Communication: Energy benchmarking with quantum Monte Carlo for water nano-droplets and bulk liquid water

D. Alfè, A. P. Bartók, G. Csányi, and M. J. Gillan

Additional information on J. Chem. Phys.

Communication: Energy benchmarking with quantum Monte Carlo for water nano-droplets and bulk liquid water

D. Alfè,[1,2,3,4] A. P. Bartók,[5] G. Csányi,[5] and M. J. Gillan[2,3,4,a)]

[1]*Department of Earth Sciences, University College London, London WC1E 6BT, United Kingdom*
[2]*London Centre for Nanotechnology, University College London, London WC1H 0AH, United Kingdom*
[3]*Thomas Young Centre, University College London, London WC1H 0AH, United Kingdom*
[4]*Department of Physics and Astronomy, University College London, London WC1E 6BT, United Kingdom*
[5]*Department of Engineering, University of Cambridge, Cambridge CB2 1PZ, United Kingdom*

We show the feasibility of using quantum Monte Carlo (QMC) to compute benchmark energies for configuration samples of thermal-equilibrium water clusters and the bulk liquid containing up to 64 molecules. Evidence that the accuracy of these benchmarks approaches that of basis-set converged coupled-cluster calculations is noted. We illustrate the usefulness of the benchmarks by using them to analyze the errors of the popular BLYP approximation of density functional theory (DFT). The results indicate the possibility of using QMC as a routine tool for analyzing DFT errors for non-covalent bonding in many types of condensed-phase molecular system. © 2013 AIP Publishing LLC.
[http://dx.doi.org/10.1063/1.4810882]

It has proved remarkably difficult to achieve a fully *ab initio* understanding of water systems, despite sustained efforts over the past 40 years (see, e.g., Refs. [1–5]). Coupled-cluster benchmarks for small water clusters[6–8] have been invaluable in calibrating the methods of density functional theory (DFT) and in creating parameterised interaction potentials. However, data on small clusters are not enough for understanding extended water systems, and several groups have recently begun to report benchmarks on clusters as large as $(H_2O)_{24}$[9,10] and on ice structures.[11,12] We show here how quantum Monte Carlo (QMC)[13,14] can provide accurate benchmark energies for large thermal samples of water clusters (nano-droplets) and for bulk liquid water itself. We illustrate the usefulness of these benchmarks by using them to analyse the errors of the popular BLYP approximation of DFT. Technical details of our calculations are reported in the supplementary material.[15]

It is well known that standard DFT approximations have significant difficulties in describing water systems.[16] The BLYP approximation,[17] which has been widely used for studying clusters,[18,19] ice structures,[20] and the bulk liquid,[21–25] significantly under-binds the dimer,[18] predicts the wrong stability ordering of isomers of the hexamer,[19] gives erroneous relative energies of ice crystal structures,[20] under-estimates the equilibrium density of the bulk liquid by ∼20%,[23–25] and makes the liquid over-structured and under-diffusive.[16,23–25] Other common DFT approximations, such as PBE,[26] suffer from comparable errors.[16,23,27] The rigorous many-body expansion[28,29] is helpful in analyzing DFT errors of molecular systems. In this scheme, the total energy $E_{tot}(1, 2, \ldots N)$ of a system of N monomers is

expressed as

$$E_{tot}(1, 2, \ldots N)$$
$$= \sum_i E^{(1)}(i) + \sum_{i<j} E^{(2)}(i, j) + E^{(>2)}(1, 2, \ldots N). \quad (1)$$

Here, i denotes the collection of variables describing the monomer i (position, orientation, internal distortion from equilibrium geometry), $E^{(1)}(i)$ is the 1-body (1B) energy of monomer i, $E^{(2)}(i, j)$ is the 2-body (2B) interaction energy of monomers i and j, and the beyond-2-body (B2B) energy $E^{(>2)}$ is everything not included in 1B and 2B energy. The importance of B2B energy is clear from the fact that the dipole moment of the H_2O monomer increases greatly (from 1.86 D to ∼2.6 D) on going from the gas phase to condensed phases,[30,31] so that the interaction between a pair of molecules is strongly affected by the presence of other molecules. B2B errors can be an important cause of inaccuracies in DFT,[5,32,33] but their effect is likely to become fully apparent only for rather large H_2O aggregates, and this is why accurate benchmarks for such aggregates are essential.[34]

There is now abundant evidence that the diffusion Monte Carlo (DMC) form of QMC approaches coupled-cluster accuracy for water systems. This was already indicated by comparisons with highly converged CCSD(T) calculations (coupled cluster with single and double excitations and perturbative triples) on the water dimer and other small clusters,[19,32,35] and a recent many-body analysis showed that DMC is accurate for the separate 1B, 2B, and B2B parts of the energy.[32] Its high accuracy is confirmed by recent calculations on ice structures,[11] which give binding energies within ∼0.2 mE_h (≃5 meV, ≃0.1 kcal/mol) per monomer of experimental values, as well as accurate equilibrium volumes and transition pressures. Furthermore, DMC can be used for both clusters and periodic systems, its basis-set convergence is rapid and automatic,[36] and its computational scaling with
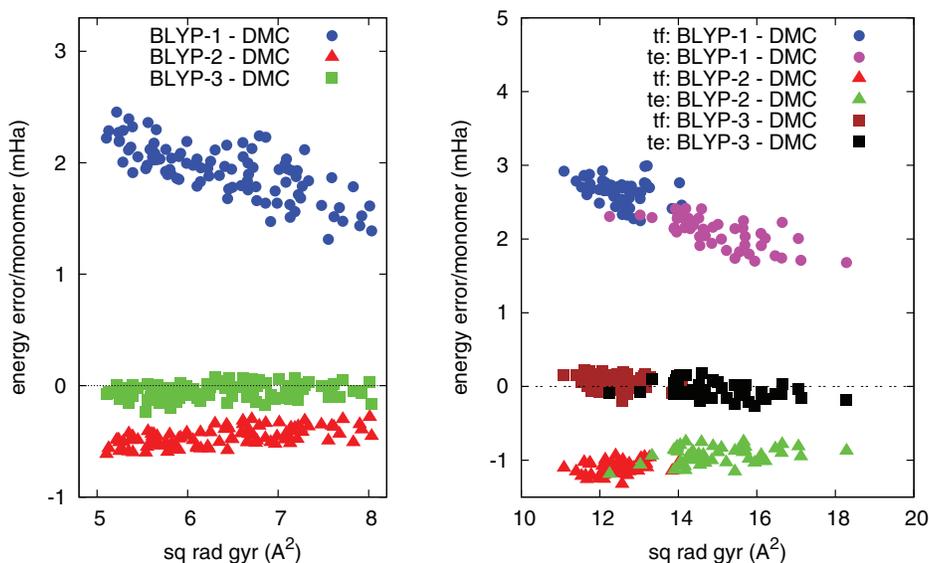
FIG. 1. Energy errors of BLYP functional after correction for 1-body (BLYP-1), 1- and 2-body (BLYP-2) and 1-, 2-, and 3-body (BLYP-3) errors for thermal samples of 100 configurations of the $(H_2O)_6$ (left panel) and $(H_2O)_{15}$ (right panel) clusters. Errors (deviations from DMC benchmarks) are plotted against squared radius of gyration R_{gyr}^2 of the cluster. Configurations are from a single simulation with the TTM3-F model[40] for the 6-mer but from two separate simulations (identified as tf and te) for the 15-mer. Units: mE_h/monomer.

system size is mild.[13,14] These facts give confidence in using DMC to generate benchmark energies for large water aggregates, including the liquid. In order to use our DMC benchmarks to analyse the errors of BLYP, we will use recently developed techniques based on Gaussian approximation potentials (GAPs)[37–39] to correct the 1B and 2B errors of this functional. The GAP techniques[37,38] allow the accurate representation (to better than 50 μE_h) of the differences between benchmark and DFT values of $E^{(1)}(i)$ and $E^{(2)}(i,j)$ for general monomer and dimer configurations. By adding these GAP corrections to the total DFT energy of any water system, we almost completely eliminate the 1B and 2B errors of the functional, so that all remaining deviations from the DMC benchmarks are due to its B2B errors.[38,39] We will denote by BLYP-1 and BLYP-2 the BLYP approximation corrected for 1-body errors and for both 1- and 2-body errors, respectively.

Our idea in studying thermal-equilibrium nano-droplets is that as we increase their size they should come to resemble the bulk liquid, so that we can systematically follow the evolution of DFT errors as we go from small clusters towards the bulk. Starting with the hexamer, we generate random samples of configurations typical of thermal equilibrium by performing molecular dynamics (m.d.) simulations using an empirical interaction model. We compute benchmark energies of these configurations using DMC, and use them to assess the errors of BLYP. As expected from earlier work,[32,38] these errors are large. We then use GAP to correct almost exactly for the 1B and 2B errors, so that the remaining differences from DMC are $E^{(>2)}$ or B2B errors. These B2B errors are significant, and grow with nano-droplet size. We then turn to a similar analysis for samples of configurations of the bulk liquid in periodic boundary conditions, again using DMC benchmarks to analyze BLYP errors into their 1B, 2B, and B2B components, and we examine their relation with the errors found for the nano-droplets. Our m.d. simulations on nano-droplets were

performed with the flexible and polarizable TTM3-F interaction model.[40] We performed our m.d. simulations on clusters at $T = 200$ K, which is high enough to ensure exploration of a wide range of configurations. In order to prevent the evaporation that would otherwise occur, we confine the molecules to the neighbourhood of a fixed point in space by a weak harmonic potential centered at this point.

Figure 1 shows the errors (deviations from DMC benchmarks) of BLYP-1 and BLYP-2 for samples of 100 configurations of the $(H_2O)_6$ and $(H_2O)_{15}$ nanodroplets. (Further details of the configuration samples and the DFT calculations, and results for $(H_2O)_9$ are in the supplementary material.[15]) The errors are plotted against the squared radius of gyration R_{gyr}^2 of the cluster, defined as $R_{gyr}^2 = N^{-1} \sum_n |\mathbf{r}_n - \bar{\mathbf{r}}|^2$, with N the number of monomers, $\mathbf{r}_n$ the O position of monomer n, and $\bar{\mathbf{r}} = N^{-1} \sum_n \mathbf{r}_n$ the centroid of O positions. The errors of BLYP-1 are positive, as expected from the well-known underbinding of BLYP,[32] but they decrease markedly with increasing R_{gyr}^2, so that the errors make the nano-droplets expand incorrectly. The BLYP-2 approximation is somewhat overbound, and its errors become more negative with decreasing R_{gyr}^2, so that they make the droplet contract. These results are consistent with our recent many-body analysis of the energetics of the hexamer.[32]

We comment briefly on the possibility of correcting the B2B errors of BLYP-2. Very recent work on large configuration samples of the water trimer shows that the BLYP 3-body energy systematically overbinds,[33] and it is natural to ask whether the B2B overbinding in the nano-droplets can be attributed to 3-body effects. To answer this, we constructed an approximate algorithm to represent short-range 3-body errors, following the methods of Paesani and co-workers,[5,33] parameterizing it by least-squares fitting to data for a large thermal sample of trimer configurations (see the supplementary material[15]). Correction of BLYP-2 using this algorithm leaves

only very small errors (deviations from DMC), as shown in Fig. 1, so that the B2B errors of BLYP do appear to be mainly 3-body errors for our nano-droplets. This merits further study.

Many DFT-based simulations have highlighted the subtle balance between low-density and high-density structures in liquid water.[24,25] To explore this balance, and its relation with our results for nano-droplets, we have computed DMC benchmark energies for configuration samples drawn from m.d. simulations of the liquid, in order to make the same many-body analysis as for the nano-droplets. To ensure the technical reliability of our DMC benchmarks, we have carefully checked for system-size errors, using configurations from 32- and 64-molecule m.d. simulations of the liquid. These tests (see the supplementary material for details[15]) show that the errors of BLYP and BLYP-2 can be reliably gauged by comparing with DMC benchmarks on 32-molecule liquid configurations. Our many-body analysis of the errors of BLYP for the liquid uses configurations drawn from two DFT m.d. simulations, one performed with the BLYP functional and the other with BLYP-2, both simulations being performed on 32-molecule systems at 350 K; the densities were chosen to give pressures close to zero (see the supplementary material[15]), and correspond to 0.778 and 1.049 g/cm^3, respectively. (These under- and over-estimates of the experimental density at 350 K by 20% and 7% are expected with the BLYP and BLYP-2 approximations.) The O–O radial distribution functions $g_{OO}(r)$ (Fig. 2 presents results from 64-molecule simulations) show, as expected from earlier work, that the higher-density BLYP-2 liquid is less ordered than the lower-density BLYP liquid, the very different values of $g_{OO}(r)$ at the first minimum and the different oscillation amplitudes at larger separations being particularly noteworthy. The comparison in Fig. 2 shows that BLYP-2 gives a liquid structure in much better agreement with experiment[41] than BLYP itself.

We drew samples of 10 configurations at random from each of the two simulations, and used DMC to compute benchmark energies for them. Their BLYP energies come di-

FIG. 3. Energy errors of the BLYP (red, green symbols) and BLYP-2 (blue, magenta) total-energy functions for two sets of configurations of liquid water. Errors (deviations from DMC benchmarks) are shown for 10 configurations each drawn from m.d. simulations performed near zero pressure with the BLYP (squares) and BLYP-2 (circles) total-energy functions.

rectly from the simulations, and the GAP representation is used to compute the 1- and 2-body corrections, exactly as for the nano-droplets. The errors of BLYP and BLYP-2 (Fig. 3) reveal a pattern that quantitatively resembles what we saw for the nano-droplets. Uncorrected BLYP shows substantial underbinding, with the errors decreasing substantially from high to low density, so that they wrongly drive the system to lower densities. The approximation BLYP-2 is overbound by between 1.2 and 2.0 mE_h/monomer, in accord with the BLYP-2 errors for the nano-droplets. The trend of the errors with BLYP-2 is consistent with its slight overestimate of the equilibrium density.

In summary, we have demonstrated that QMC can give valuable energy benchmarks for water aggregates ranging all the way from small clusters to liquid systems of up to 64 molecules.[42] Previous QMC work on water clusters and ice suggests that the accuracy of QMC is similar to that of CCSD(T) near the basis-set limit, residual errors being of order 0.2 mE_h ($\simeq$5 meV, $\simeq$0.1 kcal/mol) per monomer. To illustrate the usefulness of such benchmarks, we have used them to analyze the errors of the popular DFT functional BLYP, with accurate GAP-based representations of the 1- and 2-body errors of BLYP used to assess also its beyond-2-body errors. We have been able to trace the evolution of these errors from the dimer to the bulk liquid. The BLYP 2-body under-binding familiar from earlier work on the water dimer leads directly to a large net under-binding for the nano-droplets and the liquid, and causes all these systems to expand, giving rise to the well-known ~20% under-estimation of the equilibrium density of the liquid. However, correction for this 2-body error (and for 1-body errors) gives an approximation (BLYP-2) that over-binds for the nano-droplets and the liquid, and causes a small erroneous contraction, as noted elsewhere for ice structures. We have seen how the over-binding of BLYP-2, at least

FIG. 2. Oxygen-oxygen radial distribution function $g_{OO}(r)$ from simulations of liquid water (64 molecules in repeating cell) at $T = 350$ K performed with BLYP (dashed green curve) and BLYP-2 (solid red curve) approximations, compared with data from high-energy x-ray diffraction at 343 K[41] (dotted blue curve). The BLYP and BLYP-2 simulations were performed at densities 0.778 and 1.049 g/cm^3, respectively (see text).

for the nano-droplets, can be related to the well established 3-body errors of BLYP.

Several immediate applications of the present methods appear promising. The combination of QMC benchmarks with 1- and 2-body GAP correction for a range of cluster and extended water systems should provide a practical strategy for analyzing the errors of any chosen DFT approximation. The types of water aggregate need not be limited to those studied here, but may include the liquid and solid at different thermodynamic states, as well as extended surface systems. The idea of applying the present methods to a range of other molecular systems, for example, aqueous solutions or mixtures such as the water-methane system, is also attractive. We are now investigating some of these possibilities.

[1] H. Popkie, H. Kistenmacher, and E. Clementi, J. Chem. Phys. **59**, 1325 (1973).

[2] K. Laasonen, M. Sprik, M. Parrinello, and R. Car, J. Chem. Phys. **99**, 9080 (1993).

[3] R. Bukowski, K. Szalewicz, G. C. Groenenboom, and A. van der Avoird, Science **315**, 1249 (2007).

[4] Y. Wang, X. Huang, B. C. Shepler, B. J. Braams, and J. M. Bowman, J. Chem. Phys. **134**, 094509 (2011).

[5] V. Babin, G. R. Medders, and F. Paesani, J. Phys. Chem. Lett. **3**, 3765 (2012).

[6] W. Klopper, J. G. C. M. van Duijenveldt-van de Rijdt, and F. B. van Duijenveldt, Phys. Chem. Chem. Phys. **2**, 2227 (2000).

[7] G. S. Tschumper, M. L. Leininger, B. C. Hoffman, E. F. Valeev, H. F. Schaefer, and M. Quack, J. Chem. Phys. **116**, 690 (2002).

[8] D. M. Bates and G. S. Tschumper, J. Phys. Chem. A **113**, 3555 (2009).

[9] S. Yoo, A. Aprà, X. C. Xeng, and S. S. Xantheas, J. Phys. Chem. Lett. **1**, 3122 (2010).

[10] U. Góra, R. Podeszwa, W. Cencek, and K. Szalewicz, J. Chem. Phys. **135**, 224102 (2011).

[11] B. Santra, J. Klimeš, D. Alfè, A. Tkatchenko, B. Slater, A. Michaelides, R. Car, and M. Scheffler, Phys. Rev. Lett. **107**, 185701 (2011).

[12] P. J. Bygrave, N. L. Allan, and F. R. Manby, J. Chem. Phys. **137**, 164102 (2012).

[13] W. M. C. Foulkes, L. Mitaš, R. J. Needs, and G. Rajagopal, Rev. Mod. Phys. **73**, 33 (2001).

[14] R. J. Needs, M. D. Towler, N. D. Drummond, and P. López-Ríos, J. Phys. Condens. Matter **22**, 023201 (2010).

[15] See supplementary material at http://dx.doi.org/10.1063/1.4810882 for computational details.

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

[17] A. D. Becke, Phys. Rev. A **38**, 3098 (1988); C. Lee, W. Yang, and R. Parr, Phys. Rev. B **37**, 785 (1988).

[18] B. Santra, A. Michaelides, and M. Scheffler, J. Chem. Phys. **127**, 184104 (2007).

[19] B. Santra, A. Michaelides, M. Fuchs, A. Tkatchenko, C. Filippi, and M. Scheffler, J. Chem. Phys. **129**, 194111 (2008).

[20] O. Kambara, K. Takahashi, M. Hayashi, and J. L. Kuo, Phys. Chem. Chem. Phys. **14**, 11484 (2012).

[21] H.-S. Lee and M. E. Tuckerman, J. Chem. Phys. **125**, 154507 (2006).

[22] S. Yoo, X. C. Xeng, and S. S. Xantheas, J. Chem. Phys. **130**, 221102 (2009).

[23] J. Schmidt, J. VandeVondele, I.-F. W. Kuo, D. Sebastiani, J. I. Siepmann, J. Hutter, and C. J. Mundy, J. Phys. Chem. B **113**, 11959 (2009).

[24] J. Wang, G. Román-Pérez, J. M. Soler, E. Artacho, and M.-V. Fernández-Serra, J. Chem. Phys. **134**, 024516 (2011).

[25] Z. Ma, Y. Zhang, and M. E. Tuckerman, J. Chem. Phys. **137**, 044506 (2012).

[26] J. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

[27] P. H.-L. Sit and N. Marzari, J. Chem. Phys. **122**, 204510 (2005).

[28] S. S. Xantheas, J. Chem. Phys. **100**, 7523 (1994).

[29] J. M. Pedulla, F. Vila, and K. D. Jordan, J. Chem. Phys. **105**, 11091 (1996).

[30] C. A. Coulson and D. Eisenberg, Proc. R. Soc. London, Ser. A **291**, 445 (1966).

[31] P. L. Silvestrelli and M. Parrinello, Phys. Rev. Lett. **82**, 3308 (1999).

[32] M. J. Gillan, F. R. Manby, M. D. Towler, and D. Alfè, J. Chem. Phys. **136**, 244105 (2012).

[33] G. R. Medders, V. Babin, and F. Paesani, J. Chem. Theory Comput. **9**, 1103 (2013).

[34] H. R. Leverentz, H. W. Qi, and D. G. Truhlar, J. Chem. Theory Comput. **9**, 995 (2013).

[35] I. G. Gurtubay and R. J. Needs, J. Chem. Phys. **127**, 124306 (2007).

[36] D. Alfè and M. J. Gillan, Phys. Rev. B **70**, 161101(R) (2004).

[37] A. Bartók, M. C. Payne, R. Kondor, and G. Csányi, Phys. Rev. Lett. **104**, 136403 (2010).

[38] A. P. Bartók, M. J. Gillan, F. R. Manby, and G. Csányi, "Machine learning for predictive condensed-phase simulation," Phys. Rev. B (submitted) [preprint arXiv:1302.5680].

[39] M. J. Gillan, D. Alfè, A. P. Bartók, and G. Csányi, "First-principles energetics of water: A many-body analysis," Phys. Rev. B (submitted) [preprint arXiv:1303.0751].

[40] G. S. Fanourgakis and S. S. Xantheas, J. Chem. Phys. **128**, 074506 (2008).

[41] L. B. Skinner, C. J. Benmore, J. Neuefeind, and J. B. Parise, "A structural description of the compressibility minimum in water" (unpublished).

[42] See www.homepages.ucl.ac.uk/ucfbdxa/qmc.htm for atomic coordinates and DMC benchmark energies of the nano-droplet and liquid-state configurations studied in this work.