Machine-learning approach for one- and two-body corrections to density functional theory: Applications to molecular and condensed water

Albert P. Bartók,[1] Michael J. Gillan,[2,3,4] Frederick R. Manby,[5] and Gábor Csányi[1]

[1]*Engineering Laboratory, University of Cambridge, Trumpington Street, Cambridge, CB2 1PZ, United Kingdom*
[2]*London Centre for Nanotechnology, University College London, Gordon St., London WC1H 0AH, United Kingdom*
[3]*Department of Physics and Astronomy, University College London, Gower St., London WC1E 6BT, United Kingdom*
[4]*Thomas Young Centre, University College London, Gordon St., London WC1H 0AH, United Kingdom*
[5]*Centre for Computational Chemistry, School of Chemistry, University of Bristol, Bristol BS8 1TS, United Kingdom*

(Received 16 April 2013; revised manuscript received 11 June 2013; published 8 August 2013)

We show how machine learning techniques based on Bayesian inference can be used to enhance the computer simulation of molecular materials, focusing here on water. We train our machine-learning algorithm using accurate, correlated quantum chemistry, and predict energies and forces in molecular aggregates ranging from clusters to solid and liquid phases. The widely used electronic-structure methods based on density functional theory (DFT) by themselves give poor accuracy for molecular materials like water, and we show how our techniques can be used to generate systematically improvable one- and two-body corrections to DFT with modest extra resources. The resulting corrected DFT scheme is considerably more accurate than uncorrected DFT for the relative energies of small water clusters and different ice structures and significantly improves the description of the structure and dynamics of liquid water. However, our results for ice structures and the liquid indicate that beyond-two-body DFT errors cannot be ignored, and we suggest how our machine-learning methods can be further developed to correct these errors.

DOI: 10.1103/PhysRevB.88.054104 PACS number(s): 31.15.bw, 31.50.−x, 71.15.Pd

I. INTRODUCTION

The computer simulation of materials has become an indispensable tool across a wide range of disciplines, including materials physics and chemistry, metallurgy, the earth sciences, surface science, and biology. Simulation techniques range all the way from simple empirical force fields to the electronic structure methods based on density functional theory (DFT) and correlated quantum chemistry.[1] Electronic structure methods are capable of much greater accuracy and generality than force fields, but their computational demands are heavier by many orders of magnitude. A crucial challenge for simulation is therefore to find systematically improvable methods for casting information from accurate electronic-structure techniques into forms that are more rapidly computable. We show here how machine learning techniques[2] allow data from correlated quantum chemistry to be used to correct the DFT description of molecular materials, taking condensed-phase water as our example.

The fundamental interactions in water and other molecular materials[3] consist of exchange-repulsion, electrostatic interaction between molecular charge distributions, polarization (i.e., the electrostatic distortion of charge distributions), charge transfer, and van der Waals dispersion, together with effects due to molecular flexibility. Electron correlation plays a role in all these, and is crucial for dispersion.[4] The correlated quantum chemistry methods of MP2 (second-order Møller-Plesset) and particularly CCSD(T) (coupled-cluster with single and double excitations and perturbative triples, often referred to as the "gold standard" of molecular quantum chemistry)[5] give a very accurate description of these interactions,[6,7] but their heavy computational demands for extended systems make their routine use for condensed matter problematic. DFT techniques are less demanding and have been widely used for water,[8] but the results obtained with standard approximations often agree poorly with experiment[9] and may depend strongly on the assumed approximation.[10] There is vigorous debate about how to overcome the problems of DFT, and our point of view here is that input from correlated quantum chemistry is essential. We shall describe an approach in which machine learning[2] is used to construct representations of some of the main energy differences between correlated quantum chemistry and DFT, which can then be used to construct efficient corrected DFT schemes for simulation of large, complex systems. Our machine learning methods are partly based on the reported ideas of Gaussian Approximation Potentials (GAP),[11–13] and are also related to the way Gaussian processes were used recently to model the atomization energies of small molecules.[14]

For molecular materials, it is helpful to work with the widely used many-body representation,[15,16] in which the total energy $E_{\rm tot}(1,2,\ldots,N)$ of a system of N molecules is separated into one-body, two-body, and beyond-two-body parts:

$$E_{\rm tot}(1,\ldots N) = \sum_{i=1}^{N} E_{\rm 1B}(i) + \sum_{i<j} E_{\rm 2B}(i,j) + E_{\rm B2B}(1,\ldots N)\,. \quad (1)$$

Here, $E_{\rm 1B}(i)$ is the one-body (1B) energy of molecule i in free space, which depends on its distortion away from its equilibrium configuration. The energy $E_{\rm 2B}(i,j)$ is the two-body (2B) interaction energy of the pair of molecules (i,j) in free space, i.e., the total energy of the pair minus the sum of their 1B energies. For water, $E_{\rm 2B}(i,j)$ is a function of 12 variables specifying the separation of the molecules, their relative orientation and their internal distortions. We have grouped together the three-body (3B) and higher-body terms into the beyond-two-body (B2B) energy $E_{\rm B2B}$, which represents everything not accounted for by 1B and 2B energies.

Exchange-repulsion and first-order electrostatics are mainly or entirely 2B interactions, and in most molecular systems, E_{B2B} arises mainly from polarization. In water systems, dispersion is mainly a 2B effect,[3] though B2B contributions may not be completely negligible,[17,18] but in some molecular systems, B2B contributions to dispersion can be substantial.[19]

Our strategy of using machine learning to represent the corrections to DFT needed to obtain CCSD(T) accuracy can, in principle, be applied to all terms in the many-body series, but at the present stage, we use it only to correct the 1B and 2B terms in the energy. This means that we shall obtain an accurate description if we start from a DFT approximation that is accurate enough for B2B energy. It has long been known that there is a strong redistribution of electrons when a water monomer enters the liquid or solid phases, and its dipole moment increases from 1.86 D in the gas phase to ∼2.6 D or more in condensed phases.[20,21] The inclusion of B2B effects is therefore crucial. However, the polarizabilities of the H_2O monomer are generally quite well described by common DFT approximations (an overestimation by ∼10% is normal with nonhybrid GGA approximations, with hybrid functionals being somewhat better[22,23]) so that use of DFT for B2B energy should serve as a reasonable starting point. We note that our approach incorporates an accurate description of molecular flexibility, which is not always done even in some sophisticated force fields. Flexibility is essential, because without it one could not describe the well known lengthening of O-H bonds and the lowering of O-H vibrational frequency[24] when the H atom of a water monomer participates in a hydrogen bond with another monomer and nor could quantum nuclear effects[25] be properly treated.

The idea of working within the framework of the many-body expansion and using data from correlated quantum chemistry to develop accurate models for the 1B and 2B (and recently 3B) parts of the energy, with more approximate models used for the higher-body terms, goes back several years. This general approach has been extensively used by many groups (see, e.g., Refs. 26 and 27). By contrast with early unpolarizable models,[28] which were designed to work only for a limited range of liquid states, developments within the many-body framework have always emphasized the importance of B2B energy, which has often been represented by single- or multisite polarizabilities, usually damped at short distances. As examples of this approach, we note the series of TTM*n*-F models due to the groups of Xantheas and Burnham,[29–31] and the DPP*n* models of Jordan's group.[32] In recent years, systematic methods have been proposed for creating accurate parameterized representations of the energetics of the water dimer, using databases of CCSD(T) energies for large numbers of geometries.[26,33,34] These methods were subsequently extended to the water trimer, to produce representations of the three-body energy computed using correlated quantum-chemistry techniques.[26,27,33–36] Combined with more approximate (but still *ab initio* based) models for the higher-body energies, these approaches have achieved important successes in describing the properties of water in a variety of aggregation states from small clusters to the bulk liquid.[26,27] A noteworthy example is the very recent work from Paesani's group,[27,36] in which 1B energy is described by the very accurate Partridge-Schwenke algorithm,[37] 2B energy is represented by the HBB2 *ab initio*-parameterized model,[38] 3B energy is given by the TTM4-F model of Burnham *et al.*[31] corrected using an efficient parameterized representation of CCSD(T) 3B energies, and higher-body energies are all described by TTM4-F.

The work presented here differs from what has been done before in the following respects. First, our systematically improvable representation of energies and forces employs machine-learning techniques based on Bayesian inference in the general framework of Gaussian processes, rather than using predefined functional forms (e.g., polynomials up to a fixed degree). We shall show that this approach allows us to reproduce CCSD(T) energies with very high accuracy. Second, we use DFT rather than parameterized models for the higher-body energies. We are motivated to start from DFT by our long-term aim of developing methods for general molecular systems, including aqueous solutions of ions and small molecules. DFT is, of course, a more expensive simulation technique than parameterized force-fields, but it has the enormous advantage of not requiring a development effort every time it is applied to a new system, and this is why we adopt the strategy of using machine learning combined with correlated quantum chemistry to make systematic improvements to DFT. There is an interesting relation here with the DFT/CC method of Nachtigall *et al.*,[39,40] which is based on the idea of correcting DFT approximations for molecular systems by adding parameterized atom-atom potentials fitted to reproduce CCSD(T) energies. The approach we present here differs by using nonparametric machine learning to generate systematically improvable corrections. We note that machine learning based on neural-network techniques has recently been used to represent the energetics of the H_2O dimer[41] and more recently some larger clusters,[42] but the aim there was completely different, namely, to reproduce the DFT energies themselves directly, rather than accurate quantum-chemistry energies, and without a many-body decomposition.

In the next section, we present our machine-learning techniques for correcting the 1B and 2B parts of the energy for water systems. In the subsequent sections, we report calculations on small water clusters, on polymorphs of ice, and on liquid water, the technical details of the calculations being summarized in Sec. III and the results themselves being presented and discussed in Sec. IV. These will show that the GAP-based corrections achieve a very substantial improvement over the DFT approximation on which they are based.

II. MACHINE LEARNING WITH GAP

We start this section by outlining how we use the principles of Bayesian inference in the framework of Gaussian processes to construct our machine-learning scheme for representing corrections to DFT energies. Passing to practicalities, we then note the considerations that motivated our choice of DFT functional to be corrected, and we then describe how we computed the 1B and 2B corrections to the chosen functional. Results presented at the end of the section demonstrate the high quality of the GAP correction of the 2B energy.

A. Overview of the GAP scheme

Consider a system whose configuration is specified by points $\mathbf{R}$ in a many-dimensional configuration space. We are given the values $f(\mathbf{R}_n)$ of its energy (or corrections to its energy) at a finite set of configurations $\{\mathbf{R}_n\}$. We now ask what is the most likely value of $f(\mathbf{R})$ at a configuration $\mathbf{R}$ not in the given set $\{\mathbf{R}_n\}$? The rules of Bayesian inference[2] are used to compute this most likely value, assuming that the function f has certain smoothness properties. Here, the concept of smoothness is used in the sense of spatial correlation, and simply means that the probability of finding very different values $f(\mathbf{R})$ and $f(\mathbf{R}')$ decreases rapidly to zero as $\mathbf{R}$ and $\mathbf{R}'$ approach each other. The framework of GAP, based on Gaussian processes,[43] uses a precise formulation of smoothness in terms of a covariance function $C(\mathbf{R},\mathbf{R}')$ having the form[43]

$$C(\mathbf{R},\mathbf{R}') = \theta \exp\left\{-\sum_i [(R_i - R_i')/(2\sigma_i)]^2\right\}, \quad (2)$$

where the sum in the exponent is over the dimensions of the configuration space, θ is the typical scale of f and σ_i are the typical length scales on which $f(\mathbf{R})$ varies. The theory yields the following formula[43] for the most likely estimate of $f(\mathbf{R})$ given the data and the assumption of smoothness (often called the maximum *a posteriori* estimator):

$$f(\mathbf{R}) = \sum_n^{\text{data}} C(\mathbf{R},\mathbf{R}_n)\alpha_n \ , \quad (3)$$

where the coefficients α_n are given by inversion of the linear equations

$$f(\mathbf{R}_m) = \sum_n^{\text{data}} [C(\mathbf{R}_m,\mathbf{R}_n) + \varepsilon\delta_{mn}]\alpha_n \ , \quad (4)$$

where δ_{mn} is the Kronecker δ and the diagonal shift of magnitude ε is included to regularize the linear algebra.

When applying GAP to represent corrections to 1B and 2B energies in water, there are different ways of choosing the space in which the configurations are represented, but here it is advantageous to build in the fact that the energy function $f(\mathbf{R})$ is left unchanged by rotations and translations of the whole system, and by interchange of identical atoms. For the water monomer, the two OH distances and the angle between them provide a convenient coordinate system. For the water dimer, we ensure rotation and translation symmetry by working with the space of the 15 interatomic distances, $\mathbf{R} = \{|\mathbf{r}_i - \mathbf{r}_j|\}$, where $\mathbf{r}_i$ are the atomic positions. To ensure interchange symmetry, we symmetrize the covariance function over permutations of identical atoms and also impose a finite cutoff range:

$$\tilde{C}(\mathbf{R},\mathbf{R}') = \frac{1}{|S|}\sum_{\pi\in S} C(\pi(\mathbf{R}),\mathbf{R}') f_{\text{cut}}(R_{\text{OO}}) f_{\text{cut}}(R'_{\text{OO}}), \quad (5)$$

where π represents an element of the permutation group, S, of the water dimer, whose order $|S|$ is 8, and f_{cut} is a cutoff function whose value changes from unity to zero smoothly as R_{OO} and R'_{OO}, the distances between the pairs of oxygen atoms in the two dimer configurations $\mathbf{R}$ and $\mathbf{R}'$, approach a predefined limit. A more detailed account of our GAP formalism is given elsewhere.[11,12] The computational cost of evaluating the GAP model is linear in the size of the database $\{\mathbf{R}_n\}$, and for the case of the water dimer presented below takes about 10 ms on a single processor.

B. Choice of DFT functional

Because we wish to correct 1B and 2B energies, in selecting a DFT functional we concentrate on its accuracy for the B2B energies. In a recent study[46] on thermal samples of configurations of a range of small water clusters up to the hexamer, it was shown that for several DFT approximations the correction of one- and two-body errors does indeed yield major improvements in accuracy. Some of the DFT functionals studied there were "hybrid" functionals in which a fraction of exact exchange is included in the exchange-correlation functional. For simulations of large aggregates of water molecules, and particularly for molecular dynamics simulations, it is preferable not to use hybrid functionals, because they are computationally very demanding. Of the nonhybrid functionals studied,[46] BLYP was the one that gave the smallest root-mean-square deviations from benchmark values for the B2B energies, which is consistent with other recent work,[36] and this is why we chose to employ BLYP in the present work.

To illustrate the fact that a number of DFT functionals give a good representation of B2B energy, we show in Fig. 1 a parity plot in which the three-body energies of 50 configurations of the H_2O trimer computed with four DFT functionals are plotted against highly converged CCSD(T) benchmarks. For comparison, we also show the three-body energy computed with the polarizable and flexible interaction model TTM3-F for water published recently by the group of Xantheas[30] and for which computer code is publicly available.

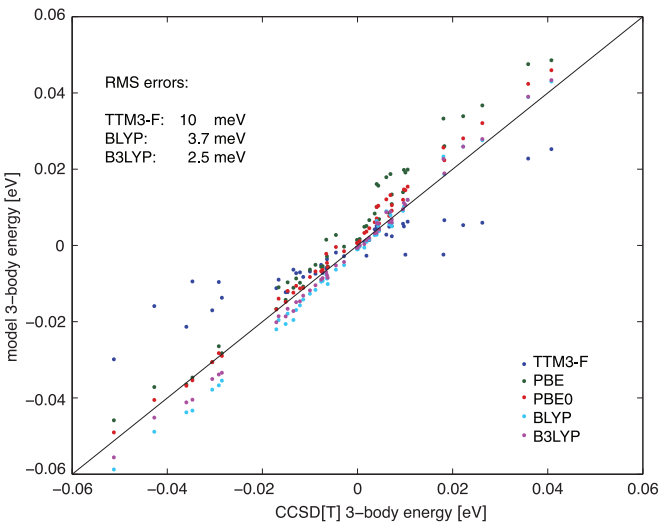

FIG. 1. (Color online) Parity plot of three-body interaction energies in a thermal sample of water trimers: approximations are compared with CCSD(T) benchmarks. Approximations shown are TTM3-F force field[30] and DFT with PBE, PBE0, BLYP, and B3LYP functionals. The thermal sample was generated by drawing trimers from an MD simulation of the bulk liquid performed using the AMOEBA forcefield[44,45] at 300 K and ambient pressure.

We note that the functionals BLYP, PBE, B3LYP, and PBE0 all reproduce the three-body energy with an accuracy that considerably exceeds that of TTM3-F. This comment is not, of course, intended as a criticism of TTM3-F, which is a highly successful model. It should be noted that the revised force field TTM4-F[31] improves the accuracy of three-body terms over its predecessor.[36]

C. One-body corrections

The energy of an isolated H_2O monomer as a function of its two O-H bond lengths r_1, r_2, and its H-O-H angle θ is very accurately represented using the parameterization due to Partridge and Schwenke (PS).[37] In order to correct the DFT energy of a water system for one-body errors, we have only to add to the DFT total energy the difference

$$\Delta E_{1\mathrm{B}} \equiv E_{1\mathrm{B}}(\mathrm{PS}) - E_{1\mathrm{B}}(\mathrm{DFT})$$

for every monomer. Since $\Delta E_{1\mathrm{B}}$ for a monomer depends on only three variables r_1, r_2, θ, many possible schemes could be used to represent $\Delta E_{1\mathrm{B}}(r_1,r_2,\theta)$ accurately. It is convenient here to use a GAP-like approach, and the method we employ uses a GAP representation for $E_{1\mathrm{B}}(\mathrm{BLYP})$ itself, so that at run-time we compute $E_{1\mathrm{B}}(\mathrm{PS})$ by the standard PS formula and $E_{1\mathrm{B}}(\mathrm{BLYP})$ from the GAP representation and then add the difference $\Delta E_{1\mathrm{B}}$ for every monomer to the total BLYP energy. Naturally, a direct GAP representation of $\Delta E_{1\mathrm{B}}$ could also be used.

To make the GAP representation of $E_{1\mathrm{B}}(\mathrm{BLYP})$, we use the MOLPRO package[47] with an aug-cc-pV5Z basis to compute the BLYP monomer energies very accurately on a uniform grid in the space of r_1, r_2 and θ (0.8 Å $< r_1,r_2 <$ 1.15 Å with 0.025 Å spacing, $72.5^\circ < \theta < 127.5^\circ$ with 5° spacing). The three symmetrized coordinates $r_1 + r_2$, $(r_1 - r_2)^2$ and $\mathbf{r}_1 \cdot \mathbf{r}_2$ are used in the GAP scheme.

D. Two-body corrections

To correct the BLYP total energy of a water system for two-body errors, we add the difference

$$\Delta E_{2\mathrm{B}} \equiv E_{2\mathrm{B}}\big(\mathrm{CCSD(T)}\big) - E_{2\mathrm{B}}(\mathrm{BLYP})$$

for all distinct monomer pairs. Our GAP representation of the two-body correction $\Delta E_{2\mathrm{B}}$ is constructed in two steps: first, we make a representation of the difference

$$\Delta E_{2\mathrm{B}}(\mathrm{MP2/aug\text{-}cc\text{-}pVTZ}) \equiv E_{2\mathrm{B}}(\mathrm{MP2/aug\text{-}cc\text{-}pVTZ}) - E_{2\mathrm{B}}(\mathrm{BLYP}),$$

where $E_{2\mathrm{B}}$(MP2/aug-cc-pVTZ) is the two-body energy calculated with the MP2 approximation and the moderately accurate aug-cc-pVTZ basis set; in the second step, we represent the difference $\Delta E_{2\mathrm{B}}(\delta\mathrm{CCSD(T)},\delta\mathrm{basis})$, which corrects both for the basis-set errors in $\Delta E_{2\mathrm{B}}$(MP2/aug-cc-pVTZ) and for the differences between CCSD(T) and MP2. All dimer calculations are done with the Molpro package,[47] and we always use counterpoise correction to suppress basis-set superposition errors. It is important to note our aim of representing the difference between CCSD(T) and BLYP as close as possible to the basis-set limit for both methods in both one-body and two-body corrections. We do this because we intend that the corrections will generally be applied to well converged plane-wave DFT calculations. This procedure is consistent with the principle noted in recent work[48] that the basis sets used in the generation of corrections to DFT should be the same as those employed when the corrections are applied.

To generate *ab initio* data for correcting DFT to MP2 level, we took 6000 water dimer configurations with $R_{\mathrm{OO}} <$ 4.5 Å and 1000 configurations with 4.5 Å $< R_{\mathrm{OO}} <$ 6.0 Å from an AMOEBA[44,45] molecular dynamics simulation of a large periodic liquid water system at 300 K. In order to achieve coverage across a wider range of configurations, the data set was augmented by 2040 configurations from a DFT/BLYP/aug-cc-pVTZ molecular dynamics simulation of a water dimer in a harmonic confining potential at 4000 K using a Langevin thermostat.[49]

Energies and forces were computed for these 9040 configurations using MP2 and BLYP with the aug-cc-pVTZ basis set. The resulting difference was fitted using the GAP framework. The basic formulation of Gaussian process regression outlined above applies to the case when the data comprises function values only. When derivatives (or equivalently forces) are also available, they should be used, since they contain much valuable information; for each dimer configuration, they supply 18 values in addition to the one value of the energy. Their treatment in the GAP framework is straightforward; one simply needs to express the covariance of any two force values and also the covariance of the observed forces with the predicted energy, and use these expressions to construct an enlarged covariance matrix that is then used in the same way as shown in equations (3) and (4). Since our covariance is a Gaussian, the covariance of the forces is just the derivative of this Gaussian with respect to the appropriate coordinate, and is easy to compute. An extended discussion and detailed formulas are given elsewhere.[11,12]

In order to correct for basis-set errors and for the errors of MP2, we exploit the fact that explicit-correlation (F12) methods greatly accelerate the basis-set convergence of correlated calculations.[50,51] Since the corrections are small, we find that it suffices to use only 1000 dimer configurations from the AMOEBA simulation mentioned above, and that energies alone (without forces) provide enough data. Our GAP correction for basis-set errors in MP2 is based on MP2-F12 computations with the aug-cc-pVTZ basis set, and the correction for the difference CCSD(T) − MP2 employs CCSD(T)-F12 and MP2-F12 calculations with the aug-cc-pVDZ basis set.

We show in Fig. 2 the two-body errors of BLYP together with the errors of GAP-corrected BLYP for a thermal sample of dimer configurations (which were not used in the construction the GAP model) drawn from a molecular dynamics simulation of liquid water. Uncorrected BLYP is too repulsive for the water dimer,[46] with unacceptably large errors of up to 50 meV at the separations of interest. However, with GAP corrections, the errors are dramatically reduced to ∼1 meV. GAP thus provides a way of virtually eliminating all errors in a chosen DFT approximation apart from those associated with B2B energy. Also shown in Fig. 2 are the errors of the approximation obtained by the popular procedure of adding the dispersion correction due to Grimme *et al.*[52] to BLYP. We note that this approximation is better than uncorrected BLYP, but is much less good than GAP-corrected BLYP.

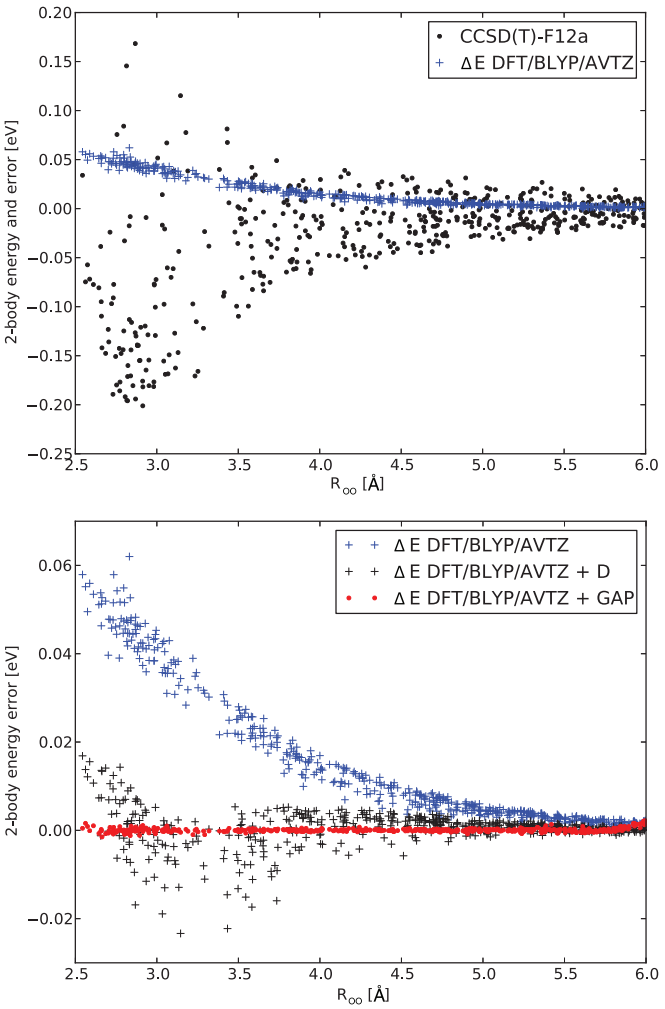

FIG. 2. (Color online) (Top) Benchmark two-body interaction (black) and errors of DFT with the BLYP functional (blue) plotted against oxygen-oxygen distance R_{OO}. Benchmarks are computed with CCSD(T) close to the basis-set limit. Bottom: two-body errors of BLYP (blue, same as in top panel) and errors of BLYP + GAP (red). The rms deviation of BLYP + GAP from benchmarks is 0.45 meV. Also shown are errors of BLYP plus Grimme D3 dispersion correction.[52] The sample of 500 dimer configurations shown here were drawn from a molecular dynamics simulation of liquid water and not used in the construction of the GAP models.

III. COMPUTATIONAL DETAILS

A. Calculations on ice structures

Our calculations on ice crystals were carried out using the CASTEP package,[53] using on-the-fly generated ultrasoft pseudopotentials, 1200 eV plane-wave cutoff and a $3 \times 3 \times 3$ Monkhorst-Pack k-point mesh in the primitive unit cell, both chosen to achieve 1 meV convergence of total energies. The geometry was relaxed (including the unit cell) until atomic displacements were smaller than 0.001 Å, and the components of the external stress tensor were less than 0.1 GPa.

B. Molecular dynamics for liquid water

Our molecular dynamics (MD) simulations on liquid water were performed with a modified version of the VASP code,[54] which includes the GAP model. The simulations employed a system of 64 D_2O molecules (heavy water) in a cubic cell with the usual periodic boundary conditions. The number of molecules per unit volume is the same as for H_2O (light water) at 0.997 g cm^{-3}, and the temperature in the production part of the run was 308 K.

Our VASP simulations employed the PAW (projector-augmented-wave) technique,[55] with a plane-wave cut-off energy of 1200 eV. The simulations using the BLYP + GAP model were performed using Born-Oppenheimer MD at constant volume and energy, corresponding to the NVE statistical-mechanical ensemble; no thermostat was employed. The value of the MD time step was 0.5 fs and at every time step convergence to the self-consistent ground state was achieved to within 1 μeV in the total energy. Under these conditions, the energy drift over a period of 10 ps was on the order of 0.3 meV/H_2O. In the BLYP + GAP simulation reported below, the system was equilibrated for 20 ps. All the results presented were computed from a subsequent production run of 25 ps. To test that the equilibration and production parts of the run were long enough, we compared the RDF from the first and second halves of the 25 ps production run, and we also varied the length of the equilibration phase from 10 to 20 ps, and observed no significant difference. To further verify that the RDF computed in the NVE ensemble does not suffer from lack of equilibration and sampling, we checked that it does not differ appreciably from the RDF obtained from a BLYP + GAP simulation in the NVT ensemble.

In addition to the main BLYP + GAP simulation, we shall also refer below to a separate simulation using the uncorrected BLYP approximation. This simulation, which was used only to calculate the liquid structure and not dynamical properties, was carried out in the NVT ensemble enforced by the Andersen thermostat, using a 2-fs time step and the same masses (16 amu) for all atoms for a total length of 10 ps. Data generated during the production of this paper, including the trajectories for the liquid water simulations and the databases used for the GAP models associated software are available at http://www.libatoms.org.

IV. RESULTS AND DISCUSSION

A. Small clusters

To illustrate the effectiveness of our GAP-corrected DFT, we start with a simple test on the ten stationary points of the water dimer.[56] These form a canonical set of configurations, which have been exhaustively studied and whose energies are extremely accurately known.[6] The global minimum structure is bound by a single hydrogen bond, but some of the less stable structures have up to four weaker hydrogen bonds.[6] We stress that none of these stationary points is included in our training set, so that the energies computed with BLYP + GAP are genuine predictions. We compare in Table I the relative energies of the 10 configurations from BLYP + GAP with the almost exact results and the predictions of the DFT functional BLYP; the table also includes the very accurate predictions of diffusion Monte Carlo (DMC).[46] As has been reported before,[57] the DFT approximation shows quite large errors of around 30 meV in some cases, while the errors of

TABLE I. Relative energies of the ten stationary points of the water dimer computed with CCSD(T) close to the basis-set limit,[6] diffusion Monte Carlo,[46] DFT with the BLYP functional, and BLYP + GAP. Numbering of the stationary states is standard.[6] Units are meV.

State	BLYP	BLYP + GAP	CCSD(T)	DMC
1	0	0	0	0
2	23	23	21	24
3	32	27	25	27
4	49	32	30	34
5	65	44	41	39
6	73	45	44	41
7	96	79	79	78
8	155	153	154	156
9	95	82	77	79
10	125	116	117	122

DMC are much smaller, being almost all less than 3 meV. Our BLYP + GAP predictions are very accurate, and indeed they compete in accuracy with DMC, at enormously reduced computational cost.

As a second test, we examine the predictions of BLYP + GAP for the energies of different isomers of the water hexamer. This system has been studied extensively for many years,[7,16,59] for a very important reason. The most stable structures of small water clusters from the trimer to the pentamer have a ringlike form in which each monomer is hydrogen bonded to two neighbors.[59] However, from the hexamer onwards, rings are less stable than compact structures in which some monomers are hydrogen bonded to three or four neighbors.[7,59,60] The energy balance for the hexamer is rather delicate, but high-precision CCSD(T) calculations leave no doubt that the compact prism and cage structures have lower energy than the more open book and ring forms.[7] However, many of the commonly used DFT approximations, including BLYP and PBE, wrongly predict that the ring or book form is most stable.[58] We compare in Fig. 3 the predictions of BLYP + GAP with CCSD(T) benchmarks and with the predictions of BLYP and DMC. We see that again the GAP-corrected DFT model is highly accurate and is comparable to DMC.

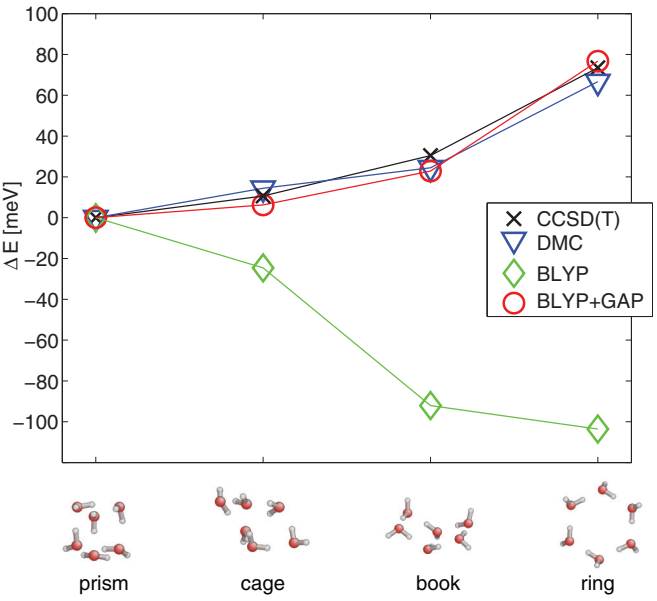

FIG. 3. (Color online) Relative energies (meV units) of four isomers of the water hexamer[58] computed with CCSD(T) close to the basis-set limit,[7] diffusion Monte Carlo,[58] BLYP, and BLYP + GAP. Geometries are depicted below the figure.

B. Ice polymorphs

For any material, crystal energetics provides a crucial test of modeling techniques. Water has a remarkably rich phase diagram, with no fewer than fifteen known ice structures.[61,62] In the common form ice Ih, found at ambient pressure, each H_2O monomer is H-bonded to four nearest neighbors at an O-O distance of 2.75 Å, the next-nearest neighbors having the much greater O-O separation of 4.5 Å. The pattern of H-bonding in ice Ih is disordered, but the closely related ordered form ice XI, stable below 72 K, has essentially the same local geometry.[61] With increasing pressure, denser structures become more stable, and we will be concerned here with (in order of increasing density) ice IX, II, XV, and VIII. The distances to non-H-bonded next-nearest neighbors decrease along this series, becoming almost exactly equal to the first-neighbor distance in ice VIII.[61] There are accurate experimental values for the zero-pressure energies and volumes of almost all these structures.

Standard DFT approximations perform poorly for ice,[22] making the energies increase far too much from ice Ih to VIII, and giving transition pressures too high by up to a factor of 10. Our machine-learning techniques allow us to correct any DFT approximation for one- and two-body errors, and enable us to discover whether these errors are responsible for the poor description of ice energetics. We have used BLYP + GAP to calculate the relaxed geometries and the equilibrium energies and volumes of the ice structures mentioned above, substituting the periodic Bernal and Fowler structure for the proton-disordered Ih.[65] The results are reported in Table II, where we also give results obtained with uncorrected BLYP.

The BLYP + GAP energies and volumes of the various structures relative to those of ice Ih are much better than those given by uncorrected BLYP. Given the large decrease of second-neighbor O-O distance from 4.5 Å to ∼2.7 Å as we pass from Ih to VIII, the experimental energy difference of 31 meV/monomer between the two structures is remarkably small. The energy difference of 223 meV given by BLYP is grossly in error, but the GAP correction brings the difference into very close agreement with experiment. Similarly, the volume differences between Ih and the other structures from BLYP + GAP are much better than from BLYP itself. Nevertheless, BLYP + GAP still suffers from some errors, since it gives a significant *uniform* overbinding in all the structures due to beyond-two-body errors, implying that BLYP + GAP overestimates the strength of cooperative H bonding, leading to an overestimate of the equilibrium density by about 5–10%. The systematic overestimation of density by BLYP + GAP would be partially compensated by zero-point effects, which increase volumes by between 1–5%.[66] Complete correctness in the relative stabilities of the low-lying ice structures is

TABLE II. Binding energies and volumes of ice polymorphs computing using DFT with the BLYP functional and BLYP + GAP compared with experimental values.[63] Zero-point vibrational contributions have been removed from the experimental energies,[64] but not from the volumes.

	Ih	II	VIII	IX	XI	XV
Binding energy [meV]						
BLYP	−540	−458	−318	−475	−544	−403
BLYP + GAP	−667	−672	−637	−670	−671	−657
EXPT	−610	−609	−579	−606		
Binding energy relative to Ih [meV]						
BLYP	0	83	223	66	−3	138
BLYP + GAP	0	−5	30	−3	−4	−11
EXPT	0	1	31	4		
Volume [Å^3]						
BLYP	31.7	26.0	21.2	28.0	32.1	24.6
BLYP + GAP	30.6	23.8	18.6	24.0	30.6	21.1
EXPT	32.0	25.3	20.1	25.6	32.0	22.9
Volume relative to Ih [Å^3]						
BLYP	0	−5.7	−10.5	−3.8	0.36	−7.9
BLYP + GAP	0	−6.7	−11.9	−6.5	0	−9.5
EXPT	0	−6.7	−11.9	−6.4	0	−9.2

a stringent test of any method, since some of the energy differences are only a few meV per monomer. Nevertheless, some comments are in order. According to the experimental phase diagram at low temperatures, ice XI should be the most stable structure, with II, XV, and VIII being successively less stable. (Ice Ih is entropically stabilized at higher temperatures, so that it is expected to be energetically slightly less stable than XI.) The clearest error in the relative energies is thus the incorrect prediction of BLYP + GAP that XV is more stable than XI by 7 meV/monomer.

C. Liquid water

Previous DFT simulations of liquid water at near-ambient conditions have encountered three main difficulties. First, the equilibrium density of the liquid sometimes differs unacceptably from the experimental value; uncorrected PBE and BLYP underestimate it by ∼10% and 10%–20% respectively.[10,48] Second, DFT approximations tend to make the liquid overstructured[10,48] as compared with neutron and x-ray diffraction data. Third, the diffusion coefficient is generally reported to be too low, by at least a factor of 2[67] and in some cases by a factor of ten or more.[9,68] In this section, we are asking whether correction of BLYP for 1B and 2B errors substantially improves the description of the liquid under those three headings. We note that some earlier simulations may have been affected by technical sources of error. Lee and Tuckerman pointed out that significantly less over structuring is obtained with very well converged basis sets.[69] On the other hand, Kühne *at al*. showed that there is a significant finite size error when using only 32 water molecules to describe bulk water.[70] See also Wang *et al*. for a discussion of finite size errors in the equilibrium pressure.[10] We have made efforts to take account of these technical problems, including correct temperature control. We also note that a full treatment of the liquid must account for quantum nuclear effects, which are neglected in the present work, and we comment below on how this neglect may alter our conclusions.

1. Pressure

Our MD simulation of liquid D_2O at 308 K and density 1.109 g cm^{-3} performed with BLYP + GAP gave an average negative pressure in the 25 ps production run of −2.6 kbar. By contrast, our simulation using uncorrected BLYP under exactly the same conditions gave the much larger positive pressure of ∼7 kbar, which is associated with the well-known 10%–20% underestimate of equilibrium density. A simple estimate based on our observed pressure together with the experimental compressibility indicates that the equilibrium density with BLYP + GAP is higher than the correct value by ∼10%, which is consistent with the uniform overbinding observed above for the ice structures.

2. Radial distribution functions

In comparing the structure of our simulated liquid with data from x-ray and neutron diffraction, we had to decide which experimental data to compare with. This is not a trivial question, because x-ray and neutron measurements have different strengths and weaknesses for water structure, and there are significant differences between different results. The experimental data that we compare with here come from a joint refinement of x-ray and neutron data, and the considerations that led us to this choice are explained in Appendix.

The well-known DFT errors of overstructuring in liquid water are most clearly seen in the oxygen-oxygen radial distribution function $g_{OO}(r)$. A comparison of the experimental and computed RDFs (using BLYP + GAP and uncorrected BLYP—see Fig. 4) shows that the GAP correction significantly

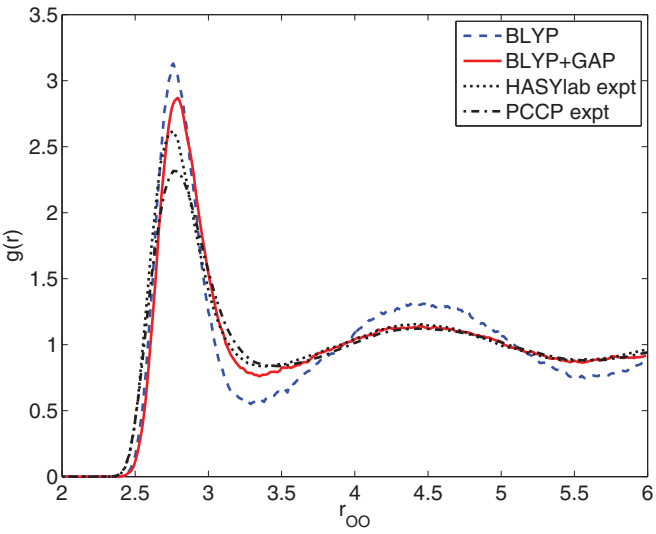

FIG. 4. (Color online) Oxygen-oxygen radial distribution function of liquid water at 308 K at experimental density using BLYP (blue dashed) and BLYP + GAP (red solid) compared with two sets of experimental data (black dotted and dash-dotted). Experimental data are from joint refinement of neutron data and two sets of x-ray data, identified as HASYlab and PCCP.[63]

improves agreement with experiment. The overstructuring of the liquid is partially corrected: the first peak in $g_{OO}(r)$ is lowered by ∼0.25 and the first trough becomes shallower by ∼0.2. However, our comparison with experimental data indicates that the liquid may still be slightly overstructured even with BLYP + GAP.

Our comparison of theory and experiment shows that the height of the first peak of $g_{OO}(r)$ in our system simulated with BLYP + GAP (2.87) is somewhat greater than than the values of 2.3–2.6 that emerge from the joint refinement, and this is why we conclude that our liquid may be a little overstructured. However, two points should be borne in mind; first, some analyses of x-ray data done without joint refinement give a first-peak height of ∼2.83;[71,72] second, quantum nuclear effects appear not to be negligible for $g_{OO}(r)$ in respect of the first peak height, and we estimate from the experimental data[73] that they could lower the height of the first peak by 0.1 or more. However, the most significant region of $g_{OO}(r)$ for characterising order in the liquid may not be the first maximum but the first minimum, because this corresponds to the "interstitial" region where *ab initio* simulations appear to have the greatest difficulty in reproducing experimental RDFs.[10] Here, there is virtual unanimity among the modern x-ray and neutron results that the value of $g_{OO}(r)$ at the first minimum is ∼0.84.[63] Our BLYP + GAP value of 0.77 is close to this. This is in strong contrast to uncorrected BLYP, which gives a very low value of 0.58. In taking all the above factors together, our conclusion is that our BLYP + GAP liquid *may* be slightly overstructured as judged by $g_{OO}(r)$, but the errors appear to be rather small.

We now turn to the comparison with experiment of our calculated $g_{OH}(r)$ and $g_{HH}(r)$. From the comparison with the joint refinement of x-ray and neutron data shown in Fig. 5, we see that the positions of the peaks and troughs in both RDFs agree closely with experiment. The heights of the first peaks and the depths of the first troughs of $g_{OH}(r)$ and $g_{HH}(r)$ in the simulated system are greater than in experiment, but we note that BLYP + GAP is in better agreement with experiment than uncorrected BLYP. However, a full comparison with experiment cannot be made without accounting for quantum nuclear effects, which appear to be significant for both RDFs.[30,76] Quantum nuclear effects can be included in the calculation of RDFs using path-integral simulation, and there have been several investigations of these effects for water using both empirical interaction models[25,30,77,78] and *ab initio* methods.[76] The path-integral simulation of liquid water using BLYP + GAP is a possibility for the future, though the computational effort would be considerable. At present, however, we simply estimate the likely size of the effects using available information.

Published path-integral simulations give conflicting indications about the influence of quantum nuclear effects on the RDFs, but we take as an example the simulations of Fanourgakis and Xantheas,[30] which were based on the empirical TTM3-F interaction model and employed large systems and extensive statistical sampling. These show that on going from classical to quantum nuclei in H_2O (light water), the heights of the first peaks in both $g_{OH}(r)$ and $g_{HH}(r)$ decrease by ∼0.1, with smaller effects in the first troughs and the second peaks. Figure 5 makes it clear that such quantum

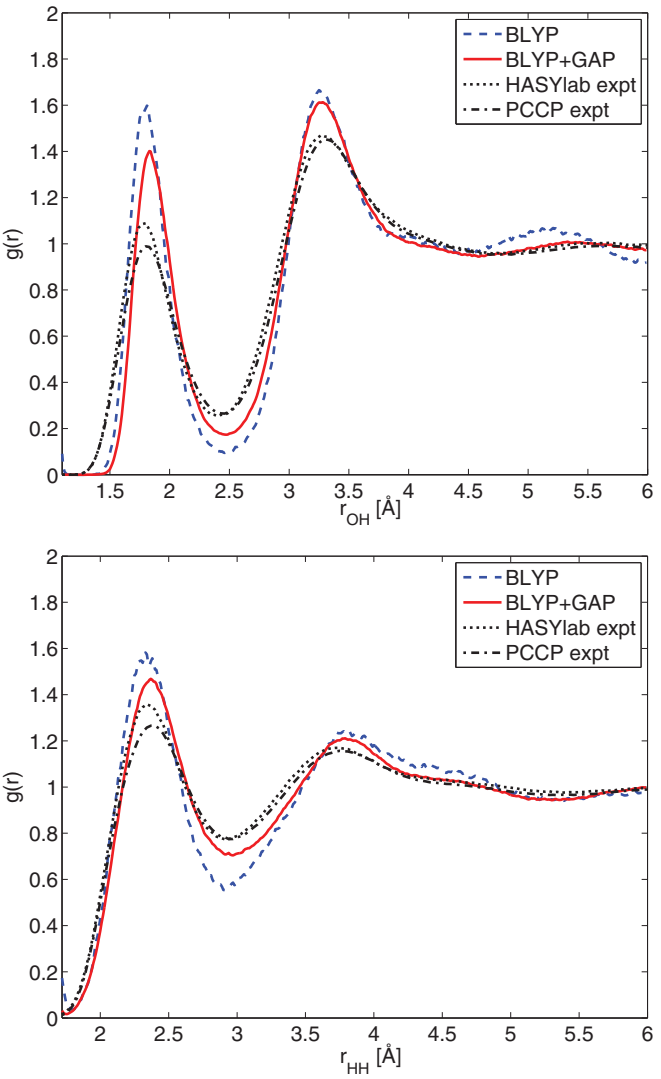

FIG. 5. (Color online) Comparison of calculated radial distribution functions $g_{OH}(r)$ and $g_{HH}(r)$ of bulk liquid water with experimental data. Calculated results are from MD simulations performed with uncorrected and GAP-corrected DFT(BLYP), both simulations performed at 308 K and experimental zero-pressure density. Experimental results are both from joint refinement[63] of x-ray and neutron diffraction measurements, the x-ray measurements being identified as HASYlab[74] and PCCP.[72] The neutron data used in both joint refinements are those of Soper.[75]

corrections would improve still further the agreement between BLYP + GAP and experiment, but noticeable discrepancies would still remain. This would confirm our earlier conclusion that our simulated liquid is slightly over-structured. However, one should note that the *ab initio* path-integral simulations of Morrone and Car[76] based on the BLYP functional suggest considerably larger quantum nuclear corrections, and we believe that definitive statements cannot yet be made.

3. *Diffusivity*

We compute the self-diffusion coefficient D of molecules in our simulation in the conventional way[49] from the slope of the time-dependent-mean-square displacement $\langle \Delta r(t)^2 \rangle$, where the squared displacement $\Delta r(t)^2 \equiv |\mathbf{r}(t_0 + t) - \mathbf{r}(t_0)|^2$

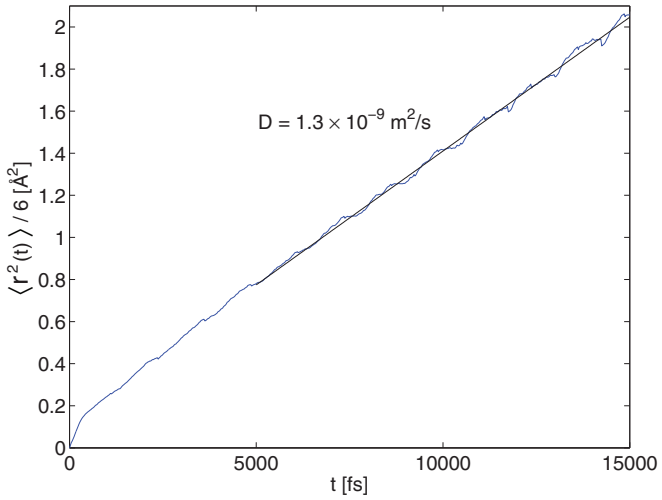

FIG. 6. (Color online) Time-dependent-mean-squared displacement of atoms as function of time interval t from MD simulation of liquid heavy water (D_2O) at 308 K and experimental zero-pressure density. Simulation was performed using GAP-corrected DFT(BLYP) approximation. Straight line is least-squares fit to the calculated data.

of an atom in time t is averaged over all atoms in the system and over all time origins t_0. The averaging is performed only for times t_0 and $t_0 + t$ that are both in the "production" period of the run, and we take time origins t_0 at intervals of 0.1 ps. The diffusion coefficient D is then obtained from the asymptotic form of $\langle \Delta r(t)^2 \rangle$ for large t:

$$\langle \Delta r(t)^2 \rangle \rightarrow A + 6D|t| \, . \tag{6}$$

Our results for $\langle \Delta r(t)^2 \rangle$ are shown in Fig. 6, from which we see that the asymptotic "long-time" form becomes established after a rather short time of only ∼1 ps.

We find a value of $D = 1.3 \times 10^{-9}$ m^2 s^{-1}. In a system of only 64 molecules, the value of D is expected to be reduced by size effects by an amount that can be estimated by standard methods.[79] It is well known that the diffusion coefficient of a liquid calculated in periodic boundary conditions converges only slowly to its value in the thermodynamic limit with increasing system size.[79,80] With the usual cubic simulation cell (length of edge L), the value $D_{\rm PBC}(L)$ obtained from the time-dependent-mean-square displacement is approximately related to the value D_∞ in the thermodynamic limit by the formula

$$D_{\rm PBC}(L) = D_\infty - \frac{k_{\rm B} T \zeta}{6\pi\eta L} \, , \tag{7}$$

where η is the shear viscosity and ζ is a numerical coefficient having the approximate value 2.837. Kühne *et al.*[70] showed that this formula gives a good fit to the size dependence of D in their DFT MD simulations of liquid water. In applying this formula to the correction of our calculated D, we take the experimental value $\eta = 1.3$ mPa s for heavy water at ambient conditions, which yields a correction of 0.4×10^{-9} m^2 s^{-1}, so that the corrected value is $D = 1.7 \times 10^{-9}$ m^2 s^{-1}. This is somewhat below the experimental value $D = 2.4 \times 10^{-9}$ m^2 s^{-1} for heavy water.[81] Since we should, in principle, correct our raw value of D using the η value of our finite-sized simulated system, rather than the experimental η, and since an underestimate of D is likely to be associated with an overestimate of η, one might argue that our correction to D is slightly larger than it should be, but we have not attempted to make an improved estimate. The values reported for BLYP itself are generally too low by a factor of at least 2 and often by more than this.[10,67,82] Once again, one- and two-body effects appear to be the main culprit in making BLYP unrealistic.

V. SUMMARY AND CONCLUSIONS

The work we have presented is motivated by the problem that the standard DFT methods widely used to model condensed matter do not have the accuracy required for a satisfactory treatment of water and some other molecular materials. Our aim has been to address this problem for water by developing machine-learning techniques based on the GAP approach to represent very accurately the one- and two-body errors of chosen DFT approximations, so that these errors can then be almost completely eliminated. To demonstrate how these ideas work in practice, we chose to correct the BLYP functional, which has been a popular choice for work on water. We found that one- and two-body corrected BLYP (referred to here as BLYP + GAP) does indeed give a considerably better description of the energetics of small water clusters and the relative energies and volumes of ice structures, and also gives significant improvements in the structure and dynamics of liquid water. Nevertheless, our work also shows that correction of one- and two-body errors does not yet achieve completely satisfactory energetics, since the resulting approximation produces significant over-binding and noticeably underestimated equilibrium volumes for ice structures and liquid water. It is clear that further developments are needed.

We note that the significant beyond-two-body (B2B) errors of BLYP + GAP are completely consistent with other very recent work. Medders *et al.*[36] have characterized the three-body errors of a number of different energy functions for water, including both parameterized force-fields and DFT approximations. For a large sample of $(H_2O)_3$ configurations, they showed that the three-body energy given by the BLYP functional systematically overbinds. In separate work,[83] the errors of BLYP have been characterized for large samples of the $(H_2O)_6$, $(H_2O)_9$, and $(H_2O)_{15}$ clusters and for periodically repeated configurations of liquid water, by comparing BLYP values of the energies with benchmarks from quantum Monte Carlo calculations. As expected, energies from uncorrected BLYP were always seriously underbound, but after correction for one- and two-body errors significant overbinding was found for all these systems, and the amount of over-binding was similar to what we have found in the present work. This confirms that the GAP methods reported here need to be supplemented by methods for correcting B2B errors. The physical origins of B2B errors of the BLYP functional need further investigation, but we note that a recent analysis for the H_2O hexamer[17] indicates that more than one mechanism may contribute, including an inaccurate description of three-body exchange repulsion.

We are currently working to extend our GAP methods to the correction of B2B errors. For water systems, a closely

related problem has already been addressed by Medders *et al.* in the work referred to above. They showed that a fairly simple parameterized functional form can be used to represent rather accurately the short-range three-body errors of the TTM4-F force field, and they further showed that this successfully corrects for most of the B2B errors of that model for water clusters and the liquid.[27] We have very recently shown[83] that essentially the same procedure succeeds in correcting the B2B errors of BLYP for water clusters up to the 15-mer. This being so, it seems highly likely that a GAP representation of B2B errors can be developed along similar lines, and we are pursuing this idea. In order to use this approach confidently for the liquid, it will be important to ensure that it cures our remaining B2B problems for the ice structures. An alternative approach to B2B errors also merits study. GAP and other related methods have already been highly successful in representing many-body interaction energies.[12,42] In those cases, the representation of the energy differed somewhat from the methods used here, since it employed a decomposition of the total energy into atomic components, machine learning then being used to describe the dependence of these components on the configurations of neighbors within a specified radius. We believe that this approach could be used to represent the errors of DFT B2B energy in water and other molecular systems, and we plan to investigate this possibility further. We also point out that for some molecular systems many-body dispersion can be a substantial effect,[19,84] which will need to be integrated into any GAP scheme for B2B errors.

We have concentrated here on water systems, but the machine-learning methods we have described are much more general, and should be useful for a wide range of molecular materials. The general idea of correcting DFT approximations by adding parameterized energy functions has become very popular in the past few years, and a variety of approaches have been proposed. The GAP methods outlined here have the advantage that they are systematically improvable, at least at the one- and two-body levels. In addition, they can readily be automated. We have stressed the need to extend the methods beyond the two-body level, but it is already interesting to explore the usefulness of one- and two-body GAP corrections for materials other than pure water. As a step in this direction, we are currently developing these corrections for methane-water mixtures, with the aim of testing their effectiveness for the industrially and environmentally important methane clathrate materials.

ACKNOWLEDGMENTS

A.B.P. was supported by a Junior Research Fellowship at Magdalene College, Cambridge. G.C. acknowledges support from the Office of Naval Research under Grant No. N000141010826 and from the European Union FP7-NMP programme under Grant No. 229205 "ADGLASS". The authors are grateful to D. O'Neill for his important contribution to an earlier incarnation of this project, and F.R.M. gratefully acknowledges funding from the Engineering and Physical Sciences Research Council (EP/F000219/1). The authors thank A. K. Soper and C. J. Benmore for useful discussions.

APPENDIX: RADIAL DISTRIBUTION FUNCTIONS FROM EXPERIMENT

The comparison of simulation and experiment for the radial distribution functions of liquid water requires us to make a choice between the many available experimental measurements. We summarize here the considerations that led us to make the comparisons presented in the paper.

Data for $g_{\mathrm{OO}}(r)$ are available from both x-ray scattering[71–74,85–93] and neutron scattering.[63,73,75,94–96] Each of these techniques has important strengths and weaknesses for water, and they can be seen as complementary.[63,73] The merit of x-ray diffraction for $g_{\mathrm{OO}}(r)$ is that the electron density is largely concentrated on the O atom, so that $g_{\mathrm{OO}}(r)$ is strongly weighted in the scattering intensity. However, the scattering cannot be interpreted in terms of an isotropic electron distribution centered on O atoms, and there has been extensive discussion about what electron density should be assumed.[63,72,86] It is recognized that the electron distribution of the H_2O molecule in the liquid is not the same as that in the gas phase, because of the well-known large change of dipole moment on going from the gas phase to the condensed phases.[20,21] One way of overcoming this problem is to adopt a parameterized model for the charge distribution.[86] Alternatively, it has been proposed that the electron distribution can be taken from *ab initio* simulations.[72] It is reassuring that these two approaches give similar results.[72]

The merit of neutron diffraction is that scattering occurs from the nuclei, whose scattering properties are characterized by the accurately known scattering lengths. The technique of isotope substitution allows, in principle, a clean separation of the three RDFs $g_{\mathrm{OO}}(r)$, $g_{\mathrm{OH}}(r)$, and $g_{\mathrm{HH}}(r)$. However, the scattering lengths are such that $g_{\mathrm{OO}}(r)$ is less strongly weighted than $g_{\mathrm{OH}}(r)$ and $g_{\mathrm{HH}}(r)$, so that it is more affected by unavoidable errors.[63] Furthermore, the usual methods of isotopic substitution assume that the structure of the liquid is the same for heavy and light water and their mixtures, and this is not in fact the case.[30,73,97]

With both x-ray and neutron diffraction, the inversion of wave-vector-dependent scattering data to obtain real-space RDFs is far from trivial, and the uncertainties due to wave-vector truncation errors are well known.[75] Recent advances in high-energy x-ray diffraction techniques help to reduce these uncertainties.

It is because of these difficulties of experimental measurement and analysis that we have opted to compare our simulated $g_{\mathrm{OO}}(r)$ with experimental results obtained from joint refinement of both x-ray and neutron data, since one would expect this approach to benefit from the merits of both techniques.[63,98,99] We use the joint refinement results of Soper,[63] based on the method of empirical potential structure refinement (EPSR).[100] This way of making the comparison is also instructive, since the joint refinement has been performed with two independent sets of high quality x-ray data,[72,74] and the differences between them give a useful indication of the uncertainties associated with the measurements themselves, rather than the analysis techniques.

[1] S. Yip, *Handbook of Materials Modeling: Models* (Springer Science + Business Media, Dordrecht, 2005).
[2] D. J. C. Mackay, *Information Theory, Inference, and Learning Algorithms* (Cambridge University Press, Cambridge, 2005).
[3] A. J. Stone, *The Theory of Intermolecular Forces* (Oxford University Press, Oxford, 2013).
[4] J. Klimeš and A. Michaelides, J. Chem. Phys. **137**, 120901 (2012).
[5] T. Helgaker, J. Olsen, and P. Jorgensen, *Molecular Electronic-Structure Theory* (Wiley, New York, 2000).
[6] G. S. Tschumper, M. Leininger, B. Hoffman, E. Valeev, H. F. Schaefer, and M. Quack, J. Chem. Phys. **116**, 690 (2002).
[7] D. M. Bates and G. S. Tschumper, J. Phys. Chem. A **113**, 3555 (2009).
[8] K. Laasonen, M. Sprik, M. Parrinello, and R. Car, J. Chem. Phys. **99**, 9080 (1993).
[9] J. C. Grossman, E. Schwegler, E. W. Draeger, F. Gygi, and G. Galli, J. Chem. Phys. **120**, 300 (2004).
[10] J. Wang, G. Román-Pérez, J. M. Soler, E. Artacho, and M. V. Fernández-Serra, J. Chem. Phys. **134**, 024516 (2011).
[11] A. P. Bartók, Ph.D. thesis, University of Cambridge, 2010.
[12] A. P. Bartók, M. C. Payne, R. Kondor, and G. Csányi, Phys. Rev. Lett. **104**, 136403 (2010).
[13] A. P. Bartók, R. Kondor, and G. Csányi, Phys. Rev. B **87**, 184115 (2013).
[14] M. Rupp, A. Tkatchenko, K.-R. Müller, and O. A. von Lilienfeld, Phys. Rev. Lett. **108**, 058301 (2012).
[15] S. S. Xantheas, J. Chem. Phys. **100**, 7523 (1994).
[16] J. M. Pedulla, F. Vila, and K. D. Jordan, J. Chem. Phys. **105**, 11091 (1996).
[17] F.-F. Wang, G. R. Jenness, W. A. Al-Saidi, and K. D. Jordan, J. Chem. Phys. **132**, 134303 (2010).
[18] N. Goldman and R. J. Saykally, J. Chem. Phys. **120**, 4777 (2004).
[19] O. A. von Lilienfeld and A. Tkatchenko, J. Chem. Phys. **132**, 234109 (2010).
[20] C. A. Coulson and D. Eisenberg, Proc. R. Soc. Lond. A **291**, 445 (1966).
[21] P. L. Silvestrelli and M. Parrinello, Phys. Rev. Lett. **82**, 3308 (1999).
[22] B. Santra, J. Klimeš, D. Alfè, A. Tkatchenko, B. Slater, A. Michaelides, R. Car, and M. Scheffler, Phys. Rev. Lett. **107**, 185701 (2011).
[23] J. R. Hammond, N. Govind, K. Kowalski, J. Autschbach, and S. S. Xantheas, J. Chem. Phys. **131**, 214103 (2009).
[24] G. A. Jeffrey, *An Introduction to Hydrogen Bonding* (Oxford University Press, Oxford, 1997).
[25] S. Habershon, T. E. Markland, and D. E. Manolopoulos, J. Chem. Phys. **131**, 024501 (2009).
[26] R. Bukowski, K. Szalewicz, G. C. Groenenboom, and A. van der Avoird, Science **315**, 1249 (2007).
[27] V. Babin, G. R. Medders, and F. Paesani, J. Phys. Chem. Lett. **3**, 3765 (2012).
[28] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, J. Chem. Phys. **79**, 926 (1983).
[29] C. J. Burnham and S. S. Xantheas, J. Chem. Phys. **116**, 5115 (2002).
[30] G. S. Fanourgakis and S. S. Xantheas, J. Chem. Phys. **128**, 074506 (2008).
[31] C. J. Burnham, D. J. Anick, P. K. Mankoo, and G. F. Reiter, J. Chem. Phys. **128**, 154519 (2008).
[32] R. Kumar, F.-F. Wang, G. R. Jenness, and K. D. Jordan, J. Chem. Phys. **132**, 014309 (2010).
[33] Y. Wang, X. Huang, B. C. Shepler, B. J. Braams, and J. M. Bowman, J. Chem. Phys. **134**, 094509 (2011).
[34] Y. Wang, V. Babin, J. M. Bowman, and F. Paesani, J. Am. Chem. Soc. **134**, 11116 (2012).
[35] E. M. Mas, R. Bukowski, and K. Szalewicz, J. Chem. Phys. **118**, 4386 (2003).
[36] G. R. Medders, V. Babin, and F. Paesani, J. Chem. Theory. Comput. **9**, 1103 (2013).
[37] H. Partridge and D. Schwenke, J. Chem. Phys. **106**, 4618 (1997).
[38] A. Shank, Y. Wang, A. Kaledin, B. J. Braams, and J. M. Bowman, J. Chem. Phys. **130**, 144314 (2009).
[39] O. Bludský, M. Rubeš, P. Soldán, and P. Nachtigall, J. Chem. Phys. **128**, 114102 (2008).
[40] M. Rubeš and O. Bludský, Phys. Chem. Chem. Phys. **10**, 2611 (2008).
[41] T. Morawietz, V. Sharma, and J. Behler, J. Chem. Phys. **136**, 064103 (2012).
[42] T. Morawietz and J. Behler, J. Phys. Chem. A (2013), doi: 10.1021/jp401225b.
[43] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning* (MIT Press, Boston, 2006).
[44] P. Ren and J. Ponder, J. Phys. Chem. B **107**, 5933 (2003).
[45] P. Ren and J. W. Ponder, J. Phys. Chem. B **108**, 13427 (2004).
[46] M. J. Gillan, F. R. Manby, M. D. Towler, and D. Alfè, J. Chem. Phys. **136**, 244105 (2012).
[47] H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, and M. Schuetz, Wiley Interdiscip. Rev.: Comput. Mol. Sci. **2**, 242 (2012).
[48] Z. Ma, Y. Zhang, and M. E. Tuckerman, J. Chem. Phys. **137**, 044506 (2012).
[49] D. Frenkel and B. Smit, *Understanding Molecular Simulation, From Algorithms to Applications* (Academic Press, San Diego, 2001).
[50] H.-J. Werner and F. R. Manby, J. Chem. Phys. **124**, 054114 (2006).
[51] T. B. Adler, G. Knizia, and H.-J. Werner, J. Chem. Phys. **127**, 221106 (2007).
[52] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, J. Chem. Phys. **132**, 154104 (2010).
[53] S. J. Clark, M. D. Segall, C. J. Pickard, P. J. Hasnip, M. I. J. Probert, K. Refson, and M. C. Payne, Z. Kristallogr. **220**, 567 (2005).
[54] G. Kresse and J. Furthmüller, Comput. Mater. Sci. **6**, 15 (1996).
[55] P. Blöchl, Phys. Rev. B Condens. Matter **50**, 17953 (1994).
[56] B. J. Smith, D. J. Swanton, J. A. Pople, H. F. Schaefer, and L. Radom, J. Chem. Phys. **92**, 1240 (1990).
[57] J. A. Anderson and G. S. Tschumper, J. Phys. Chem. A **110**, 7268 (2006).
[58] B. Santra, A. Michaelides, M. Fuchs, A. Tkatchenko, C. Filippi, and M. Scheffler, J. Chem. Phys. **129**, 194111 (2008).
[59] J. K. Gregory, D. C. Clary, K. Liu, M. G. Brown, and R. J. Saykally, Science **275**, 814 (1997).
[60] J. Kim, D. Majumdar, H. M. Lee, and K. S. Kim, J. Chem. Phys. **110**, 9128 (1999).
[61] V. F. Petrenko and R. W. Whitworth, *Physics of Ice* (Oxford University Press, Oxford, 1999).
[62] C. G. C. Salzmann, P. G. P. Radaelli, E. E. Mayer, and J. L. J. Finney, Phys. Rev. Lett. **103**, 105701 (2009).
[63] A. K. Soper, J. Phys.: Condens. Matter **19**, 335206 (2007).
[64] E. Whalley, J. Chem. Phys. **81**, 4087 (1984).
[65] J. D. Bernal and R. H. Fowler, J. Chem. Phys. **1**, 515 (1933).
[66] É. D. Murray and G. Galli, Phys. Rev. Lett. **108**, 105502 (2012).

[67]H.-S. Lee and M. E. Tuckerman, J. Chem. Phys. **126**, 164501 (2007).
[68]P. Sit and N. Marzari, J. Chem. Phys. **122**, 204510 (2005).
[69]H.-S. Lee and M. E. Tuckerman, J. Chem. Phys. **125**, 154507 (2006).
[70]T. D. Kühne, M. Krack, and M. Parrinello, J. Chem. Theory. Comput. **5**, 235 (2009).
[71]G. L. Hura, J. M. Sorenson, R. M. Glaeser, and T. Head-Gordon, J. Chem. Phys. **113**, 9140 (2000).
[72]G. L. Hura, D. Russo, R. M. Glaeser, T. Head-Gordon, M. Krack, and M. Parrinello, Phys. Chem. Chem. Phys. **5**, 1981 (2003).
[73]A. K. Soper and C. J. Benmore, Phys. Rev. Lett. **101**, 065502 (2008).
[74]R. T. Hart, C. J. Benmore, J. C. Neuefeind, S. Kohara, B. Tomberli, and P. A. Egelstaff, Phys. Rev. Lett. **94**, 047801 (2005).
[75]A. K. Soper, Chem. Phys. **258**, 121 (2000).
[76]J. A. Morrone and R. Car, Phys. Rev. Lett. **101**, 017801 (2008).
[77]R. A. Kuharski and P. J. Rossky, J. Chem. Phys. **82**, 5164 (1985).
[78]H. A. Stern and B. J. Berne, J. Chem. Phys. **115**, 7622 (2001).
[79]B. Dünweg and K. Kremer, J. Chem. Phys. **99**, 6983 (1993).
[80]I. C. Yeh and G. Hummer, J. Phys. Chem. B **108**, 15873 (2004).
[81]E. H. Hardy, A. Zygar, M. D. Zeidler, M. Holz, and F. D. Sacher, J. Chem. Phys. **114**, 3174 (2001).
[82]R. Jonchiere, A. P. Seitsonen, G. Ferlat, A. M. Saitta, and R. Vuilleumier, J. Chem. Phys. **135**, 154503 (2011).
[83]D. Alfè, A. P. Bartók, G. Csányi, and M. J. Gillan, J. Chem. Phys. **138**, 221102 (2013).
[84]R. A. DiStasio, O. A. von Lilienfeld, and A. Tkatchenko, Proc. Natl. Acad. Sci. U.S.A. **109**, 14791 (2012).
[85]A. H. Narten and H. A. Levy, J. Chem. Phys. **55**, 2263 (1971).
[86]J. M. Sorenson, G. L. Hura, R. M. Glaeser, and T. Head-Gordon, J. Chem. Phys. **113**, 9149 (2000).
[87]L. Fu, A. Bienenstock, and S. Brennan, J. Chem. Phys. **131**, 234702 (2009).
[88]G. N. I. Clark, G. L. Hura, J. Teixeira, A. K. Soper, and T. Head-Gordon, Proc. Natl. Acad. Sci. U.S.A. **107**, 14003 (2010).
[89]C. Huang, K. T. Wikfeldt, D. Nordlund, U. Bergmann, T. McQueen, J. Sellberg, L. G. M. Pettersson, and A. A. Nilsson, Phys. Chem. Chem. Phys. **13**, 19997 (2011).
[90]V. Petkov, Y. Ren, and M. Suchomel, J. Phys.: Condens. Matter **24**, 155102 (2012).
[91]J. C. Neuefeind, C. J. Benmore, J. K. R. Weber, and D. Paschek, Mol. Phys. **109**, 279 (2011).
[92]L. B. Skinner, C. J. Benmore, and J. B. Parise, J. Phys.: Condens. Matter **24**, 338001 (2012).
[93]L. B. Skinner, C. J. Benmore, B. Shyam, J. K. R. Weber, and J. B. Parise, Proc. Natl. Acad. Sci. U.S.A. **109**, 16463 (2012).
[94]A. K. Soper and R. N. Silver, Phys. Rev. Lett. **49**, 471 (1982).
[95]A. K. Soper, F. Bruni, and M. A. Ricci, J. Chem. Phys. **106**, 247 (1997).
[96]J. C. Dore, M. Garawi, and M. C. Bellissent-Funel, Mol. Phys. **102**, 2015 (2004).
[97]A. Zeidler, P. S. Salmon, H. E. Fischer, J. C. Neuefeind, J. M. Simonson, H. Lemmel, H. Rauch, and T. E. Markland, Phys. Rev. Lett. **107**, 145501 (2011).
[98]M. M. Leetmaa, K. T. Wikfeldt, M. P. Ljungberg, M. M. Odelius, J. J. Swenson, A. A. Nilsson, and L. G. M. Pettersson, J. Chem. Phys. **129**, 084502 (2008).
[99]K. T. Wikfeldt, M. M. Leetmaa, M. P. Ljungberg, A. A. Nilsson, and L. G. M. Pettersson, J. Phys. Chem. B **113**, 6246 (2009).
[100]A. K. Soper, Phys. Rev. B **72**, 104204 (2005).