LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

UCRL-JRNL-208389

Isobaric-isothermal Monte Carlo simulations from first principles: Application to liquid water at ambient conditions

M. McGrath, J. I. Siepmann, I-F. W. Kuo, C. J. Mundy, J. VandeVondele, J. Hutter, F. Mohamed, M. Krack

December 3, 2004

ChemPhysChem

Isobaric-isothermal Monte Carlo simulations from first principles: Application to liquid water at ambient conditions

Matthew J. McGrath,[a] J. Ilja Siepmann,*,[a]

I-Feng W. Kuo,[b] Christopher J. Mundy,[b]

Joost VandeVondele,[c] Jürg Hutter,[d]

Fawzi Mohamed,[e] and Matthias Krack,[e]

November 30, 2004

[**a**] Departments of Chemistry and of Chemical Engineering and Material Science, University of Minnesota, 207 Pleasant Street SE, Minneapolis, MN 55455-0431, U.S.A.

E-mail: siepmann@chem.umn.edu

[**b**] Chemistry and Materials Science Directorate, Lawrence Livermore National Laboratory, Livermore, CA 94550, U.S.A.

[**c**] Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, United Kingdom

[**d**] Physical Chemistry Institute, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland

[**e**] Laboratory of Physical Chemistry, ETH Zurich, USI-Campus, Via Giuseppe Buffi 13, CH-6904 Lugano, Switzerland

Abstract

A series of first principles Monte Carlo simulations in the isobaric-isothermal ensemble were carried out for liquid water at ambient conditions ($T = 298$ K and $p = 1$ atm). The Becke-Lee-Yang-Parr (BLYP) exchange and correlation energy functionals and norm-conserving Goedecker-Teter-Hutter (GTH) pseudopotentials were employed with the `CP2K` simulation package to examine systems consisting of 64 water molecules. The fluctuations in the system volume encountered in simulations in the isobaric-isothermal ensemble requires a reconsideration of the suitability of the typical charge density cutoff and the regular grid generation method previously used for the computation of the electrostatic energy in first principles simulations in the microcanonical or canonical ensembles. In particular, it is noted that a much higher cutoff is needed and that the most computationally efficient method of creating grids can result in poor simulations. Analysis of the simulation trajectories using a very large charge density cutoff at 1200 Ry and four different grid generation methods point to a substantially underestimated liquid density of about 0.85 g/cm^3 resulting in a somewhat understructured liquid (with a value of about 2.7 for the height of the first peak in the oxygen–oxygen radial distribution function) for BLYP-GTH water at ambient conditions.

Keywords: Isobaric-isothermal Monte Carlo simulation, density functional theory, water

1 Introduction

The scientific community's interest in water was born out of necessity. Water is the most prevalent liquid on Earth, covering approximately 70% of its surface. Water also comprises about 60% of the mass of an adult human. Consequently, many reactions important to life occur in aqueous solution, and in order to understand these processes, one must understand how water affects them. Compared to most other liquids, however, water shows many unique thermophysical properties (e.g., a temperature of maximum density at ambient pressure) related to its tetrahedral structure and the ability to act as donor *and* acceptor for two hydrogen bonds. This fact has caused the understanding of water's properties to become a grand challenge for liquid state theory [1] and molecular simulation [2]. Beginning with the first particle-based simulations of water using pairwise empirical potentials [3], much effort has been devoted to development evermore sophisticated empirical water models (e.g., [4, 5, 6, 7, 8]) that have allowed for a molecular-level understanding of some of the unusual thermophysical properties of water [9, 10, 11, 12]. Because empirical models struggle to provide an accurate description of water's complex electronic properties that greatly influence its solubility characteristics and lead to chemical events such as self-dissociation, *ab initio* methods have been turned to for the past decade to shed light on water's physical and chemical properties [13, 14, 15].

The ability to perform first principles simulations of water arrived with the introduction of the Car-Parrinello molecular dynamics (CPMD) method [16] which allows for a quantum-mechanical description for the system of interest

to be used in a molecular dynamics simulation by propagating the electronic density with a fictitious Lagrangian. The first of these simulations used 32 molecules in the microcanonical ensemble and was run for a total of ∼3.5 ps [13]. Currently *ab initio* simulations for liquid water are being extended to more than 10 ps for 64 or 128 molecule systems [17, 18, 19, 20, 21, 22], but all of these simulations were carried out in either the microcanonical or the canonical ensemble using the experimental liquid density to constrain the volume. The notable exception is a recent simulation of a slab of liquid water surrounded by a vapor region in the canonical ensemble (T = 300 K) [15]. During this simulation, the volume of a slab consisting of 216 molecules was observed to expand by about 10% compared to its initial volume with a specific density of 1.0 g/cm^3 [15].

If one is interested in true predictions of fluid properties, however, then one cannot rely on experimental knowledge of the liquid density but must allow the simulated system the freedom to find the density that minimizes its Gibbs free energy with respect to the thermodynamic constraints of temperature and pressure, that is the simulation needs to be carried out in the isobaric–isothermal ensemble [23]. In the isobaric-isothermal ensemble the volume of the simulation box is allowed to fluctuate in contact with an external pressure bath, and this ensemble has been used for quite some time in simulations of molecular fluids using empirical potentials [24]. Constant-pressure and constant-stress ensembles have also found use in *ab initio* simulations, but mostly for solids or liquids at high external pressure [25, 26, 27, 28].

The goal of the present research is to extend the use of the isobaric-isothermal ensemble to first principles Monte Carlo simulations at ambient pressure. For this purpose, a Monte Carlo module was created for CP2K [29, 30, 31], a multi-purpose program that uses a grid-based methods to integrate the exchange and correlation energy functionals and to solve Poisson's equation. Before carrying out any simulations in the isobaric-isothermal ensemble, the effect of the charge density cutoff for the auxilliary plane wave expansion on the potential energy of a few selected configurations was addressed by examining this energy as a function of the volume of the simulation cell. Once a suitable value for the cutoff was found, different methods for generating the regular grid employed in the fast Fourier transforms (FFT) used to solve the density functional Hamiltonian were explored. The next section describes the technical details for the isobaric-isothermal Monte Carlo simulations, the generation of the FFT grids, and the computation of potential energy versus volume ($U-V$) curves. Section 3 provides a discussion of the results, followed by a brief Conclusion.

2 Computational Methods

The first principle simulations of this work were performed using the computer program `CP2K`, which is developed by a large group of researchers and publicly available [29]. The energy routines of this code, called `QUICKSTEP`, utilize an atom-centered Gaussian basis set to linearly expand the Kohn-Sham (KS) orbitals [32] and a plane wave auxiliary basis set to linearly expand the electronic

density. This method is known as the Gaussian plane wave (GPW) method [30]. The CP2K code differs from its predecessor, the highly successful CPMD package [33], which uses plane waves to expand both the orbitals and the electronic density [16]. The use of Gaussian functions to represent the KS orbitals is more chemically intuitive, while the use of plane waves takes full advantage of the enforced periodicity of simulations in condensed phases.

One module of the CP2K program is known as CP2K-MC. As the name suggests, this module uses Monte Carlo algorithms to sample the configurational space while utilizing the full efficiency and power of the energy routines described above. One benefit of using Monte Carlo techniques is the explicit inclusion of the thermodynamic constraints (e.g., temperature and pressure) in the acceptance rules for the different types of trial moves [2, 34, 35]. CP2K-MC simulations of water in the NpT ensemble require the use of four distinct types of trial moves: (i) translations of rigid molecules, (ii) rigid-body rotations around the center of mass, (iii) conformational moves altering either the bond length or bond angle, and (iv) changes in the volume of the simulation box. In the present implementation, the random displacements applied to the volume are randomly selected uniform in the volume coordinate (L^3, where L is the box length). During a volume move, the center of mass position of every molecule is kept fixed in scaled coordinates, that is these center-of-mass coordinates are displaced by $L_{\mathrm{trial}}/L_{\mathrm{current}}$, where L_{trial} and L_{current} are the trial and current box lengths, respectively, in the regular Cartesian coordinate system [24].

A fifth type of move is used here to increase computational efficiency by

reducing the number of expensive *ab initio* energy calculations. This is referred to as a pre-sampling of trajectories [36, 37] and involves performing a short sequence of moves with a inexpensive classical potential [38]. This entire sequence is accepted based on the difference in the classical and *ab initio* energies at the beginning and end of the move sequence. In the present implementation, this sequence consists of eight moves of type (i) to (iii), which was determined to greatly improve the efficiency (accepted classical moves per `QUICKSTEP` calculation) under similar conditions [22]. The choice of move is done at random and helps the simulation achieve both thermal and mechanical equilibrium while guaranteeing equipartition between the various degrees of freedom. The maximum displacements of moves of type (i) to (iii) were adjusted to yield an acceptance probability of about 50% for the individual inner moves using the classical potential. The maximum displacement of the volume move was adjusted to give a 50% acceptance rate for the *ab initio* energy calculation, as no pre-sampling was done for volume moves. The length of these *ab initio* Monte Carlo simulations is given in terms of Monte Carlo cycles where one cycle involves N `QUICKSTEP` energy calculations (where N is the number of molecules in the system), i.e. a larger number of displacements of types (i) to (iii) is used during a cycle because of the use of the pre-sampling technique.

In a recent paper by this group [22], the first principles Monte Carlo approach for sampling from the canonical partition function was validated through extensive comparison with molecular dynamics simulations in the microcanonical and canonical ensembles using the programs `CPMD` and `CP2K`. This paper explored the

structural properties of water at a fixed density of 1.0 g/cm^3 and temperatures close to 315 K [22]. For these constant-density simulations, the electronic structure calculations used the Becke-Lee-Yang-Parr (BLYP) exchange and correlation energy functionals with the generalized gradient approximation [39, 40], the norm-conserving pseudopotentials of Goedecker and co-workers (GTH) [41, 42] to mimic the core electrons, and a triple-zeta valence basis set augmented with two sets of d-type or p-type polarization functions (TZV2P) for the KS orbitals on O and H. An extensive test of different basis sets was carried out recently [43], and the TZV2P basis set yields converged structural and dynamic properties for liquid water at constant volume. Consequently, it was decided that the present simulations in the isobaric-isothermal ensemble should use the same functional, pseudopotentials, and basis sets to further examine the properties of water described by this model. Although a charge density cutoff at 280 Ry for the auxiliary plane wave expansion was found to be sufficient to yield well converged energy differences (or forces) in the canonical ensemble, initial tests of energy versus volume curves (see next section) indicated that a much higher value of the cutoff is needed in simulations with fluctuating volume. For this reason, a charge density cutoff at 1200 Ry was used for the simulations in the isobaric-isothermal ensemble. Furthermore, the density derivatives for the exchange-correlation gradient correction were calculated using a quadratic spline interpolation of the density at the grid points [31].

Three different methods were tested to generate the regular grids used in the computation of the electronic density. The full theoretical basis and technical

details of the GPW method implemented in the QUICKSTEP module have been described elsewhere [30, 31] and shall not be reproduced here. In the isobaric-isothermal ensemble, the generation of the grid used to compute the Coulomb and exchange/correlation energies requires special care. When the volume of the simulation supercell changes, either the number of grid points can be kept constant or the density of the grid points can be maintained as closely as possible. Both approaches were tried in this paper, with the former implemented through the use of a reference cell (i.e., the simulation cell has the same number of grid points as a cell of the reference size would have at the default grid density used in QUICKSTEP). As the number of grid points is neccessarily discrete, but the volume itself is a continuous variable, a constant density of grid points can only approximately be enforced. The jumps in the number of grid points per unit length can be significant, if the number of grid points is chosen for optimal computational efficiency. Because of the intricacies of FFTs, this is achieved with certain grid sizes that can be factored as a product of many small prime numbers (known as prime lengths) [44]. In addition to a simulation using only prime lengths, another simulation was also carried out that tried to minimize the jumps in the number of grid points by allowing all grid sizes, which in turn increased the computational expense by about 15%.

The four simulations described here were carried out for systems consisting of 64 water molecules and the thermodynamic constraints were set to ambient condition, i.e., $T = 298$ K and $p = 1$ atm. Run PRIME-GRID, which used only prime length FFTs, was started from the final configuration of a CP2K-MC

run in the canonical ensemble at $T = 315$ K and $\rho = 1.0$ g/cm^3 (a supercell of $L = 12.42$ Å) [22]. This simulation was run for about 300 Monte Carlo cycles before the decision was made not to continue it because the use of PRIME-GRID was found to affect the volume fluctuations (see below). The other three runs were all started from the final configuration of run PRIME-GRID. At present, these three runs have only been carried out for about 180 Monte Carlo cycles; a length that unfortunately is not sufficient to yield well equilibrated samples and falls short of the 700 cycles (200 for equilibration and 500 for production) used for the simulations in the canonical ensemble [22]. One of these runs (ALL-GRID) allowed the program to select any size of regular grid (not just the prime lengths). The other two used the reference cell technique mentioned above to keep the number of grid points constant. SMALL-REF and LARGE-REF used reference cells corresponding to a liquid density of 1.0 and 0.9 g/cm^3, respectively. It should be noted that a lower specific density for the reference cell results in a larger grid density for the simulation cell.

The calculation of average properties and the structural analysis for the four simulations were carried out using the final 100 cycles of each run. The radial distribution functions (RDFs) were calculated using a bin width of 0.005 Å for separations smaller than 1.2 Å (only including the intramolecular oxygen–hydrogen bond) and a bin width of 0.02 Å for larger separations. The potential energy versus volume ($U - V$) curves were obtained by taking a configuration of a previously reported `CP2K`-MC run [22] and scanning the volume in increments of 5 Å^3, until an accurate determination of the potential energy minimum could

be made. The volume changes were performed exactly like the volume changes in the NpT simulations (i.e., only scaling the center of mass coordinates in the Cartesian frame). A single configuration was used to test for the dependence on the charge density cut off, while the energies for ten different configurations taken from a long `CP2K`-MC simulation in the canonical ensemble [22] were averaged for an initial assessment of the different grid generation techniques.

The `CP2K`-MC simulations were performed on `Thunder`, a supercomputer with Intel Itanium2 Tiger4 nodes running at 1.4 GHz [45]. Because of the relatively small system size of 64 water molecules, optimal scaling is not achieved beyond 64 and 96 processors for charge density cutoffs of 280 and 1200 Ry, respectively. For the former cutoff, a single self-consistent field (SCF) step for the electronic structure calculation requires approximately 10 s, whereas an SCF step takes about 13 s for the larger cutoff. The increase in expense for quadrupling the cutoff is only modest, thus showing the relatively small overhead associated with the auxilliary basis set in `QUICKSTEP` energy calculations.

3 Results and Discussion

3.1 Potential Energy versus Volume Curves

Figure 1 shows the affect of varying the charge density cutoff on the potential energy of the system as a function of volume. More importantly, it demonstrates a well-known peril of using grid-based methods for the computation of $U-V$ curves or for simulations in ensembles with fluctuating volume [46]. When

the grid density is kept approximately constant, an artificial energy jump is introduced between two neighboring volume values that involve a change in the total number of grid points. This situation is exaccerbated by the fact that `QUICKSTEP` performs the integration of the exchange-correlation term on the same regular grid used for solving Poisson's equation, and the choice of pseudopotential also plays a role [31]. For comparison, a $U - V$ curve calculated for the Pade functional with local density approximation [47] using the same charge density cutoff does not exhibit jumps of similar magnitude, thus pointing to the calculation of the density gradient for the GGA functional as being a major contributor to the energy jumps observed here. Clearly, large jumps would lead to a skewed volume sampling in isobaric-isothermal Monte Carlo simulations (and their discontinuity would cause additional numerical problems in molecular dynamics simulations). One can see that increasing the charge density cutoff decreases the magnitude of these jumps. The largest jump observed for a cutoff at 280 Ry is about 0.8 kJ/mol, while the largest jump for 1200 Ry is only about 0.2 kJ/mol, that is about an order of magnitude smaller than $k_{\mathrm{B}}T$, the thermal energy, at room temperature. Increasing the cutoff clearly moves the absolute energies closer toward the converged limit. Less noticeable in these $U-V$ curves is that the minimum shifts from a volume of 2040 Å^3 at 280 Ry to 2080 Å^3 at 1200 Ry. For these reasons, the cutoff for the auxiliary plane wave expansion of the electronic density was chosen to be 1200 Ry for the simulations reported in this paper.

Once the charge density cutoff was selected, efforts were made to determine

the best method to generate the regular grids. Figure 2 depicts $U-V$ curves calculated for PRIME-GRID, ALL-GRID, SMALL-REF, and LARGE-REF (using a reference cell corresponding to a specific density of 0.83 g/cm^3). The $U-V$ curve for PRIME-GRID looks very similar to the curve shown in Figure 1, demonstrating that averaging over multiple independent configurations would not have altered the conclusion that a very large cutoff is required for a density functional based simulation with fluctuating volume. The curve for ALL-GRID shows a marked decrease in the magnitude of the energy jumps (by about a factor of four) compared to PRIME-GRID, though the frequency of the jumps is increased for the former. As expected, SMALL-REF and LARGE-REF do not exhibit any jumps, since the number of grid points remains constant with respect to volume. It should be noted that for Figure 2, a larger reference cell (specific density of 0.83 g/cm^3) was used for LARGE-REF than for the NpT simulation denoted as LARGE-REF (0.90 g/cm^3). This was done to illustrate that although a large change in the size of the reference cell apparently has minimal impact on the relative shape of the $U-V$ curve, it usually leads to a change in the absolute energy. In the case studied here, the larger number of grid points resulted in a lower absolute energy. However, since the grid is only an auxilliary basis in `QUICKSTEP`, a variational principles does not apply and the energy can also increase as the number of grid points increases [30, 31]. Indeed, this is evident at $V \approx 2175$ Å^3 in Figure 2, where the potential energy jumps down for the PRIME-GRID calculation, but up for ALL-GRID.

Finally, it should be noted that these $U-V$ curves averaged over config-

urations taken from a canonical simulation at $T = 315$ K and $\rho = 1.0$ g/cm^3 point to a minimum in the potential energy that is located at about 2050 Å^3, i.e. about 7% above the volume used for the canonical simulations.

3.2 Isobaric-isothermal Simulations

The instantaneous energies and densities for the PRIME-GRID, ALL-GRID, SMALL-REF, and LARGE-REF trajectories are shown in Figures 3 and 4. Judging solely from the potential energy, one would be excused in thinking that PRIME-GRID has equilibrated after about 130 Monte Carlo cycles and is now sampling from its appropriate equilibrium distribution. However, the instantaneous density of run PRIME-GRID tells a different story. From the latter graph it is clear that something is preventing the simulation from adequately sampling phase space. The density that appears to curtail the volume fluctuations is about 0.88 g/cm^3 (which corresponds to a supercell volume of 2175 Å^3) and is precisely where the jump in the $U - V$ curve occurs (see Figure 2). This is somewhat surprising, since the magnitude of the jump is less than 0.2 kJ/mol, and the thermal energy of the simulation at this temperature is, thus, an order of magnitude larger. Nonetheless, the results of PRIME-GRID cannot be expected to accurately reflect the true properties of the model Hamiltonian, and therefore a different method must be used.

The trajectories of the two simulations using the reference cell method for the grid generation exhibit similar behavior to each other, i.e. the potential energies are heading slowly upward and the specific densities are overall de-

creasing over these relatively short trajectories. This suggests that these runs may not yet have equilibrated. Nevertheless, averaged over the last 100 cycles of the trajectory, their potential energy is found to be about 3 kJ/mol above those calculated for PRIME-GRID and their density falls about 3% below that for PRIME-GRID. However, these differences are well within the statistical uncertainty for such short trajectories. On the positive side, the two runs using the reference cell method do not suffer from similar sampling problems as run PRIME-GRID, because the reference cell keeps the number of grid points constant. This eliminates the jumps associated with changing the number of points, but it also means that changing the simulation volume changes the grid density (and correspondingly the relative convergence of the electronic structure calculation), thus potentially influencing the outcome of the simulations.

Last but not least, run ALL-GRID constrains the grid density but does not restrict the grid sizes to prime lengths. This results in a loss of computational efficiency, but lessens the magnitude of the energy jump when adding/removing grid points while avoiding the potential problems associated with the reference cell. Most importantly, the ALL-GRID run does not suffer from detrimental sampling problems associated with energy jumps as found for PRIME-GRID.

One curious aspect of these simulations is that the ALL-GRID, SMALL-REF, and LARGE-REF runs move away from the density that is associated with the sampling barrier for run PRIME-GRID. One may have expected them to oscillate around the average value of PRIME-GRID, sampling space on both sides of the barrier. It should be noted that, in particular, during the early

part of these three simulations specific densities greater than 0.88 g/cm^3 are frequently sampled.

Overall, the four simulations for the BLYP-GTH representation of water point to an ambient liquid density of about 0.85 g/cm^3, i.e., substantially below the experimental value of 1.0 g/cm^3. Thus, these simulations support the significant volume expansion observed in an *ab initio* simulation of a water slab that yielded a liquid density of about 0.9 g/cm^3 [15]. Further calculations are needed to explore the reason(s) for the discrepancies in the densities obtained from isobaric-isothermal ensemble and liquid-slab simulations. Contributing factors might include the larger cut-off used here (see also Figure 1), the difference in pseudopotentials (GTH versus Troullier-Martins [48]), and the limited duration of the isobaric-isothermal and liquid-slab simulations that might not have yielded well equilibrated samples.

Recent papers investigating *ab initio* water have emphasized structural and transport properties [20, 21, 22]. Monte Carlo simulations lack a true concept of "time", and therefore transport properties are not calculated here. Structural properties can be explored, however. Figures 5 and 6 show the oxygen–oxygen and oxygen–hydrogen radial distribution functions (RDFs) of the four NpT simulations. The statistical noise is fairly high and similar to those observed over 100 cycles in the canonical ensemble. There is no indication that the RDFs calculated for the four simulations do not fall within each other's error bars. Furthermore, very good agreement is oberved for the RDFs computed from the isobaric-isothermal simulations and a previous *ab initio* simulation of a liquid

slab at $T = 300$ K [15, 49].

The average height of the first peak in the oxygen–oxygen RDF obtained from the isobaric-isothermal simulations is about 2.7. At ambient temperature and pressure, the corresponding peak height for the TIP4P-pol2 model [7], which provides an excellent match for the experimental X-ray diffraction spectrum [50, 51], is about 3.0. Together, the isobaric-isothermal and liquid-slab simulations point toward a conclusion that the BLYP representation of water leads to an understructured liquid when allowed to expand in contact with an external pressure bath at 1 atm (or a vapor phase). In contrast, recent simulations of liquid water using the BLYP functional in the microcanonical and canonical ensembles at $\rho = 1.0$ g/cm^3 [20, 22] found an overstructured liquid phase.

4 Conclusions

This work presents a first attempt to perform first principles simulations for liquid water in the isobaric-isothermal ensemble at ambient conditions. We are able to conclude that the charge density cutoffs for the evaluation of the electronic density on a regular grid must be substantially greater in ensembles with fluctuating volume to avoid large jumps in the potential energy curve that appear when grid points are introduced/removed. The magnitude of these energy jumps is more pronounced when the regular FFT grid is constrained to prime lengths and it appears that using all possible grid numbers, while computationally less efficient (by about 15%) offers significant benefits. These jumps can be

altogether eliminated by using a constant number of grid points for all volumes, but this requires an initial guess of the density of a reference cell. Use of a reference cell alters the convergence of the electronic structure calculation for different densities and also significantly affects the absolute energies, a problem when one tries to estimate cohesive energies. Overall, constraining the grid density while allowing it to access all possible grid sizes appears to be the best way to deal with these problems. It should be noted that the integration of the exchange-correlation energy on a grid is a serious problem for these simulations, and there is a need to explore other schemes that may allow for better accuracy but still maintain computational efficiency.

For the BLYP-GTH description of water using a TZV2P basis set and a charge density cutoff at 1200 Ry, the isobaric-isothermal Monte Carlo simulations at a temperature of 300 K and a pressure of 1 atm yield a less dense and less structured liquid than observed experimentally for the same conditions.

5 Acknowledgments

We thank Michele Parrinello, Michiel Sprik, Larry Fried and Charlie Westbrook for their ongoing support of this work. Financial support from the National Science Foundation (CTS-0138393), the Engineering and Physical Sciences Research Council, and a 3M Graduate Fellowship (M.J.M.) is gratefully acknowledged. M.J.M also gratefully acknowledges a Department of Energy Computational Science Graduate Fellowship provided under grant number DE-

FG02-97ER25308. Part of this work was performed under the auspices of the U.S. Department of Energy by the University of California Lawrence Livermore National Laboratory (LLNL) under contract No. W-7405-End-48. J.I.S. and M.J.M. thank the University Relations Program (LLNL) for hosting their sabbatical visit. Part of the computer resources were provided by the Minnesota Supercomputing Institute.

References

[1] J. P. Hansen, I. R. McDonald, *Theory of Simple Liquids*, 2nd ed., Academic Press, New York, **1986**.

[2] M. P. Allen, D. J. Tildesley, *Computer Simulation of Liquids*, Oxford University Press, Oxford, **1987**.

[3] J. P. Barker, R. O. Watts, *Chem. Phys. Lett.* **1969**, *3*, 144.

[4] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, *79*, 926.

[5] M. W. Mahoney, W. L. Jorgensen, *J. Chem. Phys.* **2000**, *112*, 8910.

[6] S. W. Rick, S. J. Stuart, B. J. Berne, *J. Chem. Phys.* **1994**, *101*, 6141.

[7] B. Chen, J. Xing, J. I. Siepmann, *J. Phys. Chem. B* **2000**, *104*, 2391.

[8] P. T. Cummings, private communication.

[9] A. Wallqvist, R. D. Mountain, *Rev. Comput. Chem.* **1999**, *13*, 183.

[10] T.M Truskett, K. A. Dill, *J. Phys. Chem. B* **2002**, *106*, 11829.

[11] J. R. Errington, P. G. Debenedetti, *Nature* **2001**, *409*, 318.

[12] H.E. Stanley, S. V. Buldyrev, N. Giovambattista, *Physica A* **2004**, *342*, 40.

[13] K. Laasonen, M. Sprik, M. Parrinello, R. Car, *J. Chem. Phys.* **1993**, *99*, 9080.

[14] P. L. Geissler, C. Dellago, D. Chandler, J. Hutter, M. Parrinello, *Science* **2001**, *291*, 2121.

[15] I.-F. W. Kuo, C. J. Mundy, *Science* **2004**, *303*, 658.

[16] R. Car, M. Parrinello, *Phys. Rev. Lett.* **1985**, *55*, 2471.

[17] P. L. Silvestrelli, M. Parrinello, *J. Chem. Phys.* **1999**, *111*, 3572.

[18] B. Chen, I. Ivanov, J. M. Park, M. Parrinello, M. L. Klein, *J. Phys. Chem. B* **2002**, *106*, 12006.

[19] S. Izvekov, G. A. Voth, *J. Chem. Phys.* **2002**, *116*, 10372.

[20] J. C. Grossman, E. Schwegler, E. W. Draeger, F. Gygi, G. Galli, *J. Chem. Phys.* **2004**, *120*, 300.

[21] E. Schwegler, J. C. Grossman, F. Gygi, G. Galli, *J. Chem. Phys.* **2004**, *121*, 5400.

[22] I.-F. W. Kuo, C. J. Mundy, M. J. McGrath, J. I. Siepmann, J. VandeVondele, M. Sprik, J. Hutter, B. Chen, M. L. Klein, F. Mohamed, M. Krack, M. Parrinello, *J. Phys. Chem. B* **2004**, *108*, 12990.

[23] D. A. McQuarrie, *Statistical Mechanics*, Harper and Row, New York, **1996**.

[24] I. R. McDonald, *Mol. Phys.* **1972**, *23*, 41.

[25] R. M. Wentzcovitch, J. L. Martins, G. D. Price, *Phys. Rev. Lett.* **1993**, *70*, 3947.

[26] P. Focher, G. L. Chiarotti, M. Bernasconi, E. Tosatti, M. Parrinello, *Europhys. Lett.* **1994**, *26*, 345.

[27] M. Bernasconi, G. L. Chiarotti, P. Focher, S. Scandolo, E. Tosatti, M. Parrinello, *J. Phys. Chem. Solids* **1995**, *56*, 501.

[28] C. Cavazzoni, G. L. Chiarotti, S. Scandolo, E. Tosatti, M. Bernasconi, M. Parrinello, *Science* **1999**, *283*, 44.

[29] `CP2K`, http://cp2k.berlios.de

[30] G. Lippert, J. Hutter, M. Parrinello, *Mol. Phys.* **1997**, *92*, 477.

[31] J. VandeVondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing, J. Hutter, *Comp. Phys. Comm.* **2004**, submitted for publication.

[32] W. Kohn, L. J. Sham, *Phys. Rev. A* **1965**, *140*, 1133.

[33] CPMD, Version 3.7; copyright IBM Corp., 1990-2003, copyright MPI für Festköperforschung Stuttgart, 1997-2001; www.cpmd.org.

[34] D. Frenkel, B Smit, *Understanding Molecular Simulation*, Academic Press, New York, **1996**.

[35] A. R. Leach, *Molecular Modelling: Principles and Applications*, 2nd Edition, Prentice Hall, New York, **2001**.

[36] R. Iftimie, D. Salahub, D. Wei, J. Schofield, *J. Chem. Phys.* **2000**, *113*, 4852.

[37] L. D. Gelb, *J. Phys. Chem.* **2003**, *118*, 7747.

[38] S. Izvekov, M. C. Parrinello, C. J. Burnham, G. A. Voth, *J. Chem. Phys.* **2004**, *120*, 10896.

[39] A. D. Becke, *Phys. Rev. A* **1988**, *38*, 3098.

[40] C. Lee, W. Yang, R. C. Parr, *Phys. Rev. B* **1988**, *37*, 785.

[41] S. Goedecker, M. Teter, J. Hutter, *Phys. Rev. B* **1996**, *54*, 1703.

[42] C. Hartwigsen, S. Goedecker, J. Hutter, *Phys. Rev. B* **1998**, *58*, 3641.

[43] J. VandeVondele, F. Mohamed, M. Krack, J. Hutter, M. Sprik, M. Parrinello, *J. Chem. Phys.* **2005**, *122*, in press.

[44] J. W. Cooley, J. W. Tukey, *Math. Comput.* **1965**, *19*, 297.

[45] www.top500.org (Nov. 2005).

[46] P. Gomes Dacosta, O. H. Nielsen, K. Kunc, *J. Phys. C: Solid State Phys.* **1986**, *19*, 3163. G.P. Francis, M.C. Payne, *J. Phys.: Condens. Matter* **1990**, *2*, 4395.

[47] D. M. Ceperley, B. J. Alder, *Phys. Rev. Lett.* **1980**, *45*, 566; R. G. Parr, W. Yang, *Density Functional Theory of Atoms and Molecules*, International

Series of Monographs on Chemistry, Volume 16, Oxford University Press, 1989.

[48] Troullier, N.; Martins, J. *Phys. Rev. B* **1991**, *43*, 1993.

[49] The RDFs for the central part of the liquid slab were re-analyzed for this work using the same bin widths as described in Section 2, whereas a significantly larger bin width was used in Ref. 15 to allow for a meaningful comparison with the smaller number of surface water molecules.

[50] G. Hura, J. M. Sorenson, R. M. Glaeser, T. Head-Gordon, *J. Chem. Phys.* **2000**, *113*, 9140; J. M. Sorenson, G. Hura, R. M. Glaeser, T. Head-Gordon, *J. Chem. Phys.* **2000**, *113*, 9149.

[51] G. Hura, D. Russo,R. M. Glaeser, T. Head-Gordon, M. Krack, M. Parrinello, *Phys. Chem. Chem. Phys.* **2003**, *5*, 1981.

This work was performed under the auspices of the U.S. Department of Energy by University of California, Lawrence Livermore National Laboratory under Contract W-7405-Eng-48.

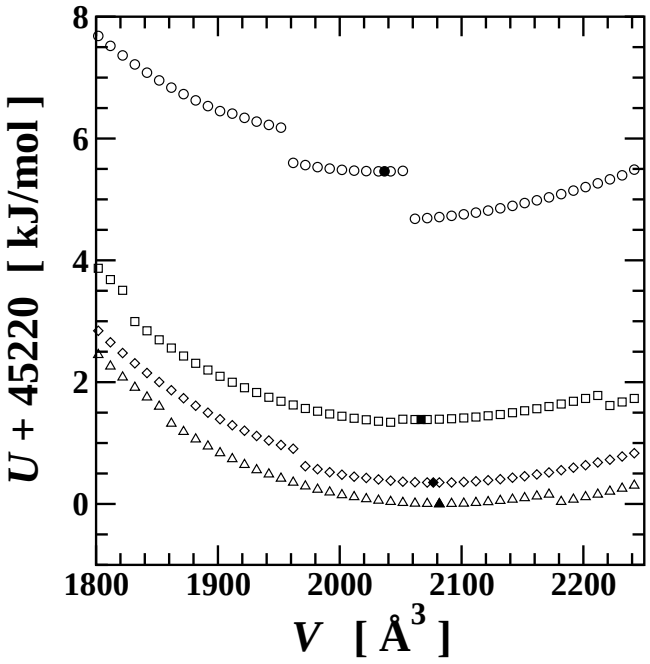

Figure 1: Comparison of the absolute potential energies (shifted by +45220 kJ/mol) of a single configuration as a function of volume for different charge density cutoffs: 280 Ry (circles), 600 Ry (squares), 900 Ry (diamonds), and 1200 Ry (triangles). The points that correspond to a slope of zero in the four $U - V$ curves are highlighted by filled symbols. For clarity only every second energy calculation is shown here (i.e., the spacing between the points is 10 Å^3). All regular grids were generated by using only prime lengths. It must be emphasized that the potential energies have been shifted in order to focus attention on the relative differences important for thermophysical properties and avoid the large energy offset arising from the atomization energy.

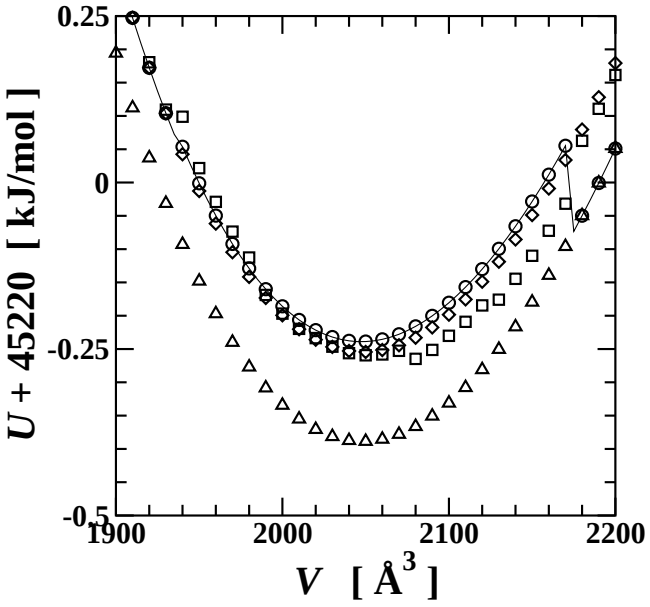

Figure 2: Comparison of the potential energies (shifted by +45220 kJ/mol) averaged over ten configurations as a function of the volume of the simulation box for the PRIME-GRID (circles and thin solid line), ALL-GRID (squares), SMALL-REF (diamonds), and LARGE-REF (triangles) grid generation methods. The charge density cutoff is 1200 Ry for all curves.

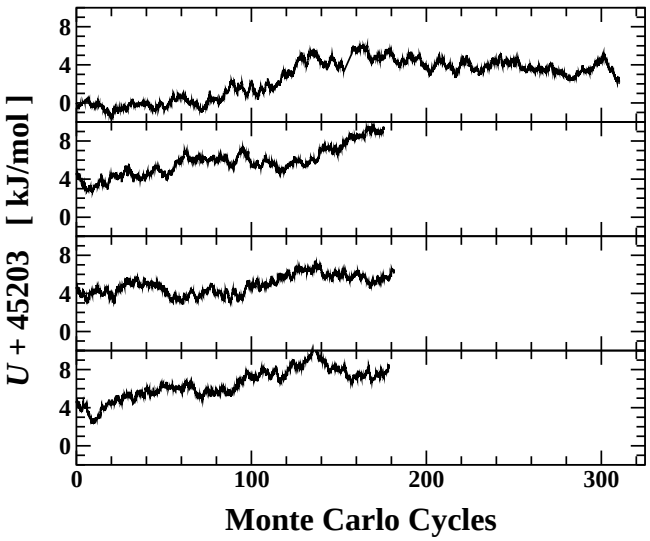

Figure 3: Comparison of the instantaneous energies for the four NpT simulations as a function of Monte Carlo cycles. From top to bottom, the values for runs PRIME-GRID, ALL-GRID, SMALL-REF, and LARGE-REF are depicted.

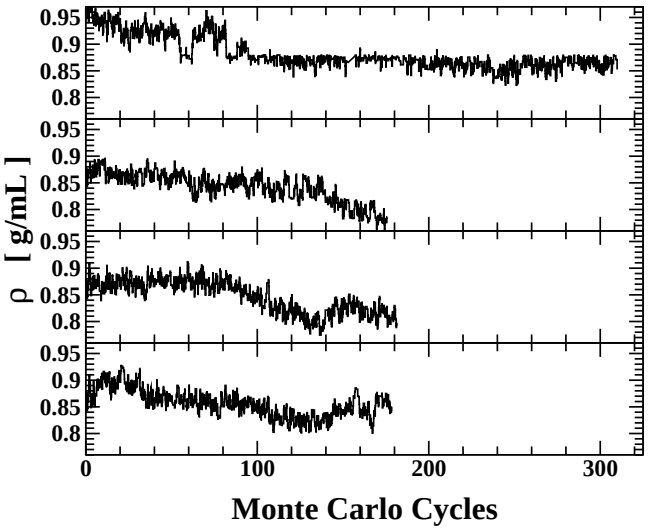

Figure 4: Comparison of the instantaneous densities for the four NpT simulations as a function of Monte Carlo. Layout as in Figure 3.

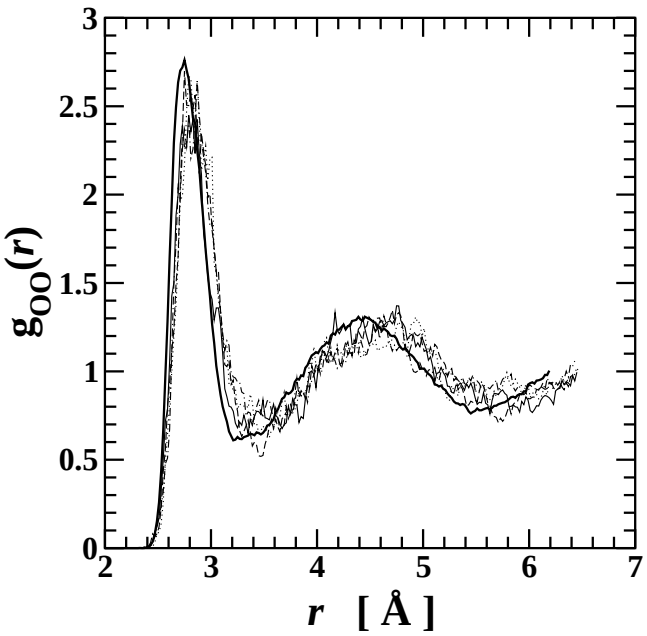

Figure 5: Comparison of the oxygen–oxygen radial distribution functions of the four NpT simulations and of a simulation using a liquid slab [15]. The thin solid, dashed, dotted, dash-dotted, and thick solid lines denote runs PRIME-GRID, ALL-GRID, SMALL-REF, LARGE-REF, and the liquid-slab simulation, respectively.

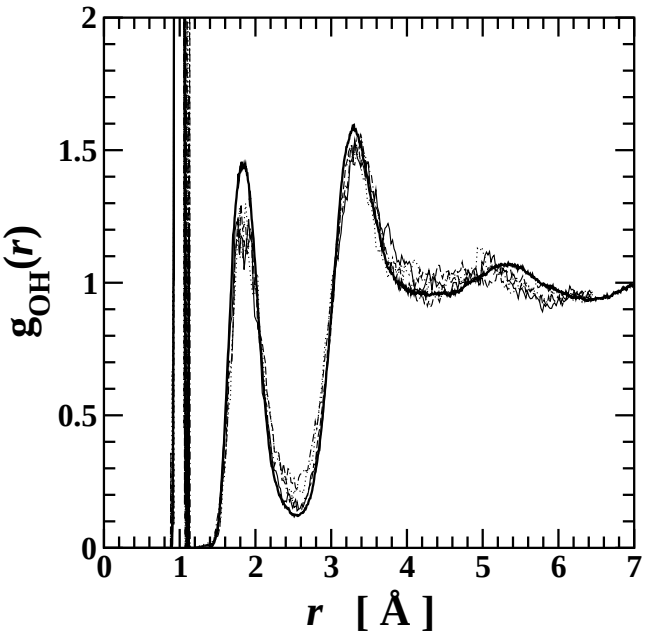

Figure 6: Comparison of the oxygen–hydrogen radial distribution functions of the four NpT simulations and of a simulation using a liquid slab [15]. The line styles are the same as in Figure 5.