Bulk Liquid Water at Ambient Temperature and Pressure from MP2 Theory

Del Ben, Mauro; Schoenherr, Mandes; Hutter, Juerg; VandeVondele, Joost

Abstract: MP2 provides a good description of hydrogen bonding in water clusters and includes long-range dispersion interactions without the need to introduce empirical elements in the description of the interatomic potential. To assess its performance for bulk liquid water under ambient conditions, an isobaric-isothermal (NpT) Monte Carlo simulation at the second-order Moller-Plesset perturbation theory level (MP2) has been performed. The obtained value of the water density is excellent (1.02 g/mL), and the calculated radial distribution functions are in fair agreement with experimental data. The MP2 results are compared to a few density functional approximations, including semilocal functionals, hybrid functionals, and functionals including empirical dispersion corrections. These results demonstrate the feasibility of directly sampling the potential energy surface of condensed-phase systems using correlated wave function theory, and their quality paves the way for further applications.

DOI: https://doi.org/10.1021/jz401931f

Posted at the Zurich Open Repository and Archive, University of Zurich
ZORA URL: https://doi.org/10.5167/uzh-87661
Accepted Version

Originally published at:
DOI: https://doi.org/10.1021/jz401931f
Bulk Liquid Water at Ambient Temperature and Pressure from MP2 Theory

Mauro Del Ben,*† Mandes Schönherr,*† Jürg Hutter,*† and Joost VandeVondele,*‡

Institute of Physical Chemistry, University of Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland, and Department of Materials, ETH Zürich, Wolfgang-Pauli-Strasse 27, CH-8093 Zürich, Switzerland

E-mail: delben@pci.uzh.ch; m.schoenherr@pci.uzh.ch; hutter@pci.uzh.ch; Joost.VandeVondele@mat.ethz.ch

*To whom correspondence should be addressed
†Institute of Physical Chemistry, University of Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland
‡Department of Materials, ETH Zürich, Wolfgang-Pauli-Strasse 27, CH-8093 Zürich, Switzerland
Abstract

MP2 provides a good description of hydrogen bonding in water clusters and includes long range dispersion interactions without the need to introduce empirical elements in the description of the interatomic potential. To assess its performance for bulk liquid water under ambient conditions, an isobaric-isothermal (NpT) Monte Carlo simulation at the second-order Møller-Plesset perturbation theory level (MP2) has been performed. The obtained value of the water density is excellent (1.02 g/mL) and the calculated radial distribution functions are in fair agreement with experimental data. The MP2 results are compared to a few density functional approximations, including semi-local functionals, hybrid functionals, and functionals including empirical dispersion corrections. These results demonstrate the feasibility of directly sampling the potential energy surface of condensed phase systems using correlated wavefunction theory, and their quality paves the way for further applications.

**keywords:** Liquids, Hydrogen Bonding, Electron Correlation, Density Functional Theory, Monte Carlo
Understanding the structural and electronic properties of liquid water at ambient conditions is a major challenge in condensed matter simulations. Water is a crucial ingredient for a large variety of systems of prime importance in basic chemistry, biology, and physics, as well as in the applied fields of catalysis and energy production. The water molecule has a large dipole moment and polarizability, is a multiple hydrogen donor and acceptor and can easily build network structures. The total cohesive energy in the condensed phase is, as a consequence of these properties, a sum of many weak interactions. Theoretical models face therefore the challenge to describe many different effects and their subtle interplay at a high precision. The development of sophisticated empirical potentials for water\textsuperscript{1–10}, allowed to gain insights into water’s behavior and its thermodynamic properties\textsuperscript{11–13}, such as, density maxima, heat capacity and effects of supercooling. However, empirical models lack transferability and might fail if used under conditions away from their fitting range. Most importantly, as soon as water takes an active role in a chemical process, either as a strongly interacting solvent, or for example as a source of protons, the electronic properties of the water molecule need to be taken into account. In this respect, first-principles methods offer the possibility to describe all the underlying physics on the same footing, simplifying the treatment of intra- and inter-molecular interactions. The capability to reproduce properties of complex systems such as liquid water can therefore be used to judge the sophistication and predictive power of a given quantum mechanical model. Density functional theory (DFT) is the most used quantum mechanical method employed for studying physical and chemical properties of condensed phase systems. Many DFT based simulation of bulk water have been reported in the literature, and in this context three main methods of sampling the phase space can be recognized\textsuperscript{14}: the Car-Parrinello molecular dynamics (CPMD) and related variants\textsuperscript{15–21}, Born-Oppenheimer molecular dynamics (BOMD)\textsuperscript{22–26} and Monte Carlo (MC) sampling.\textsuperscript{27}

Most of the CPMD and BOMD simulations were carried out in either the microcanonical (NVE) or canonical (NVT) ensemble by constraining the volume to reproduce the experimental density $\rho$. First principles MD simulations in the isobaric-isothermal (NpT) ensemble are much less common\textsuperscript{26,28}, in part due to the need for different integration schemes, the increased requirements
on plane waves basis sets in variable cell simulations, and the long simulation time that is needed to equilibrate and to sample volume fluctuations. Moreover, in the MD sampling of the NpT ensemble, the calculation of the virial and thus stresses is required to change the volume as a response to the imbalance between internal stress and external pressure.\textsuperscript{29} In this respect, the appealing feature of the MC method is that thermodynamic constraints are explicitly included into the acceptance rule for each trial move, solely based on the energy. This allows to perform simulations in different ensembles, \textit{e.g.} NpT, in a relatively straightforward manner. On the other hand, an efficient sampling of phase space in MC requires smart and system dependent trial moves\textsuperscript{30–34}, making the application of the method more intricate than molecular dynamics where configurational sampling follows a general principle. Within the framework of MC, McGrath and coworkers\textsuperscript{27} reported the first results from first-principles simulations of liquid water in the isobaric-isothermal ensemble at ambient pressure. These NpT-MC simulations performed at ambient conditions using the Becke-Lee-Yang-Parr (BLYP)\textsuperscript{35,36} functional gave a significantly less dense (about 20\%) liquid than observed experimentally. These results were confirmed and extended by Schmidt \textit{et al.}\textsuperscript{26} using very similar computational setups, but within the framework of NpT-MD. In the latter work, in addition to BLYP, also the Perdew-Burke-Ernzerhof (PBE)\textsuperscript{37} functional has been tested, and, in both cases, the influence of an empirical dispersion correction (DFT-D)\textsuperscript{38} has been investigated. It was shown that pure PBE and BLYP are indeed underestimating the water density, but the inclusion of the dispersion correction significantly improves the results giving a density as well as oxygen-oxygen radial distribution function (RDF) closer to the experimental data. From these calculations the important role played by the van der Waals interactions in bulk water clearly emerged, and it was shown that this missing interaction in standard local functionals can be efficiently included using empirical corrections.\textsuperscript{38–40} Furthermore, it became clear that NpT simulations are essential to quantify the quality of the intermolecular interaction potential and that a fixed simulation volume should be considered an influential constraint. Despite this progress, a truly first principles simulation of liquid water in the NpT ensemble is still missing.

Here, we present the results of MP2 based NpT-MC simulation of liquid water at ambient
conditions. The second-order Møller-Plesset energy\textsuperscript{41,42} is an effective correction to the Hartree–Fock (HF) ground state energy that accounts for electron correlation effects. It recovers a relatively large part of the dynamic correlation, while the HF ground state ensures the inclusion of exact non-local exchange. Most notably, MP2 introduces dispersion interactions in a completely non-empirical way. For water clusters, MP2 produces results that are in excellent agreement with coupled cluster theory\textsuperscript{43–46}, and is one of the best methods for the WATER27 benchmark.\textsuperscript{47} However, the advantages of MP2 come at a computational cost that is high compared to traditional DFT using local functionals. Furthermore, MP2 calculations need larger basis sets than DFT to reach a similar state of convergence. Therefore, only few applications of MP2 to condensed phase systems have been reported (see e.g.\textsuperscript{48–50}) and condensed phase sampling has not yet been performed. The extensive calculations reported here have become possible by combining large computer resources with an efficient algorithm for the MP2 energy calculation, which we named the Resolution of Identity Gaussian and Plane Wave (RI-GPW) approach\textsuperscript{51,52} and implemented in the CP2K\textsuperscript{53} program. In addition to the new MP2 results, we report also results of NpT-MC simulations obtained from a selection of density functional approximations, namely, BLYP, PBE and PBE0\textsuperscript{54}, including empirical dispersion corrections of the D3 type\textsuperscript{40}, which have been generated to validate the approach.

All calculations presented have been performed with the CP2K program.\textsuperscript{53} The energy at the Hartree–Fock and DFT level is computed using the Gaussian and Plane Wave (GPW)\textsuperscript{55,56} method. Within the GPW approach a Gaussian basis is used to expand molecular orbitals and an auxiliary plane wave basis for the expansion of the electronic density. In order to efficiently expand the density in plane waves, core electrons are replaced by pseudopotentials.\textsuperscript{57} The non-local HF exchange calculations have been performed employing a robust Γ-point implementation,\textsuperscript{58,59} and, only in the case of hybrid functional (PBE0), the calculations have been performed using the auxiliary density matrix method (ADMM)\textsuperscript{60} that allows for an approximate calculation of the non-local exchange energy at much reduced cost. The calculation of the MP2 energy within the RI-GPW\textsuperscript{52} approach (RI-MP2) is closely related to the original GPW method. Here, the dual representation
of the electronic density is applied to the fitting density arising from the resolution of identity approximation\textsuperscript{61–63}, where the RI fitting densities are obtained by introducing an auxiliary Gaussian basis and employing the Coulomb metric.\textsuperscript{64}

The model system consists of 64 water molecules in a cubic simulation cell under periodic boundary conditions (PBC). All simulations have been obtained with thermodynamic constraints set to ambient conditions, that is, $T = 295\text{K}$ and $p = 1\text{bar}$ and for all the theoretical models presented (DFT or RI-MP2) the MC settings are kept the same. The MC efficiency is improved with the presampling of moves.\textsuperscript{30,32} In this method an additional inexpensive (approximated) potential is introduced and used to generate a short sequence of MC moves which is entirely accepted or rejected based on the difference between the exact and approximated potential functions. The employed approximate potential is a classical but refitted force field based on Ref.\textsuperscript{65} for the DFT simulations, and a semi-local density functional method for MP2.

Representative run times for a single energy calculation on a Cray XE6 (32 cores per node) computer are 218 seconds for RI-MP2 on 400 nodes, 34 seconds for ADMM-PBE0 on 24 nodes, and 17 seconds for local density functionals on 16 nodes. The detailed description of the employed methods together with the related computational setups are reported in the supporting information (SI).

Table 1: Average density and structural data obtained from the MC NpT simulations at the various level of theory considered ($T = 295\text{K}$ and $p = 1\text{bar}$). Experimental values taken from Ref.\textsuperscript{66}. CN is the average coordination number calculated from the integral $4\pi\rho_n r^2 g_{OO}(r)$ up to the first minimum, where $\rho_n$ is the average number density of Oxygen atoms.

<table>
<thead>
<tr>
<th></th>
<th>Density [g/mL]</th>
<th>1\textsuperscript{st} Max</th>
<th>1\textsuperscript{st} Min</th>
<th>2\textsuperscript{nd} Max</th>
<th>CN</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\rho$</td>
<td>RMSD</td>
<td>$r[\text{Å}]$</td>
<td>$g_{OO}(r)$</td>
<td>$r[\text{Å}]$</td>
</tr>
<tr>
<td>BLYP</td>
<td>0.797</td>
<td>0.018</td>
<td>2.83</td>
<td>2.44</td>
<td>3.46</td>
</tr>
<tr>
<td>BLYP-D3</td>
<td>1.066</td>
<td>0.018</td>
<td>2.78</td>
<td>3.01</td>
<td>3.51</td>
</tr>
<tr>
<td>PBE-D3</td>
<td>1.055</td>
<td>0.015</td>
<td>2.73</td>
<td>3.24</td>
<td>3.15</td>
</tr>
<tr>
<td>PBE0-ADMM-D3</td>
<td>1.023</td>
<td>0.013</td>
<td>2.74</td>
<td>3.23</td>
<td>3.30</td>
</tr>
<tr>
<td>RI-MP2</td>
<td>1.020</td>
<td>0.015</td>
<td>2.76</td>
<td>3.12</td>
<td>3.32</td>
</tr>
<tr>
<td>exp.</td>
<td>1.00</td>
<td>-</td>
<td>2.80</td>
<td>2.57</td>
<td>3.45</td>
</tr>
</tbody>
</table>
A central result of this work is the computed value of the density of liquid water. This quantity is obtained from averaging the fluctuating instantaneous density as obtained during the MC simulation. The instantaneous density and the corresponding running average are depicted in Figure 1, while the calculated average value with the associated root mean square deviations (RMSD) are reported in Table 1. Our MP2 result for the density of liquid water at ambient conditions is \( \approx 1.020 \) g/mL. We consider this to be in excellent agreement to the experimental value, since it implies an error in the lattice parameters of less than 1%. We attribute this to the known quality of MP2 for describing water hydrogen bonds, and the presence of dispersion interactions. Moreover, it is known from classical
and DFT based simulations that, for flexible water models, the inclusion of nuclear quantum effects leads to less structured liquid and improve the density behavior.\textsuperscript{67–70} By observing that these effects are less pronounced in \text{D}_2\text{O} than in \text{H}_2\text{O} and that the former has a molar density 1.3\% higher than the latter, we expect the inclusion of this correction to reduce further the density.\textsuperscript{69,71} This result stands also in sharp contrast to that of BLYP, a dispersion free functional that yields 0.797 g/mL. The reason for this is that the water density depends crucially on the medium to long range part of the potential.\textsuperscript{23,72} Dispersion corrected density functionals, BLYP-D3, PBE-D3, and PBE0-ADMM-D3 all have densities that are much closer to experiment, albeit slightly too dense. The density obtained for PBE0-ADMM-D3 (1.023 g/mL) is of MP2 quality, but this might be in part fortuitous, as the small basis employed in the ADMM approach could influence this result. Nevertheless, it is known that the PBE0 functional improves PBE results for structural, spectroscopic and thermodynamics properties\textsuperscript{54,73,74}, and yields good results for water dimer interactions when compared to high level coupled cluster calculations.\textsuperscript{75}

Computing a first principles estimate of the density is challenging, and long simulations are essential to sample fluctuations and to equilibrate the system. If a significant structural reorganization would be needed, equilibration times could easily exceed simulation times. In this respect, our RI-MP2 simulation is a ‘best effort’ simulation, with a length constrained by the high computational cost. The number of MC cycles is larger in the corresponding DFT simulations. However, two observations enhance our confidence in the computed MP2 value. First, the total number of accepted MC cycles is similar for all reported calculations. This can be attributed to the high quality of DFT presampling in the MP2 case, which results in a much higher acceptance rate (∼50\%) compared to the other simulations using presampling based on a classical force field (15 – 17\%). Second, those methods that remain close to the experimental density (RI-MP2 and PBE0-ADMM-D3) are likely, and observed, to require shorter equilibration periods.

Further support for the employed methodology, but also insight in the typical error bars, comes from a comparison with literature DFT results. At the BLYP level, the calculated density is 0.797 g/mL, in agreement with the results reported by McGrath and coworkers\textsuperscript{27} (0.8 g/ml), obtained
by NpT-MC, Schmidt et al.\textsuperscript{26} (0.73-0.78 g/mL) calculated employing NpT-BOMD, and Wang et al.\textsuperscript{72} (0.76-0.85 g/mL). In the latter case, the water density is inferred from pressure density curves obtained from series of NVT-MD simulations performed at different volumes. A larger deviation is observed when the BLYP density is compared to the value calculated by Ma and coworkers\textsuperscript{28} (0.92 g/mL) from NpT-CPMD using a DVR basis set. In that work, the large deviation is attributed to the usage of a converged basis set that contributes to softening the structure and improving diffusivity.\textsuperscript{76} However, even though the basis set used in this work is significantly larger than in Ref.\textsuperscript{26} and \textsuperscript{27}, a strong basis set dependence of the density at the BLYP level is not observed. The value calculated at the BLYP-D3 level (1.066 g/mL) is in agreement with that reported by Ma et al.\textsuperscript{28} (1.07 - 1.13 g/mL). When comparing our BLYP-D3 and PBE-D3 with BLYP-D2 and PBE-D2 results reported by Schmidt and coworkers\textsuperscript{26} deviations of 7\% and 11\% are observed, respectively. These deviations have to be attributed to the use of the D3 correction instead of D2, or potentially the use of larger basis sets and simulation lengths in this work.

The more detailed structure of the liquid is summarized in Figure 2 by the radial distribution functions for Oxygen-Oxygen ($g_{OO}(r)$) and Oxygen-Hydrogen ($g_{OH}(r)$), and quantified in Table 1. The MP2 results are in good agreement with experiment, in particular the mid and long range parts. The first minimum is relatively shallow, suggesting that a diffusive liquid is obtained. However, MD is required to obtain a precise value for the diffusion constant, as time correlation functions can not be extracted from these MC simulations. The maximum of the first peak is too pronounced, i.e. 3.12 vs 2.57. This difference is in part explained by the fact that our simulations ignore nuclear quantum effects, which influence this property.\textsuperscript{67–70} The coordination number, which condenses the shape of the first peak into a single number, is in fair agreement with experiment 4.6 vs. 4.3. The DFT results are less satisfactory, in particular BLYP without dispersion produces a very structured liquid as quantified by the low value of the first minimum. Note that the $g_{OO}(r)$ obtained under NpT conditions is significantly different from the one obtained in the NVE ensemble (see e.g. Ref.\textsuperscript{24}) at experimental density. In particular, a low coordination number is observed. The D3 dispersion correction leads to a $g_{OO}(r)$ that is almost featureless after the first peak, such a smearing out of the
Figure 2: Radial distribution functions for Oxygen-Oxygen (left) and Oxygen-Hydrogen (right) distances obtained from the NpT-MC simulations at the various level of theory considered ($T = 295K$ and $p = 1$ bar). (a),(d) BLYP (green dashed) and BLYP-D3 (red dotted); (b),(e) PBE-D3 (green dashed) and PBE0-ADMM-D3 (red dotted); (c),(f) RI-MP2 (green dashed line). The most recent experimental Oxygen-Oxygen RDF, obtained from x-ray diffraction and taken from Ref. 66, is depicted as a solid black line.
second peak with inclusion of dispersion has been discussed by Møgelhøj et al. in Ref.\textsuperscript{25}. The large coordination number can be considered an artefact of the shallow minimum observed, and depends strongly on the precise location of the minimum. PBE-D3 and PBE0-ADMM-D3 give very similar $g_{oo}(r)$, as was previously observed\textsuperscript{58} in simulations without the dispersion correction. Whereas these functionals lead to a slightly more structured liquid, they are similar to the MP2 results, in particular PBE0-ADMM-D3. The latter functional could thus be a computational expedient alternative to MP2, for example to investigate nuclear quantum effects.

Results have been reported for liquid water at ambient conditions as obtained from NpT-MC simulations at the MP2 level theory, which is free from empirical parameters. The obtained density of 1.02 g/mL is in excellent agreement with the experimental value, and the $g_{oo}(r)$ generally agrees well with the most recent experimental observations, albeit with a slightly too high first maximum. Other popular DFT methods have been tested, in particular showing good results for the water density at the PBE0-ADMM-D3 level, while being generally in agreement with previously reported work for other DFT methods. It is clear that the inclusion of dispersion interactions and an accurate description of the hydrogen bond are crucial ingredients for the correct simulation of liquid water at ambient conditions. The quality and accuracy of the obtained MP2 results paves the way to further applications of this technique for challenging questions in the field of aqueous solutions, e.g. the structures of the solvated electron and hydroxyl ion.

**Acknowledgement**

J.V. acknowledges financial support by the European Union FP7 in the form of an ERC Starting Grant under contract no. 277910. We acknowledge that the results of this research have been achieved using the PRACE Research Infrastructure resource Hermit based in Germany at Stuttgart (HLRS). Additional calculations were enabled by the Swiss National Supercomputer Centre (CSCS) under project IDs s441 and h05. The research leading to these results has received funding from the Swiss University Conference through the High Performance and High Productivity Computing (HP2C) Programme.
Supporting Information Available

Detailed description of the employed methods with the related computational setups, pseudopotentials, basis-set parameters and initial geometry. This material is available free of charge via the Internet at http://pubs.acs.org. This material is available free of charge via the Internet at http://pubs.acs.org/.

References


(74) Santra, B.; Michaelides, A.; Scheffler, M. On the Accuracy of Density-functional Theory
