LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

UCRL-JRNL-208404

Toward a Monte Carlo program for simulating vapor-liquid phaseequilibria from first principles

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

December 6, 2004

Computer Physics Communication

Toward a Monte Carlo program for simulating vapor–liquid phase equilibria from first principles

Matthew J. McGrath, [a] J. Ilja Siepmann, [a,*] I-Feng W. Kuo, [b] Christopher J. Mundy, [b] Joost VandeVondele, [c] Michiel Sprik, [c] Jürg Hutter, [d] Fawzi Mohamed, [e] Matthias Krack, [e] and Michele Parrinello [e]

[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.*

[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

Efficient Monte Carlo algorithms are combined with the `Quickstep` energy routines of `CP2K` to develop a program that allows for Monte Carlo simulations in the canonical, isobaric-isothermal, and Gibbs ensembles using a first principles description of the physical system. Configurational-bias Monte Carlo techniques and pre-biasing using an inexpensive approximate potential are employed to increase the sampling efficiency and to reduce the frequency of expensive *ab initio* energy evaluations. The new Monte Carlo program has been validated through extensive comparison with molecular dynamics simulations using the programs `CPMD` and `CP2K`. Preliminary results for the vapor–liquid coexistence properties ($T = 473$ K) of water using the Becke-Lee-Yang-Parr exchange and correlation energy functionals, a triple-zeta valence basis set augmented with two sets of *d*-type or *p*-type polarization functions, and Goedecker-Teter-Hutter pseudopotentials are presented. The preliminary results indicate that this description of water leads to an underestimation of the saturated liquid density and heat of vaporization and, correspondingly, an overestimation of the saturated vapor pressure.

Key words: Gibbs ensemble Monte Carlo, density functional theory, water
PACS: 05.20.-y, 05.70.-a, 61.20.Ja, 64.70.Fx

1. Introduction

Water holds a unique role among liquids, not only because of its ubiquity and importance on earth, but more so because of its anomalous liquid properties. Thus, understanding its properties has been a grand challenge for liquid state theory [1] and molecular simulation [2]. The first particle-based simulations of liquid water using pairwise empirical potentials were carried out almost 40 years ago [3]. However, the strong dipole moment and large polarizability of water and its participa-

* Corresponding author: siepmann@chem.umn.edu

tion in many chemical processes, particularly its self-dissociation, pose a challenge for empirical potentials. Although great strides have been made in the development of empirical force fields for water, none of these has yet succeeded to yield a quantitative description of the thermodynamic, structural, and dynamic properties of water over its entire liquid range [4]. In contrast, an *ab initio* representation of water affords the opportunity to study both physical and chemical properties.

The first Car-Parrinello molecular dynamics (CPMD) simulation [5] for liquid water employing a quantum-mechanical description of the molecular interactions was performed in the early 1990s, but available computer resources limited this CPMD simulation to a small system size (32 molecules) and short simulation length (1.5 and 2 ps for equilibration and production, respectively) [6]. Large increases in the available computer power and in the efficiency of the simulation approaches are now enabling more systematic studies of the liquid properties of water using first principles approaches. In particular, using the CPMD approach it is now possible to follow the trajectories of 64 water molecules for more than 10 ps [7–11] or of 216 water molecules for about 5 ps [12]. Although these simulations have contributed to a better understanding of structural aspects of neat liquid water, aqueous solutions [13–16], and the air-water interface [12], little research has been done to evaluate the thermodynamic properties of first principles descriptions of water. The notable exception is a recent work by Asthagiri *et al.* [17] aimed at estimating the free energy of liquid water using a combination of *ab initio* molecular dynamics and quasichemical theory. Using this somewhat indirect approach, Asthagiri *et al.* computed the free energy of hydration of liquid water at a fixed density of 1.0 g/cm^3 and reported values that range from -21 kJ/mol (at $T = 314$ K) for the revised PBE functional [18] to -60 kJ/mol (at $T = 337$ K) for the original PBE functional [19].

The aim of the present research is to use Monte Carlo simulations in the Gibbs ensemble [20–22] to directly calculate the vapor–liquid coexistence properties of water. The remainder of this manuscript gives a brief description of the newly developed Monte Carlo program, its validation through comparison with molecular dynamics simulations in the microcanonical and canonical ensembles, and preliminary results for the coexistence properties at $T = 473$ K.

2. Simulation Methods and Details

The first principles Monte Carlo simulations of this work were performed using the computer program `CP2K` [23,24] which is a general purpose program that builds upon the success of `CPMD` [25]. The electronic structure part of `CP2K`, called `Quickstep`, uses the Gaussian plane wave (GPW) method [26] for the calculation of forces and energies. The GPW method is based on the Kohn-Sham formulation [27] of density functional theory. For the GPW method, the Kohn-Sham orbitals are expanded using a linear combination of atom-centered Gaussian-type orbital functions, and an auxiliary basis set of plane waves describes the electronic charge density [26]. This combination is chemically more intuitive and computationally more efficient than sole use of a plane wave basis set. First principles simulations with `CP2K` sample directly from the Born-Oppenheimer surface [24].

The computational efficiency of the `Quickstep` energy routine is exploited in `CP2K-MC`, a module of `CP2K` that allows for Monte Carlo sampling from various statistical-mechanical ensembles [2], including those with fluctuating particle numbers. Use of a Monte Carlo framework is advantageous because temperature, pressure, and chemical potential are explicitly accounted for in the acceptance rules [2], and the momentum part of the Hamiltonian can be removed by integration [28], i.e. the configurational properties of the classical phase space are independent of any choice of mass. Our Monte Carlo implementation for water in the canonical ensemble employs three different types of trial moves: (i) translations of rigid molecules, (ii) rigid-body rotations around the molecular center of mass, and (iii) conformational moves altering either bond length or angle. The type of move is selected at random and the separate move types ensure equipartition of the translational, rotational, and vibrational degrees of freedom of the system.

Pre-sampling of trajectories [29,30] with an inexpensive approximate potential [31] for a short sequence of moves is carried out to reduce the number of expensive *ab initio* energy evaluations, thereby enhancing the computational efficiency.

Two additional types of Monte Carlo moves are required for a simulation in the Gibbs ensemble that utilizes two separate simulation boxes for the vapor and liquid phases, i.e. the phases are in thermodynamic contact but do not share an explicit interface. Mechanical equilibrium is established by volume exchanges between the two simulation boxes. The internal structure of every molecule is kept rigid during a volume exchange and only the center-of-mass positions are scaled [32]. Pre-sampling is not used for the volume exchange moves. A configurational-bias Monte Carlo approach [33–36] is used to efficiently swap molecules between the two phases, thereby equalizing the chemical potential of water in each phase. The particle swap move uses a two-step split-energy configurational-bias technique [37,38] involving 640 trial positions that are all evaluated using the approximate potential function. Thereafter, one suitable trial position is selected using the relative Rosenbluth weights calculated for the approximate potential and its true energy is evaluated using the `Quickstep` routine. The final acceptance step for this swap move then corrects for the energy difference between the approximate and *ab initio* energies [37,38].

The Gibbs ensemble Monte Carlo simulation at $T = 473$ K is preformed for a system consisting of a total of 64 water molecules. The initial configuration for the Gibbs ensemble simulation was prepared using the final configuration of a canonical ensemble simulation at $T = 315$ K and $\rho = 1.0$ g/cm^3 that was scaled to the experimental saturated liquid density at $T = 473$ K. The volume of the initially empty vapor box was set equal to the initial liquid volume. The electronic structure parameters used in the Gibbs ensemble simulation are as follows: the Becke-Lee-Yang-Parr (BLYP) exchange and correlation energy functionals [39,40] together with a triple-zeta valence basis set augmented with two sets of *d*-type or *p*-type polarization functions (TZV2P), the norm-conserving pseudopotentials of Goedecker *et al.* (GTH) [41], and a relatively large charge density cut-off at 1200 Ry for the auxiliary plane wave basis set.

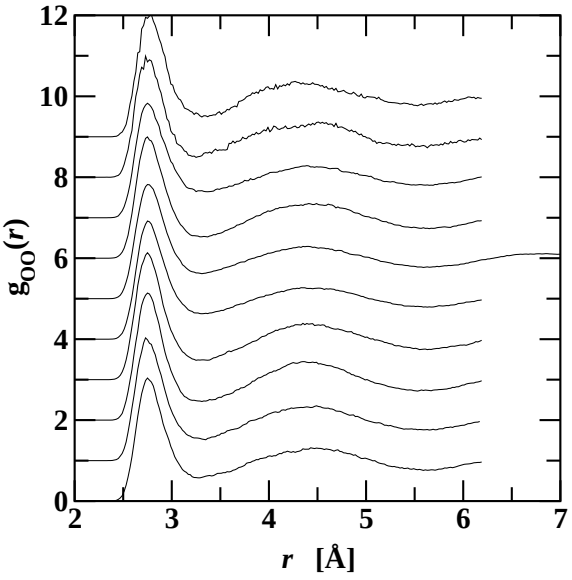

Fig. 1. Comparison of the oxygen–oxygen radial distribution functions for liquid water at $\rho = 1.0$ g/cm^3 and $T \approx 315$ K obtained from first principles simulations using the BLYP exchange and correlation energy functionals. Shown from bottom to top (and off-set from each other by one unit) are the RDFs for runs `CPMD`-NVE-BO, `CP2K`-MD-NVE-1, `CP2K`-MD-NVE-2, `CP2K`-MD-NVE-3, `CPMD`-NVE-400, `CPMD`-NVE-400-128, `CPMD`-NVT-i-400, `CPMD`-NVT-ie-800, `CP2K`-MC-NVT-1, and `CP2K`-MC-NVT-2. A complete description of the simulation parameters can be found in Ref. [11].

3. Validation for Monte Carlo simulations in the canonical ensemble

In light of two recent reports [10,17] that have questioned the reproducibility of structural parameters of liquid water obtained from first principles simulations, significant effort was devoted to assess this question and to validate the current Monte Carlo approach for simulations in the canonical ensemble. A detailed report of this validation study has been published elsewhere [11], thus only a brief account is given here.

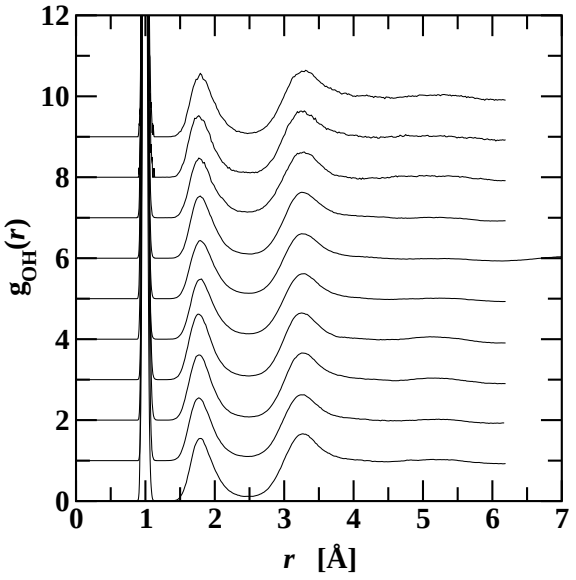

Fig. 2. Comparison of the oxygen–hydrogen radial distribution functions for liquid water at $\rho = 1.0$ g/cm^3 and $T \approx 315$ K obtained from first principles simulations. Lines as in Fig. 1.

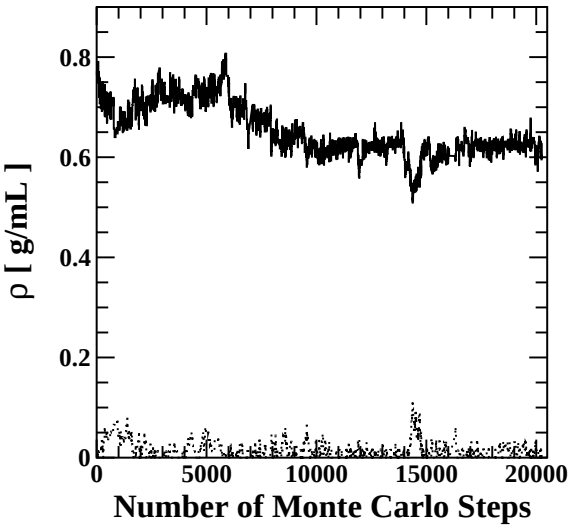

Fig. 3. Instantaneous values of the vapor and liquid densities (depicted as solid and dashed lines, respectively) obtained from Gibbs ensemble simulations for BLYP-GTH-TZV2P-1200 water at $T = 473$ K.

Figures 1 and 2 show the oxygen–oxygen and oxygen–hydrogen radial distribution functions (RDFs) for ten different first principles simulations. Clearly, these simulations do not show the striking non-uniformity discussed by Asthagiri *et al.* [17] and Grossman *et al.* [10]. The average values for height and position of the first peak in the oxygen–oxygen RDFs are 3.0 ± 0.1 and 2.75 ± 0.02 Å, respectively, and the coordination numbers are close to the tetrahedral value of 4 for all simulations. Here it should be noted that some scatter in the RDFs is unavoidable because of (i) statistical uncertainties arising from "short" simulation length and "small" system size compared to simulations using empirical potentials, and (ii) unavoidable variations in the ionic temperature for simulations in the microcanonical ensemble. Most importantly, the structural properties calculated for the `CP2K-MC` simulations in the canonical ensemble agree very well with those from molecular dynamics simulations. Furthermore, the values of the classical constant-volume heat capacity obtained from the Monte Carlo and molecular dynamics simulations were also found to be in good agreement [11].

4. Gibbs ensemble Monte Carlo simulations

Figures 3 and 4 show the instantaneous values of the vapor and liquid densities and the heat of vaporization, respectively, observed throughout the Gibbs ensemble trajectory at $T = 473$ K. It appears that the system has reached equilibrium after about 10^5 Monte Carlo steps (or about 150 Monte Carlo cycles) where it should be noted that this number corresponds to the `Quickstep` energy calculations, i.e. the number of pre-sampling moves using the approximate potential is significantly larger. The fluctuations in the heat of vaporization are relatively large compared to simulations for rigid water molecules using empirical potential, but it should be recognized that the current simulation uses classical sampling for the vibrational degrees of freedom, i.e. a significant amount of heat is stored in these internal degrees of freedom. A given pre-sampling sequence can involve molecules in both phases, which means a

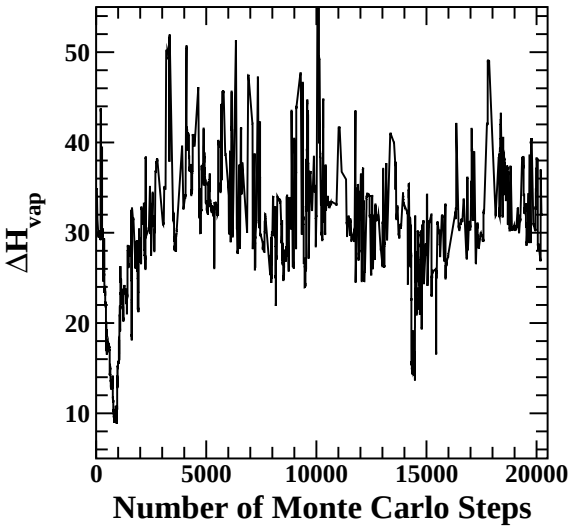

Fig. 4. Instantaneous values of the heat of vaporization obtained from Gibbs ensemble simulations for BLYP-GTH-TZV2P-1200 water at $T = 473$ K.

decrease in the total energy of the vapor phase might be compensated by an increase in the total energy of the liquid phase and *vice versa* to lead to the acceptance of the move sequences but these two changes can add together for a drastic change in the heat of vaporization.

When the properties of the Gibbs ensemble simulations are averaged over the final 150 Monte Carlo cycles, the following estimates of the saturation properties at $T = 473$ K are obtained (with their experimental counterparts listed in parenthesis): $\rho_{\text{liquid}} = 0.61$ g/cm^3 (0.865 g/cm^3), $\rho_{\text{vapor}} = 0.01$ g/cm^3 (0.0079 g/cm^3), and $\Delta H_{\text{vap}} = 32$ kJ/mol (34.9 kJ/mol).

5. Conclusions

The main conclusions of the current work are (a) that reproducible structural properties for liquid water can be obtained from constant-volume (N,V,E and N,V,T) simulations using different first principles sampling approaches [11], (b) that the BLYP description of water yields a somewhat overstructured liquid at $T \simeq 315$ K and $\rho = 1.0$ g/cm^3 [11], and (c) that the BLYP-GTH-TZV2P-1200 description of water leads to an underestimation of the saturated liquid density and heat of vaporization and, correspondingly, an overestimation of the saturated vapor pressure at $T = 473$ K.

6. Acknowledgments

We thank 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 Department of Energy Computational Science Graduate Fellowship (M.J.M.) is gratefully acknowledged. 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] Hansen, J. P.; McDonald, I. R. *Theory of Simple Liquids*, 2nd ed. (Academic Press: New York, 1986).

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

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

[4] Chen, B.; Potoff, J. J.; Siepmann, J. I. *J. Phys. Chem. B* **2000**, *104*, 2378.

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

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

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

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

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

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

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

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

[13] Molteni, C.; Parrinello, M. *J. Am. Chem. Soc.* **1998**, *120*, 2168.

[14] Bruge, F.; Bernasconi, M.; Parrinello, M. *J. Am. Chem. Soc.* **1999**, *121*, 10883.

[15] Tuckerman, M. E.; Marx, D.; Parrinello, M. *Nature* **2002**, *417*, 925.

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

[17] Asthagiri, D.; Pratt, L. R.; Kress, J. D. *Phys. Rev. E* **2003**, *68*, 041505.

[18] Zhang, Y.; Yang, W. *Phys. Rev. Lett.* **1998**, *80*, 890.

[19] Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865.

[20] Panagiotopoulos, A.Z. *Mol. Phys.* **1987**, *61*, 813.

[21] Panagiotopoulos, A.Z.; Quirke, N.; Stapleton, M.; Tildesley, D.J.; *Mol. Phys.* **1988**, 63, 527.

[22] Smit, B.; de Smedt, P.; Frenkel, D.; *Mol. Phys.* **1989**, 68, 931.

[23] CP2K, cp2k.berlios.de

[24] VandeVondele, J. REVIEW NEED DETAILS

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

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

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

[28] McQuarrie, D. A. *Statistical Mechanics* (Harper and Row: New York, 1996).

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

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

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

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

[33] Siepmann, J.I.; Frenkel, D.; *Mol. Phys.* **1992**, 75, 59.

[34] Frenkel, D.; Mooij, G.C.A.M.; Smit, B.; *J. Phys.: Cond. Matt.* **1992**, 4, 3053.

[35] Siepmann, J. I.; Karaborni, S.; Smit, B. *Nature* **1993**, *365*, 330.

[36] Martin, M. G.; Siepmann, J. I. *J. Phys. Chem. B* **1999**, *103*, 4508.

[37] Smit, B.; Karaborni, S.; Siepmann, J. I. *J. Chem. Phys.* **1995**, *102*, 2126.

[38] Vlugt, T. J. H.; Martin, M. G.; Smit, B.; Siepmann, J. I.; Krishna, R. *Mol. Phys.* **1998**, *94*, 727.

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

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

[41] Goedecker, S.; Teter, M.; Hutter, J. *Phys. Rev. B* **1996**, *54*, 1703; Hartwigsen, C.; Goedecker, S.; Hutter, J. *Phys. Rev. B* **1998**, *58*, 3641.

This work was performed under the auspices of the U.S. Department of Energy by University of California, Lawrence Livermore National Laboratory under contract no. W-7405-Eng-48.