto

Dipartimento di Chimica Fisica ed Elettrochimica, Università degli Studi di Milano, via Golgi 19, 20133 Milano, Italy

Gary S. Ayton and Gregory A. Voth*

Department of Chemistry and Center for Biophysical Modeling and Simulation, University of Utah, Salt Lake City, Utah 84112

Received December 1, 2007

Abstract: An extension of superposition state molecular dynamics (SSMD) [Venkatnathan and Voth *J. Chem. Theory Comput.* **2005**, *1*, 36] is presented with the goal to accelerate timescales and enable the study of "long-time" phenomena for condensed phase systems. It does not require any a priori knowledge about final and transition state configurations, or specific topologies. The system is induced to explore new configurations by virtue of a fictitious (free-particle-like) accelerating potential. The acceleration method can be applied to all degrees of freedom in the system and can be applied to condensed phases and fluids.

I. Introduction

Molecular dynamics (MD) simulation models the time evolution of a given atomistic-level system by integrating Newton's equations of motion.[1–3] Extensions to MD employ dynamics described by a Langevin equation[4–6] or variations of the previous differential equations with additional parameters (e.g., friction coefficients) or degrees of freedom (e.g., Nosé−Hoover thermostats[1,7–9]). The integration of the underlying equations of motion are limited in time; in fact, most MD simulations are far too limited in duration to examine many important biomolecular processes which can occur on timescales longer than tens of nanoseconds. For example, lateral diffusion in lipid bilayers occurs on the second scale,[10] while protein folding occurs on the millisecond scale.[11–15] On the other hand, relying on the rapid growth in computer technology employing a "brute force" MD approach is also not feasible as a speedup of 6 orders of magnitude will be required in order to access the relevant timescales.

In some complex systems, the multiple time- and length-scales can lead to the so-called effect of "broken ergodicity", where mechanical observables (e.g., internal energy, pressure) calculated via accumulated time averages differ (sometimes greatly) from the ensemble average (where all points in phase space are considered). In other words, the timescale used for the measurement is much shorter than the actual relaxation time for the system. As a result, in this particular case, the MD simulation performs a sum of subaverages of isolated phase space subsets and misses some others. As an example of an approach to overcome these issues, Andricioaei and Straub[16] have introduced new generalized Monte Carlo (MC) and MD algorithms inspired by Tsallis statistics (q-jumping) and later, "smart walking" MC.[17] Other approaches include the multicanonical algorithm,[18,19] where extensive macrovariables are added, simulated tempering,[20,21] and replica exchange (REX),[22–25] where intensive thermodynamic state quantities (pressure, temperature, chemical potentials, etc.) are varied. All the above methods have been implemented both in MC and MD simulation algorithms. A main limitation involves the number of replicas; these can grow unmanageably large when many macrostates are required for the simulation to satisfy ergodicity. Fenwick et al.[26] have combined MC replica methods with biased force-field parameters in order to directly modulate the specific molecular interaction responsible for the kinetic traps and, in the same spirit, others[27,28] have modified force-field

* Corresponding author. E-mail: voth@chem.utah.edu.

10.1021/ct7003275 CCC: $40.75
Published on Web 03/04/2008

parameters (except those related to solvent−solvent interactions) in an MD protocol.

Another approach for timescale acceleration is the parallel replica method (PRD),[29–31] where the system is replicated in parallel and independent MD trajectories are generated via different initial velocity distributions. Whenever a successful trajectory is obtained, all processors are stopped. This state is then replicated over all processors, and the whole process is restarted. A further implementation of this method is the parallel sequential synchronization (PSS)[32] which shows how PRD is easy to combine with other techniques.

To accelerate timescales and enable the study of "long-time" phenomena within the MD framework, one can perform simulations at higher temperatures. Voter[33] has developed, and later improved,[29] a method called temperature accelerated dynamics (TAD) which raises the temperature and corrects for this bias by filtering out transitions that would not occur at the original temperature. A completely different approach for "rare events" dynamics (especially for passages over high barriers) employs accelerating potentials.[34–45] Here, the potential energy surface is modified for a small set of degrees of freedom and the original state is recovered on-the-fly by means of non-Boltzmann weights. Conformational flooding[46] first selects a subconformational space and then destabilizes the initial conformation and consequently lowers the free energy barrier of structural transitions. Similarly, the hyperdynamics method[29,35,36] focuses on infrequent transition events from one potential energy basin to another and then constructs a bias potential such that the original potential energy surface changes are done without affecting the transition state regions where the rate is calculated using the harmonic limit of transition state theory. Significant boosts have been found for surface diffusion dynamics;[29,34–36,42] however, for more complex systems, the construction of such a bias potential is not trivial and may not solve the low barrier problem when the system is trapped by a set of states connected by low barriers. A simpler recipe is one offered by Tully and co-workers,[37,41] where the bias potential is chosen so that the system evolves in a flat "puddle region" instead of sinking into a local minimum. This method has been applied to dihedral degrees of freedom in small peptide dynamics[37–39,41,47] and has been implemented in MC simulation with a bias in momentum space.[37,48,49] Recent work has employed a "boost potential" to modify the original potentials that govern the system.[44,50–52] The scheme raises the wells depths in a continuous manner and has been successfully applied, for example, to alanine dipeptide in an explicit solvent. In conjunction with a quasiharmonic analysis, this approach has been used to calculate the entropy for an eight residue peptide in explicit water.[44]

A different solution to the direct kinetic dynamics of rare events is transition path sampling.[53–55] In this method reactants and products are known a priori, and path ensembles between these states are generated by constructing a random walk in path space with a MC algorithm.

In solid-state simulations, a popular choice to accelerate MD is feature activated molecular dynamics (FAMD),[56] which creates localized regions around a defect atom or cluster (active region) sites. A full MD simulation is subsequently performed while the rest of the solid lattice is kept static. This is done by the use of a thermal activity function, which decays to zero in a sigmoid fashion once out of the active region.

Although the above accelerating potential procedures partially removes the problem of broken ergodicity and accelerates the crossing of barriers, it is not suitable for calculating equilibrium thermodynamic properties as it undersamples low-energy states. Self-guided molecular dynamics approaches[57–61] can avoid this problem and can be implemented without any a priori knowledge of the system; rather they are based on the cumulative history of a system's trajectory. In particular, when this idea is combined with replica exchange,[61] the copies are self-regulating and compete during the simulation to overcome to any undersampled region. Along the same line of an adaptive algorithm, Laio and Parrinello[62–65] introduced the so-called "metadynamics" method in which a history-dependent potential, given as the sum of Gaussians centered on the trajectory in a reduced "collective variable" space, fills the free energy surfaces minima and drives the systems to explore new wells.

In condensed phases, the complexity of the system is such that it is not possible to isolate a priori a subset of degrees of freedom responsible for the long-time properties. For example, a simple picture of a double well coupled to a bath with a starting and ending configuration is, most of the time, an inadequate description of the system. Consequently, a state-to-state transition model with a transition state configuration, which follows a first order kinetic picture, is often out of the question. An example problem that enters into this category is the bilateral diffusion of phospholipids in membranes; the possible configurations are so numerous that trying to isolate any subsets of degrees of freedom will mostly turn out to be impossible, if not counterproductive. In fact, the transition state that separates two multidimensional basins in a double well picture is misleading for such systems since no specific event and reaction subset of coordinates can be identified. To deal with these issues, a more generalized version of enhancing sampling must be adopted, which is not restricted to first order kinetics. Instead, the system should enchance its ergodic properties without any specific instructions or a priori topological constraints, and should be left to explore new configurations in an unrestricted fashion.

In light of the above considerations, the goal of the present work is to develop an approach within the MD framework that can deal with all the degrees of freedom of complex condensed phase systems, for example fluids, by not requiring any a priori knowledge of the system (such as the final state), and by guaranteeing a full recovery of the unbiased potential statistics. The starting point of the present work is the superposition state molecular dynamics (SSMD) method.[45] The main idea underlying SSMD is to create fictitious potentials whose dynamics are accelerated with respect to the original (physical) one. These potentials are built up as a superposition of the upper and lower states with respect the physical one and they are coupled in such a way that the overall dynamics smoothly and continuously switch

from one potential surface to another to achieve acceleration. The coupling follows the superposition state rules of quantum mechanics, and the construction of the fictitious potential and the location of the coupling term are crucial aspects of this approach.

In previous work,[45] SSMD was developed and applied to a one-dimensional rough potential energy landscape and it was proven to enhance ergodic behavior. Here, it is extended to a multidimensional potential energy landscape and is applied to the case of an isotropic fluid. An isotropic fluid is an example of the kind of complex system described previously as all degrees of freedom equally contribute to the ergodic behavior of the system.

The remaining sections of this paper are organized as follows: section II reviews the sampling techniques utilized in this work, while in section III the extension of superposition state molecular dynamics (SSMD) to condensed phase simulations is introduced. In section IV, SSMD is implemented for Coulombic interactions. Some closing comments are then given in section VI.

II. Non-Boltzmann Sampling

The aim of this section is to review non-Boltzmann sampling and later show its limitation to condensed phase simulations. A system of N particles with phase space coordinates $\mathbf{p} = (\mathbf{p}_1, \ldots, \mathbf{p}_N)$ and $\mathbf{r} = (\mathbf{r}_1, \ldots, \mathbf{r}_N)$, where $\mathbf{p}_i$ is the momenta and $\mathbf{r}_i$ is the position, of particle i is considered. The canonical phase space probability density is given by

$$P(\mathbf{r}, \mathbf{p}) = \frac{\mathrm{e}^{-\beta H(\mathbf{r},\mathbf{p})}}{Q} \tag{1}$$

where

$$Q = \iint \mathrm{d}\mathbf{r}\, \mathrm{d}\mathbf{p}\, \mathrm{e}^{-\beta H(\mathbf{p},\mathbf{r})} \tag{2}$$

is proportional to the canonical partition function. Here,

$$H(\mathbf{r}, \mathbf{p}) = \sum_{i=1}^{N} \frac{|\mathbf{p}_i|^2}{2m} + V(\mathbf{r}) \tag{3}$$

is the classical Hamiltonian and $\beta = 1/k_B T$, where k_B is Boltzmann's constant, T is the temperature, and $V(\mathbf{r})$ is the total potential energy function. In terms of just the coordinate space, eq 1 can be expressed in the usual way

$$P(\mathbf{r}) = \frac{\mathrm{e}^{-\beta V(\mathbf{r})}}{\int \mathrm{d}\mathbf{r}\, \mathrm{e}^{-\beta V(\mathbf{r})}} \tag{4}$$

The physical meaning of the probability density $P(\mathbf{r})$ is that the canonical ensemble average of a position-dependent observable $O(\mathbf{r})$ is given by

$$\langle O \rangle = \int \mathrm{d}\mathbf{r}\, P(\mathbf{r}) O(\mathbf{r}) \tag{5}$$

An MD estimate of this integral is given when eq 5 is replaced by the following sum:

$$\langle O \rangle \cong \frac{1}{N_T} \sum_k O_k(\mathbf{r}) \tag{6}$$

where the k index stands for the MD averaging step, and N_T is the total number of averaging steps. If a different "accelerated" MD potential $V'(\mathbf{r})$ is used, then the observable average in the original canonical ensemble can still be obtained by rewriting eq 5 in the following way,

$$\langle O \rangle = \frac{\int \mathrm{d}\mathbf{r}\, \mathrm{e}^{-\beta V(\mathbf{r})} O(\mathbf{r})}{\int \mathrm{d}\mathbf{r}\, \mathrm{e}^{-\beta V(\mathbf{r})}} = \frac{\int \mathrm{d}\mathbf{r}\, \mathrm{e}^{-\beta V'(\mathbf{r})} O(\mathbf{r}) \mathrm{e}^{-\beta(V(\mathbf{r})-V'(\mathbf{r}))}}{\int \mathrm{d}\mathbf{r} \mathrm{e}^{-\beta V'(\mathbf{r})} \mathrm{e}^{-\beta(V(\mathbf{r})-V'(\mathbf{r}))}} =$$

$$\frac{\langle O(\mathbf{r}) \mathrm{e}^{-\beta \Delta V(\mathbf{r})} \rangle}{\langle \mathrm{e}^{-\beta \Delta V(\mathbf{r})} \rangle} \tag{7}$$

or by the equivalent sum from the MD simulation as the potential $V'(\mathbf{r})$:

$$\langle O \rangle \cong \frac{\sum_k O_k(\mathbf{r}) \mathrm{e}^{-\beta \Delta V_k(\mathbf{r})}}{\sum_k \mathrm{e}^{-\beta \Delta V_k(\mathbf{r})}} \tag{8}$$

where $\Delta V_k(\mathbf{r}) = V_k(\mathbf{r}) - V_k'(\mathbf{r})$ is the difference between the original and the accelerated potential at the average step k and the dynamics is carried out on the new potential $V'(\mathbf{r})$. This approach has been in used for many decades now, both in MC and MD calculations. In essence, most of the methods described in the Introduction can be reduced to this essential idea. It has the significant advantage that all the terms, both in the numerator and denominator, of eq 8, are reweighted on-the-fly in the MD simulation.

Recall that the overall aim of the present work is to develop an acceleration scheme that is capable of treating condensed phase systems. As such, a first test employs a benchmark system and consisted of a Lennard-Jones fluid of $N = 1372$ particles characterized by the pair-interaction potential

$$V(r_{ij}) = 4\varepsilon\left[\left(\frac{\sigma}{r_{ij}}\right)^{12} - \left(\frac{\sigma}{r_{ij}}\right)^{6}\right] \tag{9}$$

where $r_{ij} = |r_i - r_j|$ is the interparticle separation, σ is the fundamental unit of distance, and ε is the fundamental unit of energy. The fundamental unit of energy ε, can be expressed as $\varepsilon = \gamma\varepsilon_c$, where γ is a dimensionless scaling parameter and ε_c is a constant. The same is valid for the fundamental unit of distance σ, i.e., $\sigma = \gamma_\sigma \sigma_c$. In all simulations $N\sigma_c^3/L^3 = 0.8$, the mass, m, is unity, and $k_B T/\varepsilon_c = 0.8$.

In a straightforward application of eq 8, an accelerated dynamics can be directly employed by using a new potential $V'(\mathbf{r})$ with $\gamma = 0.5$ and is shown by the dashed line in panel a of Figure 1. All other parameters were kept constant. This dynamics was performed for *all degrees of freedom*, and it is equivalent to the dynamics of the original reduced potential ($\gamma = 1$) but at double the temperature. Simulations employed 500 000 time steps in the constant NVT ensemble.[7,8] The reduced time step was set to $\Delta t_c = 0.0025$, which is related to the unscaled time step by the relation $\Delta t = \sqrt{\varepsilon_c/m\sigma_c}\Delta t_c$. An examination of the radial distributions functions, as shown in panel b of the same figure (with the same line convention as in a), shows that the original radial distribution

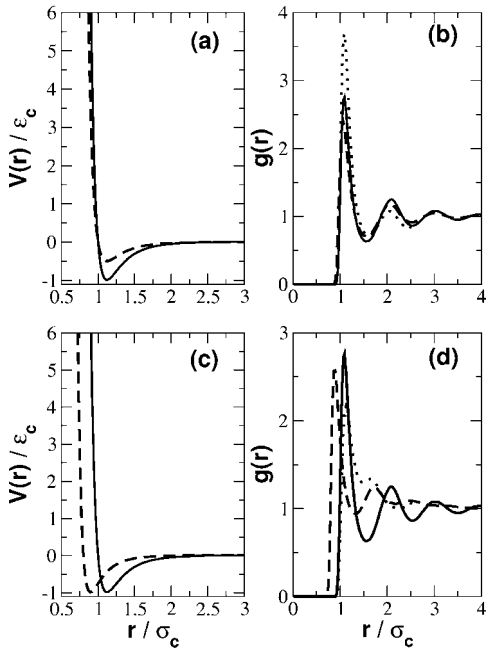

Figure 1. Potentials and radial distribution functions. (a) Standard Lennard-Jones potential ($\gamma = 1$) in the solid line and for $\varepsilon = 0.5\varepsilon_c$ and $\sigma = \sigma_c$ in the dashed line. (b) Radial distribution function at a reduced temperature equal to $k_BT/\varepsilon_c = 0.8$ and reduced density $N\sigma_c^3/L^3 = 0.8$ after 500 000 time steps for the standard Lennard-Jones potential in the solid line, for the half-deep well in the dashed line, and the reweighted one in a dotted line as an estimate of the standard distribution function as obtained by eq 8. (c) Same as a but with $\varepsilon = 1$ and $\sigma = 0.8\sigma_c$ in the dashed line. (d) Same as b where the dashed line is used for the corresponding potential in panel c.

function (continuous line) is not recovered by rescaling on-the-fly according to eq 8. In this and in subsequent instances, any rescaled quantities found from a direct application of eq 8 will be denoted as "re-weighted". In fact, the reweighted radial distribution function (dotted line) largely overestimates the first peak of $g(\mathbf{r})$. The dynamics was also examined, where here the core radius was scaled to $\gamma_\sigma = 0.8$ (dashed line in panel c of Figure 1) and the reweighted radial distribution function was obtained (dotted line in panel d of the same figure). Here, the correct location of the first shell–core is observed, but the magnitude is off and the longer-ranged correlations are not accounted for correctly. Since non-Boltzmann sampling is exact regardless of the number of degrees of freedom, a much longer simulation will fully recover the original radial distribution function, but at the same time, any advantage due to the new accelerating potential is lost. In fact, if the same simulation of the dashed potential in panel a of Figure 1 is performed at a much higher temperature ($k_BT/\varepsilon_c = 4.0$) and lower density ($N\sigma_c^3/L^3 = 0.2$) (i.e., a gas phase), the original radial distribution function is recovered as shown in Figure 2 by the dotted lines, even if it slightly misses the small second solvation shell at $r_{ij} = 2.5\sigma_c$.

The above results indicate that the sources of errors are both topological and numerical in nature. They are topological because the potential landscape is multidimensional, and with a higher dimensionality, a larger number of sampling

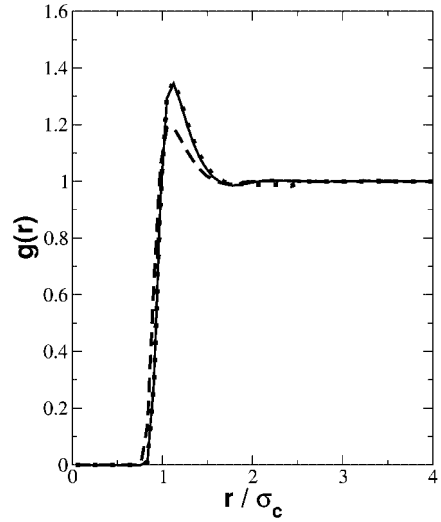

Figure 2. Same plot as panel b of Figure 1 but with the simulation performed at a reduced temperature of $k_BT/\varepsilon_c = 4.0$ and a reduced density $N\sigma_c^3/L^3 = 0.2$.

events are necessary to yield converged statistics. They are numerical because at high temperatures β is smaller, and the errors made in the exponential sum and its propagation in the exponentiation are accumulated.[43] Instead, at lower temperatures, a complex potential energy landscape may not be recovered by the reweighting expression in eq 8 in standard run lengths. This shows that an indiscriminate use of non-Boltzmann sampling techniques is not fruitful, even if these are in principle exact. This was also reported recently by de Oliveira et al.[51] for the case of a rigid water molecule condensed phase model, where only the oxygen−oxygen radial distribution function was considered. For this reason in section III, a still exact but at the same time practical and more efficient approach is presented.

III. Superposition State Molecular Dynamics (SSMD) for Condensed Phase Simulations

In superposition states molecular dynamics (SSMD),[45] an accelerating potential $V'(\mathbf{r})$ is obtained by a superposition of "states", which include the original (physical) one characterized by the potential $V(\mathbf{r})$, along with fictitious states characterized by an additional potential, $V_f(\mathbf{r})$. In the case of a two-state combination with one fictitious and one real potential, the effective SSMD potentials can be extended to multidimensional systems by superimposing each pairwise interaction at a time in the way described above. The resulting effective potential is derived by (analytically) diagonalizing a two-state matrix as

$$V'(r_{ij}) = \frac{V(r_{ij}) + V_f(r_{ij})}{2} \pm \frac{1}{2}\sqrt{4V_{12}^{\,2}(r_{ij}) + (V(r_{ij}) - V_f(r_{ij}))^2} \tag{10}$$

where $V_f(r_{ij})$ is the fictitious (not physical and in this case) pairwise potential, $V_{12}(r_{ij})$ is the coupling, and $V'(r_{ij})$ can be given both by the sum (the highest eigenenergy) or the difference (the lowest eigenenergy) in eq 10. In this paper, the lowest eigenenergy will always be used to perform the dynamics and the fictitious and coupling potentials will be varied.

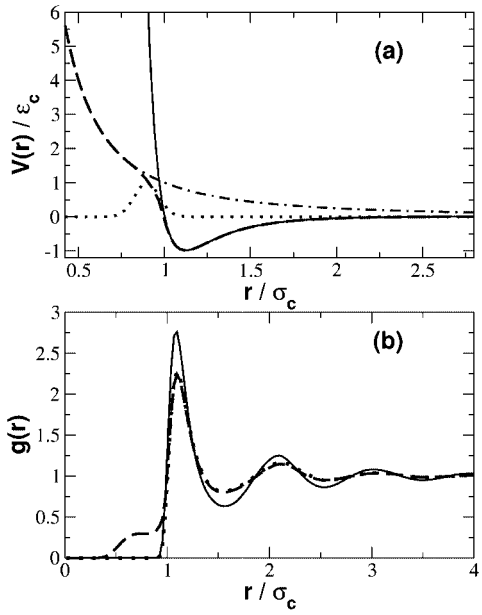

Figure 3. SSMD potential and radial distribution function at the same temperature and density as in Figure 1. (a) Standard Lennard-Jones potential in the solid line; upper fictitious state potential in the dashed-dot line; coupling potential in the dotted line; and the resulting SSMD potential in the dashed line. (b) Standard Lennard-Jones fluid radial distribution function in the solid line; SSMD potential radial distribution function in the dashed line; and the reweighted one in the dotted line. Note that the reweighting is exact on the left side of the first solvation shell peak.

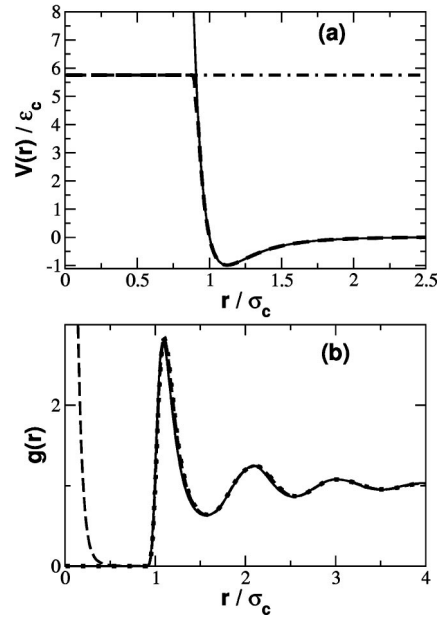

Figure 4. Potential and radial distribution functions at the same temperature and density as in Figure 1. (a) Original Lennard-Jones potential in the solid line; the upper fictitious SSMD state $V'(r_{ij}) = 5.75\varepsilon_c$ in the dashed-dot line; and the resulting SSMD potential in the dashed line. (b) Original radial distribution function in the solid line; SSMD potential radial distribution function in the dashed line; and reweighted radial distribution line in the dotted line. The reweighting fully recovers the original distribution function.

However, one also has the flexibility to use the upper eigenenergy or to employ more fictitious states in an *N*-state SSMD scheme.

As an example, the original LJ potential $V(r_{ij})$ was plotted in panel a of Figure 3 with a continuous line, whereas the fictitious potential (shown as the dot-dashed line) is given by

$$V_f(r_{ij}) = \varepsilon\left(\left(\frac{\sigma}{r_{ij}}\right)^2 + \delta\right) \tag{11}$$

where the arbitrary constant is $\delta = 4$, and with the dotted line, the coupling potential is given by

$$V_{12}(r_{ij}) = V_{12}(r_0)e^{-\alpha(r_{ij} - r_0)^2} \tag{12}$$

where $V_{12}(r_0) = \varepsilon$, $\alpha = (10/\sigma)^2$, and r_0 is the pairwise length where the switching between the real and the fictitious potentials occurs. The resulting SSMD potential $V'(r_{ij})$ is plotted in Figure 3 with a dashed line and it clearly follows the original potential up to the beginning of the repulsive wall. The corresponding radial distribution function (dashed line), the reweighted (dotted line) and the exact one (continuous line) are shown on panel b of Figure 3. It can be seen how the original one is completely recovered on the repulsive wall side, but it does not accurately describe the long-range correlation part (rdf structure).

In light of this result, the avoided crossing between the fictitious and the real potential located at r_0 was pushed deeper into the repulsive region than the value shown in Figure 3. Then, in order to achieve an eventual acceleration for the entire fluid (i.e., for all degrees of freedom) at the same time, the fictitious potential was chosen to be a free particlelike one as shown in panel a of Figure 4; in other words, the interaction is such that if two particles come closer than the critical distance r_0, then their pairwise potential gradually flattened to a constant, while other interactions with the rest of the fluid remain unchanged. This choice exactly recovers the original radial distribution function as shown by the dotted line in panel b of Figure 4, and it was verified to be independent of system size, total simulation time, and choice of time step. When the distance between the particles is $\mathbf{r} < r_0$, the accelerated fluid contains an additional amount of kinetic energy as compared with the ordinary LJ fluid. This amount is the difference between the potential cut off and the original potential at $\mathbf{r}$, and it is the origin of the resulting acceleration. However, this kinetic surplus is quickly redistributed between all degrees of freedom via the thermostat, bath interactions, or particle collisions. For the choice of potential shown in dashed line on panel a of Figure 4, an increase in the diffusion constant of the order of 25–30% was achieved.

A rough evaluation of the area under the radial distribution function for $\mathbf{r} < r_0$ in Figure 4 gives the number of particles in this region after being correctly normalized for spherical symmetry. This number is never larger than 10% of the total number of particles, and typical values for $\langle e^{-\beta\Delta V}\rangle$ (the denominator of eq 8) were never being observed to be smaller than 0.98. Interestingly, the velocity autocorrelation function (not reported here) reaches the asymptotic zero value plateau in a smoother way and in half of the time with respect to the original one, suggesting that the dynamics is more

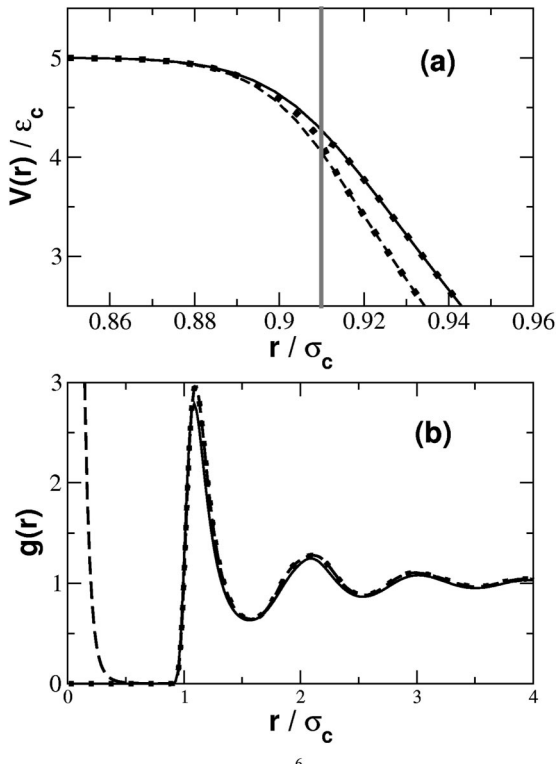

Figure 5. Potential and radial distribution function of a charged electroneutral Lennard-Jones fluid. (a) Gray vertical line indicating the cutoff radius and delimiting the accelerating region to the left; SSMD potential for attractive and repulsive interactions in the dashed and solid line, respectively; the same potential as before but switching off the charge interactions for $r_{ij} < r_0$. Note the overlap between the repulsive and attractive curves in this case once r_{ij} is smaller that r_0. (b) Same plot as panel b of Figure 4.

diffusive and less inertial. This is shown also in panel a of Figure 7, where an example of the mean square displacement is reported. The advantage of the present method is that the key parameter involves varying the value of the fictitious potential cutoff height. The disadvantages of this approach are related to the fact that the particles will remain in the $\mathbf{r} < r_0$ region if the potential cutoff is too low, which could alter the radial distribution function and decrease the diffusion constant.

IV. Implementation for Coulombic Interactions

The next level of complexity explored in this work was to include long-range electrostatics into the underlying model. Point charges were added to the previously described Lennard-Jones fluid such that the whole system was electroneutral. The total potential acting on each pair of particles was then given by the sum of a Coulombic and a Lennard-Jones interaction. A standard switching function (which modulates the transition from a given asymptotic value to zero) was used for cutting off the long-range Coulomb interaction. A direct application of eq 10 gives the attractive (dashed) and repulsive (solid) SSMD potentials. The details of these at the crossing region are shown in panel a of Figure 5. The dynamics for this resulted in pairs of particles

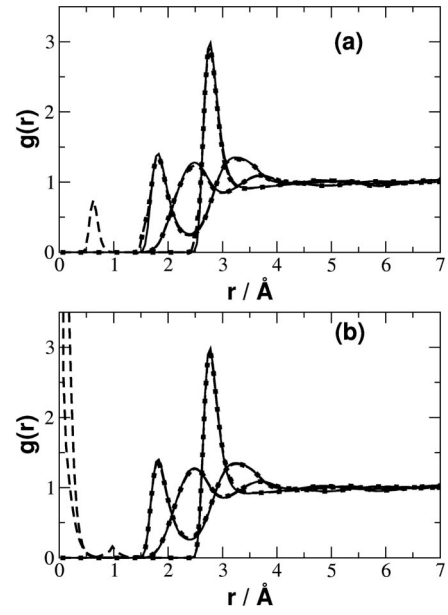

Figure 6. Radial distribution functions for the oxygen−oxygen interaction (curve with the highest peak), oxygen−hydrogen interaction (curve closer in), and hydrogen−hydrogen interaction at $T = 298$ K for a simulation of 256 water molecules and 250 ps time length. (a and b) Original distributions in solid lines; SSMD distributions in dashed lines; and reweighted distributions in the dotted line. (a) Pairwise SSMD procedure for each atoms at a time (see main text). (b) Fictitious potential switched on for the entire molecular pairwise interactions whenever a pair of atoms of these molecules are in a distance less than r_0.

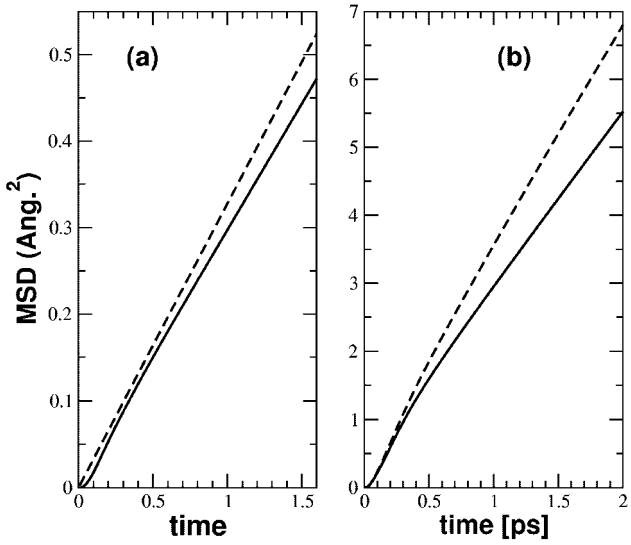

Figure 7. Mean square displacements for a Lennard-Jones fluid (a) and for liquid water (b). Accelerated results are shown in dashed lines. Notice that for the LJ fluid the inertial transition is reduced upon acceleration. In this plot, the self-diffusion coefficient shows a percentage increment of 20% and 25%, respectively, for the LJ fluid and for water.

remaining bound within the $\mathbf{r} < r_0$ region. Oppositely charged particles were strongly attracted to each other. Thus, by switching off their charges once their distance is $\mathbf{r} < r_0$, their pair interaction is the same as in the standard LJ fluid. This last choice is represented by the dotted potential in the

same figure where two dotted lines for repulsive and attractive interactions are seen for $\mathbf{r} > r_0$, and one dotted line for $\mathbf{r}$ smaller than the critical distance. The corresponding radial distribution functions, displayed on the bottom panel, are once again fully recovered and with the result that the fluid dynamics is accelerated.

The next system studied selected involved 256 water molecules at $T = 298$ K in a cubic box with periodic boundary conditions and density equal to 1000 kg/m^3. The potentials were modeled according to the flexible TIP3P force field.[66] A typical simulation length ranged from 150 to 250 ps. SSMD was implemented in DL_POLY 2.14[67] and the time-step was equal to 0.5 fs. As also noted for the LJ fluid, the acceleration increment does not change with time-step size and total simulation time (as long as the simulation results are converged). The long-range Coulombic interactions are calculated employing both the reaction field method (with a dielectric constant equal to 78) and the Ewald summation method. The explicit implementation of this last method for SSMD is reported in the Appendix.

Two acceleration approaches were employed: an atomic and a molecular approach. In the atomic approach, each atom within the water molecule was subjected to the acceleration cutoff. In the molecular approach, the entire water molecule was treated as a single site (analogous to the previous LJ simulations).

In the first approach, each intermolecular pairwise potential was cut off in the same way as was done for the Lennard-Jones fluid example both for attraction and repulsion. In this case, hydrogen−hydrogen interactions take most advantage of the potential cut off, simply because of the smaller van der Waals radius of hydrogen−hydrogen interaction. Nevertheless, once the hydrogen atoms were closer than $\mathbf{r} < r_0$, the oxygen atoms were still out of the $\mathbf{r} < r_0$ region and thus their strong repulsive interaction pulled the water molecules back. The effect of this acceleration scheme is shown by the artificial shells in $g(\mathbf{r})$ (dashed line of Figure 6a). The reweighted SSMD curve correctly removes it in order to recover the original radial distribution function, while all other pair interactions were left unchanged as expected by this kind of acceleration scheme. Even if the system acquired new phase space regions represented by the artificial shell in Figure 6a, the overall acceleration improvement was found to be only a few percent (i.e., the diffusion constant increased by about 5%).

However, it is possible to achieve an improvement with the second type of acceleration where a molecular-based cutoff scheme is employed. Whenever atoms of two different molecules come closer than a certain cut off radius r_0, then all pair interactions between each component of the two molecules can be thought to smooth out to a free particle potential; the result is shown by the radial distribution function in panel b of Figure 6. The correct equilibrium distribution function is recovered and the diffusion constant is increased by 33%. An example of the mean square displacement for the accelerated dynamics is shown in panel b of Figure 7 (dashed line). This plot suggests that the reduced number of ballistic collisions is mainly responsible for the acceleration. This result is promising if it is taken into consideration that the method has been applied to all degrees of freedom without distinction between them of any sort and that no specific rare events, such that ones that arise in a double well picture, are targeted. In fact, the fluids described above are in thermodynamic equilibrium, as opposed to a multidimensional double well where a major boost in probability transitions between the two basins can be obtained only after selecting a subset of coordinates.

Between the two schemes, the first atom-based approach is easier to implement, and this could play a role when more complex (e.g., biomolecular) systems are considered. The second scheme embodies the original idea of creating a fictitious phase space region which allows entire molecules to funnel through new configurations and, as a consequence, gain an entropic boost.

V. Concluding Remarks

Standard accelerated molecular dynamics via non-Boltzmann sampling has been shown to be problematic for condensed phase fluid systems, which do not adhere to a transition state theory (TST) picture along one or a few chosen coordinates. When applied indiscriminately to all fluid degrees of freedom, standard non-Boltzmann sampling results are poor, as shown on Figure 1. Here, an alternative approach is proposed based on the insertion of extended states into the interactions with the goal to enhance phase space sampling. The present methodology is the natural extension to complex systems of SSMD (which was previously presented for a one-dimensional example[45]). This approach, which in the present case couples the original potential to a fictitious free particle one, has been proven to recover the exact statistics via the non-Boltzmann sampling. It can be applied to selected types of pairwise interactions if one thinks these to be more relevant than others or equally to all degrees of freedom. In this case, the SSMD approach has be shown to achieve acceleration for both a standard LJ fluid and one with Coulombic interactions added. These systems are considered as good benchmarks for a condensed phase acceleration technique because their physical properties are well-known and any anomalies in the SSMD implementation can be easily recognized.

A resulting acceleration of 33% is also achieved for water after applying the SSMD implementation to all of the degrees of freedom. This can be considered a promising result if one takes into consideration that, in hyperdynamics,[29,36] a boost factor of 1400 for one degree of freedom reduces already to a factor of 2 if applied to a 20-dimensional system[68] and that the present SSMD implementation can recover the original statistical properties via non-Boltzmann sampling.

Acknowledgment. This research was supported by the Office of Naval Research (N00014-05-1-0457) and the National Science Foundation (CHE-0317132 and CHE-0719522). We thank Drs. Yeshitila Gebremichael, Vinod Krishna, and Satoru Iuchi for helpful discussions and comments.

Appendix: SSMD for Ewald Summations

The Ewald method[1,69] estimates Coulombic interactions for an electrically neutral system of charged point particles under periodic boundary conditions in a given volume L^3, where L is the length of the cubic simulation cell. These interactions are given by the sum of three contributions: the first is the electrostatic potential generated by the periodic sum of charges with density distributed as Gaussians

$$V_g(r_{ij}) = \frac{2\pi}{L^3} \sum_{(k\neq 0)} \frac{e^{-k^2/4\alpha}}{k^2} \sum_{i,j} \frac{q_i q_j}{4\pi\varepsilon_d} \cos(\mathbf{k}\cdot r_{ij}) \quad \text{(A1)}$$

where α is the Gaussian exponential coefficient of the charges distributions, ε_d the dieletric constant of the medium, k is the reciprocal space vector, q_i and q_j are the point charges at respective locations r_i and r_j, and $r_{ij} = |r_i - r_j|$. The second term is given by the correction of the spurious self-interactions included into the Fourier sum

$$V_{sl}(r_{ij}) = -\sum_j \frac{q_j}{4\pi\varepsilon_d} \sqrt{\frac{\alpha}{\pi}} \quad \text{(A2)}$$

and finally, the third one is the short-range interaction due to the point charges screened by oppositely charged Gaussians,

$$V_{sr}(r_{ij}) = \frac{1}{2} \sum_{(i\neq j)} \frac{q_i q_j}{4\pi\varepsilon_d} \frac{1 - \mathrm{erf}(\sqrt{\alpha} r_{ij})}{r_{ij}} \quad \text{(A3)}$$

where erf stands for "error function" and the cell summation has been taken out from all the above potentials.

In order to smooth the hard core potential, both the repulsive and attractive ones, the Fourier contribution has to be compensated by summing the exact opposite amount via real space. For this reason, eq A3 is rewritten as follows

$$V_{sr}(r_{ij}) = \frac{1}{2} \sum_{(i\neq j)} \frac{q_i q_j}{4\pi\varepsilon_d} \frac{f(r_{ij}) - \mathrm{erf}(\sqrt{\alpha} r_{ij})}{r_{ij}} \quad \text{(A4)}$$

where, given a cut off radius r_0, $f(r_{ij})$ interpolates between 1, when $r_{ij} > r_0$, and r_{ij}/r_0 when $r_{ij} < r_0$, in a SSMD fashion. On doing so, the resulting short-range potential in the forbidden region is

$$V_{sr}(r_{ij}) = \frac{1}{2} \sum_{(i\neq j)} \frac{q_i q_j}{4\pi\varepsilon_d} \left[\left(\frac{1}{r_0} - \frac{1}{r_{ij}} \right) + \frac{\mathrm{erfc}(\sqrt{\alpha} r_{ij})}{r_{ij}} \right] \quad \text{(A5)}$$

where erfc is the complementary error function. The forces used can be calculated by performing the first derivatives of each potential piece. This implementation allows one to apply SSMD to the Ewald summation approach and exactly corrects the short-range potential such that, after the summation over all components, it brings the forces to be equal to zero for $r_{ij} < r_0$.

References

(1) Frenkel, D.; Smit, B. *Understanding Molecular simulation: From Algorithms to Applications*, 2nd ed.; Academic: Boston, 2002.

(2) Allen, M. P.; Tildsley, D. J. *Computer Simulations of Liquids*; Oxford University Press: New York, 1987.

(3) Heermann, D. W. *Computer Simulation Methods in Theoretical Physics*; Springer: Berlin, 1986.

(4) Brunger, A.; Brooks, C. L., III.; Karplus, M. *Chem. Phys. Lett.* **1984**, *105*, 495.

(5) Schlick, T.; Barth, E.; Mandziuk, M. *Annu. Rev. Biophys. Biomol. Struct.* **1997**, *26*, 181.

(6) Schlick, T.; Skeel, R. D.; Brunger, A. T.; Kale, L. V.; Board, J. A.; Hermans, J.; Schulten, K. *J. Comput. Phys.* **1999**, *151*, 9.

(7) Nosé, S. *J. Chem. Phys.* **1984**, *81*, 511.

(8) Martyna, G. J.; Tuckerman, M. E.; Tobias, D. J.; Klein, M. L. *Mol. Phys.* **1996**, *87*, 1117.

(9) Hoover, W. G. *Phys. Rev. A* **1985**, *31*, 1695.

(10) Gawrisch, K. *The Dynamics of membrane Lipids*; CRC Press LLC: Boca Raton, FL, 2005.

(11) Daggett, V. *Curr. Opin. Struct. Biol.* **2000**, *10*, 160.

(12) Mayor, U.; Christopher, M. J.; Daggett, V.; Fersht, A. R. *Proc. Natl. Acad. Sci.* **2000**, *97*, 13518.

(13) Duan, Y.; Kollman, P. A. *Science* **1998**, *282*, 740.

(14) Shea, J.-E.; Brooks, C. L., III. *Annu. Rev. Phys. Chem.* **2001**, *52*, 499.

(15) Sheinerman, F. B.; Brooks, C. L., III. *Proc. Natl. Acad. Sci.* **1998**, *95*, 1562.

(16) Andricioaei, I.; Straub, J. E. *J. Chem. Phys.* **1997**, *107*, 9117.

(17) Andricioaei, I.; Straub, J. E. *J. Chem. Phys.* **2001**, *114*, 6994.

(18) Berg, B. A.; Neuhaus, T. *Phys. Lett. B* **1991**, *267*, 249.

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

(20) Marinari, E.; Parisi, G. *Europhys. Lett.* **1992**, *19*, 451.

(21) Lyubartsev, A. P.; M, A. A.; Sheykunov, S. V.; Vorontsov-Velyaminov, P. N. *J. Chem. Phys.* **1992**, *96*, 1776.

(22) Sugita, Y.; Okamoto, Y. *J. Phys. Soc. Jpn.* **1999**, *65*, 1604.

(23) Tesi, M.; E, J. J. v. R.; Orlandini, E.; Whittington, S. G. *J. Stat. Phys.* **1996**, *82*, 155.

(24) Hansmann, U. H. E. *Chem. Phys. Lett.* **1997**, *281*, 140.

(25) Hukushima, K.; Nemoto, K. *J. Phys. Soc. Jpn.* **1996**, *65*, 1604.

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

(27) Jang, S.; S, S.; Pak, Y. *Phys. Rev. Lett.* **2003**, *91*, 058305.

(28) Affentranger, R.; I, T.; Di Iorio, E. E. *J. Chem. Theory Comput.* **2006**, *2*, 217.

(29) Montalenti, F.; Voter, A. F. *J. Chem. Phys.* **2002**, *116*, 4819.

(30) Voter, A. F. *Phys. Rev. B* **1998**, *57*, 13985.

(31) Zagrovic, B.; E, J. s.; Pande, V. *J. Mol. Biol.* **2001**, *313*, 151.

(32) Uberuaga, B. P.; Anghel, M.; Voter, A. F. *J. Chem. Phys.* **2004**, *120*, 6363.

(33) Sorensen, M. R.; Voter, A. F. *J. Chem. Phys.* **2000**, *112*, 9599.

(34) Miron, R. A.; Fichthorn, K. A. *J. Chem. Phys.* **2003**, *119*, 6210.

(35) Voter, A. F. *Phys. Rev. Lett.* **1997**, *78*, 3908.

(36) Voter, A. F. *J. Chem. Phys.* **1997**, *106*, 4665.

(37) Corcelli, S. A.; Rahman, J. A.; Tully, J. C. *J. Chem. Phys.* **2003**, *118*, 1085.

(38) Hamelberg, D.; Mongan, J.; McCammon, J. A. *Protein Sci.* **2004**, *13*, 76.

(39) Hamelberg, D.; T, S.; McCammon, J. A. *J. Chem. Phys.* **2005**, *122*, 241103.

(40) Steiner, M. M.; Genilloud, P.-A.; Wilkins, J. W. *Phys. Rev. B* **1998**, *57*, 10236.

(41) Rahman, J. A.; Tully, J. C. *J. Chem. Phys.* **2002**, *116*, 8750.

(42) Madun, R. A.; Fichthorn, K. A. *Phys. Rev. Lett.* **2004**, *93*, 128301.

(43) Xing, C.; Andricioaei, I. *J. Chem. Phys.* **2006**, *124*, 034110.

(44) Minth, D. D. L.; Hamelberg, D.; McCammon, J. A. *J. Chem. Phys.* **2007**, *127*, 154105.

(45) Venkatnathan, A.; Voth, G. A. *J. Chem. Theory Comput.* **2004**, *1*, 36.

(46) Grubmuller, H. *Phys. Rev. E* **1995**, *52*, 2893.

(47) Hamelberg, D.; J, M.; McCammon, J. A. *J. Chem. Phys.* **2004**, *120*, 11919.

(48) MacFadyen, J.; Andricioaei, I. *J. Chem. Phys.* **2005**, *123*, 074107.

(49) Mella, M. *J. Chem. Phys.* **2005**, *122*, 204106.

(50) Hamelberg, D.; de Oliveira, C. A. F.; McCammon, J. A. *J. Chem. Phys.* **2007**, *127*, 155102.

(51) de Oliveira, C. A. F.; Hamelberg, D.; McCammon, J. A. *J. Phys. Chem. B* **2006**, *110*, 22695.

(52) de Oliveira, C. A. F.; Hamelberg, D.; McCammon, J. A. *J. Chem. Phys.* **2007**, *127*, 175105.

(53) Dellago, C.; Bolhius, P. G.; Chandler, D. *J. Chem. Phys.* **1998**, *108*, 9236.

(54) Bolhuis, P. G.; Dellago, C.; Geissler, P. L.; Chandler, D. *J. Phys.: Condens. Matter* **2000**, *12*, A147.

(55) Bolhuis, P. G.; Chandler, D.; Dellago, C.; Geissler, P. L. *Annu. Rev. Phys. Chem.* **2002**, *53*, 291.

(56) Prasad, M.; Sinno, T. *J. Chem. Phys.* **2004**, *121*, 8699.

(57) Wu, X.; Brooks, B. R. *Chem. Phys. Lett.* **2003**, *381*, 512.

(58) Wu, X.; Wang, S. *J. Phys. Chem. B* **1998**, *102*, 7238.

(59) Wu, X.; Wang, S. *J. Chem. Phys.* **1999**, *110*, 9401.

(60) Andricioaei, I.; Dinner, A. R.; Karplus, M. *J. Chem. Phys.* **2003**, *118*, 1074.

(61) Bitetti-Putzer, R.; Dinner, A. R.; Yang, W.; Karplus, M. *J. Chem. Phys.* **2006**, *124*, 174901.

(62) Bussi, G.; Laio, A.; Parrinello, M. *Phys. Rev. Lett.* **2006**, *96*, 090601.

(63) Laio, A.; Parrinello, M. *Proc. Natl. Acad. Sci.* **2002**, *99*, 12562.

(64) Micheletti, C.; Laio, A.; Parrinello, M. *Phys. Rev. Lett.* **2004**, *92*, 170601.

(65) Oganov, A. R.; Martonak, R.; Laio, A.; Raiteri, R.; Parrinello, M. *Nature* **2005**, *438*, 1142.

(66) Durell, S. P.; Brooks, B. R.; Ben-Naim, A. *J. Phys. Chem.* **1994**, *98*, 2198.

(67) Smih, W.; Forester, T. *J. Mol. Graphics* **1996**, *14*, 136.

(68) Henkelman, G.; Jonsson, H. *J. Chem. Phys.* **2001**, *115*, 9657.

(69) Ewald, P. P. *Ann. Phys.* **1921**, *64*, 253.

CT7003275