Polarizable interaction potential for water from coupled cluster calculations. I. Analysis of dimer potential energy surface

Robert Bukowski,[1] Krzysztof Szalewicz,[1,a)] Gerrit C. Groenenboom,[2] and Ad van der Avoird[2,b)]

[1]*Department of Physics and Astronomy, University of Delaware, Newark, Delaware 19716, USA*
[2]*Theoretical Chemistry, Institute for Molecules and Materials, Radboud University Nijmegen, Toernooiveld 1, 6525 ED Nijmegen, TheNetherlands*

A six-dimensional interaction potential for the water dimer has been fitted to *ab initio* interaction energies computed at 2510 dimer configurations. These energies were obtained by combining the supermolecular second-order energies extrapolated to the complete basis set limit from up to quadruple-zeta quality basis sets with the contribution from the coupled-cluster method including single, double, and noniterative triple excitations computed in a triple-zeta quality basis set. All basis sets were augmented by diffuse functions and supplemented by midbond functions. The energies have been fitted using an analytic form with the induction component represented by a polarizable term, making the potential directly transferable to clusters and the bulk phase. Geometries and energies of stationary points on the potential surface agree well with the results of high-level *ab initio* geometry optimizations. © *2008 American Institute of Physics.*
[DOI: 10.1063/1.2832746]

I. INTRODUCTION

Due to its abundance and overwhelming importance for living organisms, water has been the subject of intense studies by generations of scientists, both on the experimental and theoretical sides. It might be therefore somewhat surprising that water clusters and condensed phases are still far from being fully understood. The reason is that van der Waals interactions between water molecules are capable of assembling these molecules into objects with very peculiar properties, exemplified by the unusual temperature profile of the liquid density with a maximum at 4° C, the anomalous isothermal compressibility, and the multiple phases of ice. Invaluable tools for investigations of water on the molecular level are the Monte Carlo (MC) and molecular dynamics (MD) simulation techniques,[1] which allow one to derive the macroscopic properties from intermolecular interaction potentials. The nature of the interactions between water molecules turns out to be rather complicated. First, the existence of the deep minimum in the hydrogen-bonded configuration of the dimer combined with the geometry of the water monomer leads to strong orientational preferences in forming clusters and the liquid phase, resulting in dynamical "hydrogen-bonded networks." A further complication is brought by nonadditive effects which make it impossible to sufficiently accurately represent the energy of water as a sum of pair contributions. Moreover, the usually neglected dependence of the interaction potential on monomer internal coordinates (the so-called monomer-flexibility effects) may play an important role in the determination of seemingly static properties, such as the liquid structure. An additional difficulty ap-

pearing in simulations of water is the quantum nature of the motion of water molecules. Due to the small moments of inertia, quantum effects show up not only in the "stiff" intramonomer degrees of freedom but also in the "soft" ones, describing the relative motions of the molecules. All these issues are often implicitly taken care of by applying so-called empirical potentials—usually pairwise-additive functions fitted to reproduce bulk-phase experimental data in a classical simulation. Popular empirical models of this kind include extended simple point charge (SPC/E),[2] four-point transferable intermolecular potential (TIP4P),[3] five-point transferable intermolecular potential (TIP5P),[4] and Reimers–Watts–Klein (RWK) (Ref. [5]) potentials. Since the early 1990s, "polarizable" empirical potentials have also become popular.[6–10] Such potentials contain the classical induction terms which explicitly account for many-body induction effects. Both types of empirical potentials describe the properties of bulk water remarkably well, particularly taking into account their very simple functional form, but always have a limited range of applicability, determined by the range of thermodynamic parameters used in the fit. Moreover, simulations with empirical potentials do little to answer questions about the physical origins of the phenomena being investigated. This is because the effective nature of such potentials obscures different components of the interaction and often distorts one component to compensate for another. For example, empirical nonpolarizable potentials fitted to thermophysical data are not able to reasonably reproduce pure dimer properties, such as the second virial coefficient, since these potentials implicitly account for nonadditive interactions not present in a dimer. Instead of classical MC/MD simulations, one can perform path integral MC (PIMC),[11–13] path integral MD (PIMD),[14–16] or the so-called centroid MD simulations[17–25] which account for quantum effects. Most of

a)Electronic mail: szalewic@udel.edu.
b)Electronic mail: avda@theochem.ru.nl.

128, 094313-1

quantum MC/MD simulations to date used empirical potentials which already implicitly include quantum effects. However, recently some simulations with nonempirical potentials have been performed.[16,26,27]

An alternative to empirical potentials is to calculate the interaction energies *ab initio*. The conceptually simplest approach is the "on the fly" molecular dynamics where calculations are performed at the whole multimolecule system for each point on the trajectory visited by a classical MD simulation. The most popular implementation of such a method was developed by Car and Parrinello[28] (CPMD), but many other implementations are used.[29,30] Quantum extensions of the original CPMD method have been proposed,[31] based on path integral ideas. An on the fly method straightforwardly accounts for both the nonadditive and monomer-flexibility effects, requires calculation for a relatively small number of points compared to the dimensionality of a typical simulated system, and does not require any fitting of potential surfaces. On the other hand, each calculation has to be performed on about $N=32$ or more molecules, so that only the simplest and not very accurate *ab initio* methodology based on density functional theory (DFT) can be used in practice. Even then, the computational cost of the simulations only allows for small system sizes and short time scales, typically on the order of a few picoseconds. Moreover, it is well known that DFT methods, at least in their present form, are unable to reproduce the long-range dispersion component of the interaction energy, important for weakly bound van der Waals systems.

Another *ab initio* approach is the one based on the many-body expansion, where the interaction energy is given as a sum of the two-body term (itself a sum of pair interactions) and the three-, four-, and higher-body (pairwise-nonadditive) terms. The potentials representing the first few terms, much simpler than the complete *N*-body potential, can be calculated using accurate *ab initio* methods, fitted to analytic functions, and used in a simulation. The dimensionality of the potentials grows quickly with the number of bodies and in practice it is not possible to move beyond three-body potentials with rigid monomers. Thus, the applicability of this approach depends critically on the rate of convergence of the many-body expansion. See Ref. 32 for a more extensive discussion of these issues.

A prerequisite of a successful *ab initio* simulation using the many-body expansion is the availability of the two-body or pair potential, describing the interactions between just two water molecules. Such a potential itself is of considerable interest for the interpretation of the microwave and far-infrared spectra of the dimer,[33–38] measured in supersonic molecular beams. Another interesting purely two-body property is the second virial coefficient of steam, important for the accurate thermodynamic description of various systems and in the construction of humidity standards.[39]

Interactions between two water molecules have been a subject of *ab initio* studies since the early days of quantum chemistry. Most often calculations were carried out using the supermolecular method for characteristic points on the potential surface or local optimizations were performed to locate such points. An important study of this type is the one

by Smith *et al.*,[40] who used second-order many-body perturbation theory (MBPT2) to locate ten stationary points, including the global minimum and transition structures on the interconversion paths between different equivalent minima. For selected configurations found in this way, more accurate calculations were performed at the fourth-order (MBPT4) level. More recently, the analysis of Ref. 40 was extended by Tschumper *et al.*,[41] who refined the ten stationary points using the coupled-cluster method with single, double, and noniterative triple excitations [CCSD(T)] in a triple-zeta quality [TZ2P(f,d)+dif] basis set. For all the optimized stationary points, very accurate interaction energies were obtained including the MBPT2 component extrapolated to the complete basis set (CBS) limit, the CCSD(T)/[TZ2P(f,d)+dif] contribution, as well as higher correlation contributions obtained from the Brueckner doubles (BD) method[42] with perturbative treatment of triple and quadruple excitations [BD(TQ)].[43] Core-valence, core-core, and relativistic contributions were also considered. The energies reported in Ref. 41 are believed to be accurate to about 0.05 kcal/mol. Most attention in the literature, however, has been devoted to the global minimum of the dimer. Very accurate optimizations and interaction energy calculations for this structure have been reported in Refs. 44–46.

The local explorations of the potential energy landscape described above provide important information about the characteristic structures and interconversion paths. However, to perform a simulation of liquid water or ice, a representation of the complete potential energy surface (PES) is necessary. The first *ab initio* pair potential for rigid monomers, created in 1976 by Matsuoka, Clementi, and Yoshimine (MCY),[47] was a simple site-site form including Coulomb and double exponential terms, fitted to 66 dimer energies computed using a variant of the configuration interaction method. The relatively good performance of this potential in predicting bulk observables should be considered fortuitous taking into account the small number of grid points and the rather inaccurate values of the *ab initio* energies used in the fit. Later attempts to improve the accuracy of the potential led, in fact, to less satisfactory predictions.[48,49] Continuation of this work in the early 1990s by Niessar, Corongiu, Clementi *et al.* (NCC) resulted in a yet another pair potential fitted to 350 energies computed at the MBPT4 level.[50–52] The NCC potential contained a polarization term, fitted to three-body nonadditive energies computed at the self-consistent field (SCF) level for 250 trimers.

Stone and co-workers[53,54] developed a set of rigid-monomer anisotropic site potentials (ASP), based on elaborate asymptotic models of electrostatic, induction, and dispersion interactions, and a fit to exchange energies calculated using the Hayes-Stone intermolecular perturbation theory (IMPT),[55,56] a variant of symmetry-adapted perturbation theory (SAPT).[57–59] The dispersion asymptotic coefficients used in these potentials were obtained by Rijks and Wormer[60,61] (ASP-W) and by Szczesniak *et al.*[62] (ASP-S). Modifications of the ASP-W potential by Millot *et al.*[54] led to the ASP-W2 and ASP-W4 potential energy surfaces. Although the ASP potentials are formally two-body potentials, the representation of the induction energy in terms of perma-

nent multipoles and polarizabilities also allows the recovery of asymptotic nonadditive induction effects in clusters or condensed phases. More recently, Burnham and Xantheas[63,64] developed several potentials for water. In particular, they obtained a fairly good agreement with experiment for structural parameters of liquid water and ice using the so-called TTM2-R (Thole-type model rigid) polarizable model,[63] fitted to scaled MBPT2/aug-cc-pVTZ dimer interaction energies computed along the minimum energy path of C_s symmetry. Recently, the TTM2 model was revised,[65] and the new version is denoted as TTM2.1. Whereas this potential continues to work very well for clusters and bulk water, its two-body part does not properly reproduce[66] the dimer vibration-rotation-tunneling (VRT) spectrum. This may indicate that the small numbers of grid points and parameters used to generate TTM2.1 do not allow an accurate representation of the complicated dimer potential energy surface.

As it was first shown in Refs. 67 and 68, the VRT spectrum of the water dimer is very sensitive to the details of the potential surface near the minimum and to the characteristics of the tunneling barriers. Most of the popular *ab initio* and empirical potentials produce the VRT transition frequencies several hundred percent different from the experimental ones.[66–68] Consequently, Fellers *et al.*[69] developed the VRT(ASP-W) potential—obtained by explicitly refitting a number of parameters of the ASP-W surface[53] to the VRT spectroscopic data—which reproduced the spectra very well. Recently, two new fits of this kind have been obtained by Goldman *et al.*,[70] dubbed VRT(ASP-W)II and VRT(ASP-W)III, and used in diffusion Monte Carlo simulations of water clusters up to the hexamer[71] and in simulations of liquid water.[72] Since the starting point for all these fits were *ab initio* potentials, such types of potentials should perhaps be called semiempirical ones.

In Refs. 73 and 74, SAPT (Refs. 57–59) was used to calculate interaction energies for a large number of water dimer geometries, 1008 and 2510, respectively. The potential developed in Ref. 73, called SAPT-pp, separately fitted each main component of the interaction energy (such as electrostatic, induction, dispersion, and exchange) using an expansion in terms of functions of the Euler angles and center of mass separation. Although the fit was accurate, it was too complex to be used in MC/MD studies. Therefore, in Ref. 74 the extended set of energies was fitted to a flexible site-site formula involving five symmetry-distinct sites on each molecule. The resulting six-dimensional potential surface, called SAPT-5s, was then employed in calculations of VRT spectra of the H_2O and D_2O dimers, producing results in almost quantitative agreement with experiment.[75–77] SAPT-5s was the first purely *ab initio* potential for water to give such good predictions of the VRT spectrum. The discrepancies from experiment were further reduced by "tuning," i.e., refitting the *ab initio* data while subjecting the fit parameters to a single constraint enforcing the desired splitting between two of the dimer energy levels. The "tuned" potential was dubbed SAPT-5st.[76]

An approach similar to that used to develop SAPT-pp was recently applied by Torheyden and Jansen[78] who analyzed the physical components of the water dimer interaction energy. These authors computed SAPT interaction energies at nearly 500 grid points with rigid monomers using the aug-cc-pVTZ basis and bond functions. They also computed CCSD(T) interaction energies in aug-cc-pVTZ and aug-cc-pVQZ bases (without bond functions) and extrapolated them to CBS limits. Torheyden and Jansen developed two different sophisticated component-by-component fits to the potential energy surface which accurately reproduced the *ab initio* values. However, the forms of these fitting functions are probably too complicated for use in MC/MD simulations. The virial coefficients computed in Ref. 78 deviated from experiment somewhat more than SAPT-5s virials did.

In recent years, 12-dimensional *ab initio* potentials for the water dimer with flexible monomers have also appeared. A SAPT potential,[79] consistent with SAPT-5s for the rigid monomer geometry, was based on nearly 250 000 grid points. This potential was used to predict for the first time the anharmonic effects in the shifts of monomer vibrational frequencies upon dimerization. Huang *et al.*[80] obtained a potential energy for the water dimer with flexible monomers using the CCSD(T) approach, the aug-cc-pVTZ basis set, and nearly 20 000 grid points.

In another development, a three-body nonadditive potential was created[81] by fitting a physically motivated 12-dimensional site-site formula to nonadditive energies of more than 7500 water trimers, calculated using the three-body version of SAPT (Ref. 82) at the SCF level of theory. This level of nonadditive theory was shown to be sufficiently accurate for water[81] and predicted water trimer spectra very well.[75] This is in contrast to systems such as rare gas trimers which are dominated by post-SCF terms.[82–86] With the two- and three-body potentials of spectroscopic accuracy at hand, it became possible to probe the fundamental question of the convergence of the many-body expansion of the interaction energy by performing liquid phase simulations with the SAPT-5s+3B(SAPT) potential.[87] The higher-than-three-body induction effects were also accounted for by means of an asymptotic polarizable model. Although the overall agreement with experiment was good, the oxygen-oxygen radial distribution function did not possess some features characteristic for the tetrahedral arrangement of water molecules visible in experimental curves.[88,89] Several explanations were suggested to account for the discrepancy between the simulated and experimental structures, including the quantum and monomer-flexibility effects, both neglected in the simulations of Ref. 87. Insufficient representation of the nonadditive effects, especially the higher-body ones, was also suspected. Residual deficiencies of the SAPT-5s pair potential were considered less likely due to the very good performance of this potential in the VRT spectrum calculations. However, the basis set and theory level truncation error of SAPT-5s, estimated in Ref. 74 to be about 0.3 kcal/mol in the minimum region, certainly left room for improvements of the two-body component. It also left open a possibility that the accuracy of SAPT-5s was uneven in different regions of configurational space. Thus, accuracy could be high in the region probed by VRT spectra, but lower in some regions relevant for predictions of liquid structure.

The suspicion that inaccuracies of the SAPT-5s pair po-

tential are responsible for the discrepancies with experiment for liquid water was confirmed by the development of a recent pair potential for water.[90] This potential was created to test the ability of generating the complete potential surfaces by the new version of SAPT based on the DFT description of monomers and denoted by SAPT(DFT)—see Refs. [91–93] and references therein. Somewhat surprisingly, the resulting potential, named SDFT-5s, was found to be overall predicting the experimental properties better than SAPT-5s, although some VRT transitions were recovered significantly worse. In particular, the new potential recovered the structure of liquid water substantially better than SAPT-5s. This indicated that inaccuracies of the two-body part are actually the major reason for the discrepancies with experiments seen in Ref. [87]. To check this possibility, we decided to develop one more *ab initio* pair potential for water. To reach as high accuracy as possible, we applied the CCSD(T) approach, a higher-level method than the currently programed SAPT. Although, for most dimers, SAPT and CCSD(T) energies are very close, the small differences may be relevant for simulations of water. In order to make our potential nearly basis set independent, we obtained it as a sum of the MBPT2 interaction energy extrapolated to the CBS limit and the contribution beyond MBPT2 from CCSD(T) calculations in an augmented triple-zeta quality basis set. The analytic, six-dimensional representation is similar to that used for SAPT-5s, except that it contains an induction term in the form of the classical polarization model, which makes it possible to automatically include asymptotic many-body induction effects when considering clusters or condensed phases. In this respect, the potential described here, which we shall refer to as CC-pol, may be placed in the category of "polarizable potentials." The CC-pol potential has been briefly described in Ref. [66]. In the current paper, we will present the *ab initio* calculations and estimate their accuracy as well as develop the asymptotic part of the potential surface. The main part of the present work will be a test of the CC-pol potential in predictions of characteristic points on the surface. The CC-pol potential has been applied in calculations of VRT spectra of the water dimer, calculations of virial coefficients, and in simulations of liquid water.[66] A more detailed account of these applications is given in the paper accompanying the present one.[94]

The outline of the paper is as follows. In Sec. II we present the details of our *ab initio* calculations, followed in Sec. III by the description of the fitting function and strategy. Section IV discusses the stationary points found on the CC-pol surface and compares them to those from the high-level *ab initio* optimizations of Tschumper *et al.*,[41] as well as to those given by the SAPT-5s (Ref. [74]) and VRT(ASP-W) (Ref. [69]) potentials. Comparison of the CC-pol potential with SAPT-5s is continued in Sec. V. Finally, Sec. VI contains a summary and conclusions.

II. METHOD

A. Interaction energies

The potential reported in this work describes the interaction between two rigid water molecules. As argued in

Ref. [95], if the nuclear dynamics of the dimer is to be approximated by rigid molecule dynamics, the preferred rigid-monomer structure is the one corresponding to the geometry averaged over the vibrational ground state of the monomer. Thus, following Ref. [96], we make the O–H distance equal to 0.971 625 7 Å and the HOH angle equal to 104.66°. The set of 2510 dimer geometries used in our calculations is the same as in Ref. [74]. It contains a broad spectrum of configurations, covering the vicinities of the minima and saddle points, regions giving large contributions to the second virial coefficient, as well as configurations that are characteristic for liquid water and ice.

For each dimer geometry, the interaction energy is calculated in the supermolecular approach using the counterpoise (CP) correction scheme of Boys and Bernardi,[97–100] i.e., as the difference between the energy of the dimer and the sum of the energies of the monomers, all obtained in the full basis set of the dimer. We used the correlation-consistent augmented aug-cc-pVXZ basis sets of Dunning and co-workers,[101,102] with $X=T,Q$. For a number of geometries, the $X=5$ basis was also used to assess the convergence of the results. In each case, a 14-term set of $3s2p1d$ bond functions[73] was added at the point midway between the centers of mass of the monomers. In all our calculations, all electrons were correlated.

Let $E_{\mathrm{int}}^{\mathrm{SCF}}(X)$, $E_{\mathrm{int}}^{\mathrm{MBPT2}}(X)$, and $E_{\mathrm{int}}^{\mathrm{CCSD(T)}}(X)$ denote the supermolecular interaction energies at the SCF, MBPT2, and CCSD(T) levels, respectively, calculated at a given dimer geometry in the aug-cc-pVXZ basis set (with the bond functions specified above). With the MBPT2 and CCSD(T) correlation contributions given by

$$\Delta E_{\mathrm{int}}^{\mathrm{MBPT2}}(X) = E_{\mathrm{int}}^{\mathrm{MBPT2}}(X) - E_{\mathrm{int}}^{\mathrm{SCF}}(X), \qquad (1)$$

$$\Delta E_{\mathrm{int}}^{\mathrm{CCSD(T)}}(X) = E_{\mathrm{int}}^{\mathrm{CCSD(T)}}(X) - E_{\mathrm{int}}^{\mathrm{MBPT2}}(X), \qquad (2)$$

we define the interaction energy at a given dimer geometry as

$$E_{\mathrm{int}}(\mathrm{extr}) = E_{\mathrm{int}}^{\mathrm{SCF}}(Q) + \Delta E_{\mathrm{int}}^{\mathrm{MBPT2}}(TQ) + \Delta E_{\mathrm{int}}^{\mathrm{CCSD(T)}}(T), \qquad (3)$$

where the MBPT2 correlation contribution has been extrapolated to the infinite basis set limit from the $X=T$ and $X=Q$ results using the X^{-3}-type scheme ($X=2,3,4$ is traditionally denoted by D, T, Q, respectively),

$$\Delta E_{\mathrm{int}}^{\mathrm{MBPT2}}(TQ) = \Delta E_{\mathrm{int}}^{\mathrm{MBPT2}}(Q)$$
$$+ \tfrac{27}{37}[\Delta E_{\mathrm{int}}^{\mathrm{MBPT2}}(Q) - \Delta E_{\mathrm{int}}^{\mathrm{MBPT2}}(T)]. \qquad (4)$$

All the quantities on the right-hand side (rhs) of Eqs. (1)–(4) were computed using the MOLPRO package.[103]

To assess the accuracy of the SCF component $E_{\mathrm{int}}^{\mathrm{SCF}}(Q)$ and of the extrapolation [Eq. (4)], we performed additional calculations at the MBPT2 level for a few points in the region of the potential well using the aug-cc-pV5Z basis set. Based on the differences between the $X=Q$ and $X=5$ results, the SCF component may be considered accurate to better than 0.01 kcal/mol. For the MBPT2 component, we have used $X=Q$ and $X=5$ energies in an extrapolation formula analogous to Eq. (4) to obtain $\Delta E_{\mathrm{int}}^{\mathrm{MBPT2}}(Q5)$. The differences
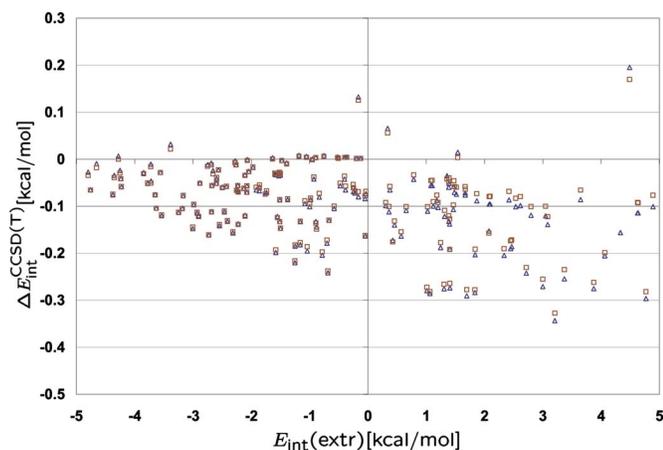
FIG. 1. (Color online) Basis set convergence of $\Delta E_{int}^{CCSD(T)}$ for a set of 189 dimer geometries randomly selected from among 2510. Squares: $\Delta E_{int}^{CCSD(T)}(TQ)$, triangles: $\Delta E_{int}^{CCSD(T)}(T)$.

between the $\Delta E_{int}^{MBPT2}(TQ)$ and $\Delta E_{int}^{MBPT2}(Q5)$ values averaged around 0.03 kcal/mol, which may be considered a reasonable estimate of the error of the MBPT2 correlation component of our potential. This estimate is consistent with the one proposed in Ref. 41, based on similar extrapolations. One may add that the Q5 extrapolations performed with and without bond functions gave CBS limits to within 0.005 kcal/mol of each other. This is an additional indication that these limits are likely converged to about 0.01 kcal/mol, confirming our estimates of accuracy.

The basis set saturation of the contribution $\Delta E_{int}^{CCSD(T)}$ was assessed by calculating this contribution in the aug-cc-pVQZ basis (plus bond functions) for 189 dimer geometries randomly selected from the set of 2510 and extrapolating as in Eq. (4) to obtain the quantity $\Delta E_{int}^{CCSD(T)}(TQ)$. The latter quantity and $\Delta E_{int}^{CCSD(T)}(T)$ are shown in Fig. 1 for the 189 points as functions of the total interaction energy $E_{int}(extr)$ of Eq. (3). The immediate observation is that, for most test configurations, the correlation effects beyond MBPT2 are attractive and non-negligible, often exceeding −0.2 kcal/mol. Closer inspection shows that near the global minimum and other low-energy stationary points (*vide infra*), these effects are smaller, on the order of −0.05 kcal/mol or less, consistent with literature findings.[41,46] However, the larger magnitude of high-level correlation contributions for other dimer configurations indicates that the use of the CCSD(T) method for obtaining the complete potential surface is needed at the desired level of accuracy. Similar conclusions can be reached based on the work of Torheyden and Jansen.[78] The discrepancies between $\Delta E_{int}^{CCSD(T)}(T)$ and $\Delta E_{int}^{CCSD(T)}(TQ)$ are generally much smaller than either of the latter quantities. It is safe to conclude that, for important configurations, the high-order (beyond MBPT2) CCSD(T) correlation contributions to $E_{int}(extr)$ computed in the aug-cc-pVTZ basis differ by less than about 0.02 kcal/mol from the true values. It is also interesting to note that, except for several attractive configurations, the extrapolated quantity $\Delta E_{int}^{CCSD(T)}(TQ)$ is smaller in absolute value than its counterpart obtained in the aug-cc-pVTZ basis.

Taking into account the saturation of all quantities on the

rhs of Eq. (3) with basis set size, one can estimate the total interaction energy $E_{int}(extr)$ in the well region to be within 0.05 kcal/mol of the exact CCSD(T) value. Whereas this error estimate has been obtained by adding the 0.03 kcal/mol error in $\Delta E_{int}^{MBPT2}(TQ)$ to the 0.02 kcal/mol error in $\Delta E_{int}^{CCSD(T)}(T)$, the two errors are in most cases of opposite sign, so that the 0.05 kcal/mol estimate may be too conservative. The difference between the CCSD(T) interaction energy and the full configuration interaction (FCI) result is not known for the water dimer, not even in a small basis set. The only information about the effects beyond CCSD(T) is from the BD(TQ) method. The difference between the BD(TQ) and CCSD(T) interaction energies, which mainly accounts for the effects of "true" quadruple excitations, is only about 0.02 kcal/mol.[41] Thus, the potential given by Eq. (3) should be quite close to the exact Born-Oppenheimer surface, to within 0.07 kcal/mol or better. For comparison, the accuracy of the SAPT-5s potential was about 0.3 kcal/mol. The accuracy estimate of our CCSD(T) calculations is essentially the same as that proposed by Tschumper *et al.*[41] These authors computed ΔE_{int}^{MBPT2} in larger basis sets, up to $X=6$ (although without midbond functions), but the extrapolations based on the TQ results appear to be almost as accurate as those based on the 56 ones. The $\Delta E_{int}^{CCSD(T)}(T)$ contribution was obtained in both cases from $X=T$ calculations. Note that the present work achieves this accuracy for the complete surface, whereas Ref. 41 gets it only for a few points on the surface.

Results of all *ab initio* calculations described above are available as a part of the Supporting Information.[104]

B. Asymptotic region of the potential

In the region of large intermolecular separations R, the multipole expansion of the interaction operator can be used to obtain the interaction energy as an asymptotic series in inverse powers of R with coefficients (the so-called van der Waals constants) dependent only on the properties of the monomers: multipole moments and polarizabilities (static and dynamic). The asymptotic information can be built into the analytic representation of the potential energy surface. To this end, one needs to compute multipoles and van der Waals constants at the level consistent with the short-range calculations.

The tesseral[105,106] multipole moments Q_{lm} through $l=7$ were calculated in a manner analogous to Eq. (3) as

$$Q_{lm} = Q_{lm}^{SCF}(Q) + \Delta Q_{lm}^{MBPT2}(TQ) + \Delta Q_{lm}^{CCSD(T)}(T), \qquad (5)$$

where the correlation contributions in basis set X are defined as

$$\Delta Q_{lm}^{MBPT2}(X) = Q_{lm}^{MBPT2}(X) - Q_{lm}^{SCF}(X), \qquad (6)$$

$$\Delta Q_{lm}^{CCSD(T)}(X) = Q_{lm}^{CCSD(T)}(X) - Q_{lm}^{MBPT2}(X), \qquad (7)$$

and $\Delta Q_{lm}^{MBPT2}(TQ)$ denotes the MBPT2 contribution extrapolated as in Eq. (4). All multipole moments in the expressions above are calculated with the inclusion of orbital relaxation, i.e., as derivatives of the total electronic energy of the molecule computed using a given method with respect to the

proper multipolar field. In the case of the SCF and MBPT2 methods, such a differentiation can be performed analytically (in the former case, it is equivalent to calculating the multipole moments as average values with the SCF determinant) and we used the program MOMDRE from the POLCOR suite of codes by Wormer and Hettema[107,108] for this purpose. To obtain the $\Delta Q_{lm}^{\text{CCSD(T)}}$ contributions, the finite-field technique was used to get $Q_{lm}^{\text{CCSD(T)}}$, and then Q_{lm}^{MBPT2} computed by MOMDRE (via analytic differentiation) was subtracted. The value of the field parameter in the finite-field CCSD(T) calculations ranged from 10^{-5} to 10^{-6} a.u., which ensured the stability of the numerical derivative to within a few percent. The dipole moment of the water molecule resulting from this procedure is equal to -0.7321 a.u., virtually identical to the experimental value of -0.73 a.u.[109] Our CCSD(T) values of the quadrupole moment components Q_{20} and Q_{22} were equal to -0.1174 and 2.1969 a.u., respectively, which can be compared to the experimental estimates[110] of -0.20 and 2.20 a.u., respectively.

The expressions for the van der Waals constants describing the dispersion and induction interactions in the asymptotic region are not available at the level consistent with the CCSD(T) calculations. We therefore used the POLCOR suite[107,108] to compute these coefficients at the level consistent with the currently implemented SAPT approach (see Ref. 73 for details), asymptotically equivalent to the supermolecular MBPT4 level. Here again, the SCF (more strictly, coupled or uncoupled Hartree–Fock for induction and dispersion components, respectively) part of each coefficient was computed in the aug-cc-pVQZ basis set, whereas the correlation part was extrapolated from the aug-cc-pVTZ and aug-cc-pVQZ results using the X^{-3} scheme. All the multipole moments and the asymptotic coefficients computed in this work are available in the Supporting Information.[104]

III. FIT

A. Functional form

The interaction energies $E_{\text{int}}(\text{extr})$ of Eq. (3) were fitted to a function consisting of a site-site component and a term representing the induction interaction,

$$V = \sum_{a \in A, b \in B} u_{ab}(r_{ab}) + V_2^{\text{ind}}(A,B). \tag{8}$$

The individual site-site "potentials" u_{ab}, depending on the distances r_{ab} between two sites located on different monomers, have the same form as in Ref. 74,

$$u_{ab}(r_{ab}) = \left(1 + \sum_{m=1}^{3} a_m^{ab} r_{ab}^m\right) e^{\alpha_{ab} - \beta_{ab} r_{ab}} + f_1(\delta_1^{ab}, r_{ab}) \frac{q_a q_b}{r_{ab}}$$

$$+ \sum_{n=6,8,10} f_n(\delta_n^{ab}, r_{ab}) \frac{C_n^{ab}}{r_{ab}^n}. \tag{9}$$

The functions u_{ab} can be viewed as extensions of the Buckingham-type potential, containing the exponential term, the electrostatic term with site charges q_a and q_b, and the r_{ab}^{-n} terms with coefficients C_n^{ab} responsible for the asymptotic part of the dispersion interaction as well as the remainder of

asymptotic induction energy not recovered by V_2^{ind}. To prevent the asymptotic terms from diverging at short intermolecular distances, these terms (except for V_2^{ind}) are multiplied by the Tang-Toennies damping functions,[111]

$$f_n(\delta, r) = 1 - e^{-\delta r} \sum_{m=0}^{n} \frac{(\delta r)^m}{m!}, \tag{10}$$

which are close to 1 for large r, but continuously decay to zero when r decreases.

The sites involved in the summation in Eq. (8) have been described in detail and listed in Table I of Ref. 74. These are the O and H atoms, four additional off-atomic sites located in the plane perpendicular to the molecule (two D1 and two D2 sites), and one site near the center of mass (D3). The total number of sites on each monomer is thus 8, and these sites can be classified in five symmetry-unique types. Not all types of sites carry all types of interactions implied by formula (9). Specifically, the charges of the D2 and D3 sites are set equal to zero, which implies that the damping parameters δ_1^{ab} with either a or b corresponding to any of these site types need not be considered. Likewise, only the O and H atomic sites participate in the modeling of the long-range dispersion interactions, so that the only nonvanishing C_n^{ab} parameters correspond to $ab=$OO, HH, or OH, and similarly for $\delta_n^{ab}, n = 6,8,10$. The D1 site type is excluded from the exponential-type interactions. Overall, the site-site part of V is determined by 77 adjustable parameters (3 charges q_a, 10 each of α_{ab} and β_{ab} parameters, 9 asymptotic coefficients C_n^{ab}, 15 damping parameters δ_n^{ab}, and 30 polynomial coefficients a_m^{ab}).

The term $V_2^{\text{ind}}(A,B)$ (where A,B stands for the positions and orientations of the two molecules) describes the induction component of the interaction energy by means of a one-center polarization model. In this model, an induced dipole is generated on a single polarizable site located on each molecule as a result of interaction with the electric field of the other molecule(s). In our model, the position of the polarizable center is the same as in the three-body potential of Ref. 81, i.e., near the center of mass of the molecule, 0.194 690 6 Å away from the oxygen atom towards the hydrogens (this value was optimized in the process of fitting the three-body potential). Rather than to define V_2^{ind}, we shall consider the general case of N water molecules. In such a system, the N-body induction energy in the polarization model is given by

$$V_N^{\text{ind}} = -\frac{1}{2} \sum_{i=1}^{N} \boldsymbol{E}_i \cdot \boldsymbol{\mu}_i^{\text{ind}}, \tag{11}$$

where it is understood that V_N^{ind} depends on the positions and orientations of all molecules. Note that there is no damping involved in formula (11). The electrostatic field $\boldsymbol{E}_i$ on the polarizable center of molecule i is generated by the *static* charges [q_x in Eq. (9)] of all the other molecules, and the induced dipole moments $\boldsymbol{\mu}_i^{\text{ind}}$ satisfy the set of coupled equations

$$\boldsymbol{\mu}_i^{\text{ind}} = \alpha \left(\boldsymbol{E}_i + \sum_{j \neq i}^{N} \boldsymbol{T}_{ij} \boldsymbol{\mu}_j^{\text{ind}} \right). \tag{12}$$

In the equation above, $\boldsymbol{T}_{ij}$ is the dipole-dipole interaction matrix,

$$\boldsymbol{T}_{ij} = -\frac{1}{r_{ij}^3} \left(1 - 3 \frac{\boldsymbol{r}_{ij} \otimes \boldsymbol{r}_{ij}}{r_{ij}^2} \right), \tag{13}$$

where $\boldsymbol{r}_{ij}$ is the vector connecting the polarizable centers of molecules i and j, the symbol $\otimes$ denotes a tensor product, and $\alpha = 9.922$ bohr^3 is the experimental average dipole-dipole polarizability of the molecule given by Murphy.[112] The value of this quantity computed at the theory level consistent with the currently implemented SAPT approach,[73] i.e., the coupled Hartree–Fock polarizability plus the (true) MBPT2 contribution,[107,108] after applying the (TQ) CBS extrapolation is equal to 9.5615 bohr^3. The most accurate theoretical value was computed at the CCSD(T) level by Maroulis[113] using the equilibrium monomer geometry and amounts to 9.62 ± 0.03 bohr^3. To precisely compare theory with experiment, one has to add a correction for zero-point energy (ZPE) vibrations to theoretical results and to extrapolate the experimental dynamic polarizability[112] measured at 514.5 nm to zero frequency. The ZPE correction was computed by Maroulis[113] and by Avila[114] who obtained values of 0.29 and 0.31 bohr^3, respectively. An often quoted extrapolation to zero frequency, resulting in a static polarizability of 9.83 ± 0.02 bohr^3, was performed by Russell and Spackman[115] from refractivities measured by Cuthbertson and Cuthbertson.[116] However, if newer measurements by Newbound[117] are used, the extrapolated value is 9.703 ± 0.013 bohr^3. Note that Newbound's extrapolation of the Cuthbertson and Cuthbertson data resulted in a static polarizability of 9.742 bohr^3. The dynamic polarizability at 514.5 nm given by Newbound's fit [Eq. (6) in Ref. 117 with the second coefficient divided by 10] is 9.907 bohr^3, 0.015 bohr^3 different from Murphy's value. If the static polarizability of Ref. 117 is shifted by this difference, one obtains 9.718 bohr^3 as an alternative experimental estimate of this quantity.

The set of coupled equations [Eq. (12)] may be solved using a matrix inversion technique, or, as it was done here, in an iterative manner until the induced dipoles converge. As the convergence criterion, we used a threshold of 10^{-20} (a.u.)^2 on the sum of squared differences between the induced dipoles in two consecutive iterations.

The employed fitting strategy (described further on) implies that the physical interpretation of the individual terms in Eq. (9) is possible only in the large-R asymptotic region. For shorter intermolecular distances, such as those near the minimum, *all* terms in Eq. (9) together with the V_2^{ind} term are used to model the potential and only the *sum* of all these terms is meaningful.

The functional form of the fit, Eq. (8), differs from the one used in the SAPT-5s potential of Ref. 74 in that the bulk of the induction component of the interaction energy has been separated out as the "polarizable" V_2^{ind} term instead of being bundled with the dispersion component in a site-site

expansion in powers of r_{ab}^{-1}. Such an approach has two advantages. First, V_2^{ind} exactly reproduces the leading isotropic (R^{-6}) term in the asymptotic expansion of the induction energy [since our CCSD(T) dipole moments agree to all digits with experiment and we use the experimental polarizability] and the leading anisotropic terms are correct within a few percent [only a small error results here from using the isotropic polarizability instead of the complete polarizability tensor, as this tensor is nearly isotropic]. In fact, we have checked that the use of V_2^{ind} improves a separate fit to the SAPT induction energy also for finite separations. Thus, although the r_{ab}^{-n} expansion still has to compensate for the approximations involved in V_2^{ind}, the main task of this expansion is to describe only the dispersion asymptotics, which improves the accuracy of the fit. The second advantage of using an explicit "polarizable" representation of the induction component is the ease of generalization to systems larger than the dimer. The total N-body polarization energy V_N^{int} of Eq. (11) automatically accounts for the majority of nonadditive induction effects, of crucial importance in water.

B. Fitting procedure

Adjustments of the parameters in Eq. (9) proceed in three stages. First, the charges q_x are obtained for $x = $O, H, and D1 by fitting to the set of multipole moments of the water molecule, calculated as described in Sec. II A. All multipoles through $l = 7$ were used in the fit, with extra weight put on the total charge (i.e., the multipole $l = 0$) to keep the molecule electrically neutral.

The asymptotic parameters C_n^{ab} were obtained in the second stage of the fit. The sum of the asymptotic dispersion and induction energies was first computed *ab initio* on a set of dimer geometries obtained from the original set of 2510 points by adding 3 Å to the R coordinate of each geometry. To perform this computation, we used the one-center asymptotic expansion[118] with van der Waals constants through R^{-12} obtained as described in Sec. II B and Ref. 73. From the asymptotic energies computed in this way on the grid of geometries, the term V_2^{ind} was subtracted, and the result was least-squares fitted to the (undamped) r_{ab}^{-n} expansion with C_n^{ab} as free parameters. In this way, the part of induction energy already reproduced by V_2^{ind} is not accounted for by C_n^{ab}. The main task of the latter parameters is to model the dispersion energy in the asymptotic region. During the fitting of the site-site asymptotic coefficients, it was ensured that the leading isotropic (R^{-6}) term of the *ab initio* dispersion plus induction asymptotics is reproduced exactly. This was accomplished by imposing a constraint $C_6^{\text{OO}} + 4(C_6^{\text{OH}} + C_6^{\text{HH}}) = 47.0829$ a.u., where the number on the rhs is the sum of the *ab initio* values of the isotropic dispersion and induction C_6 van der Waals constants (47.053 232 and 10.665 173 a.u., respectively) minus $2\alpha\mu^2 = 10.6355$ a.u. The latter quantity (with the experimental isotropic polarizability $\alpha = 9.922$ bohr^3, Ref. 112, and the dipole moment $\mu = 0.732\,090\,2$ a.u. given by our static charges) is the coefficient of the isotropic R^{-6} component of V_2^{ind}.

All the remaining parameters in Eq. (9) were adjusted in the third and final stage, by least-squares fitting to the

TABLE I. Parameters of the global minimum of water dimer. Distances in Å, angles in degrees, energies in kcal/mol. See Ref. 41 for definitions and atom labeling. Square brackets around $r_{H_2O_1}$ indicate that this parameter was fixed during the optimization.

	$r_{O_1O_2}$	$r_{H_2O_1}$	$\theta_{O_2O_1H_2}$	$\theta_{O_1O_2Bisec}$	Energy
CC-pol	2.9107	[0.9716]	6.48	122.66	−5.103[a]
CCSD(T)[b]	2.909	0.9653	4.47	124.92	−5.04 ± 0.08
CCSD(T) + $\delta_{BD(TQ)}$[c]					−5.02 ± 0.08
CCSD(T)[d]	2.912 ± 0.005	0.9639	5.5	124.4	−5.02 ± 0.05
CCSD(T)-PES[e]	2.937	0.9675	5.9		−4.762
SAPT-5s[f]	2.9549	[0.9716]	6.36	127.16	−4.861
SAPT-5st[g]	2.9244	[0.9716]	6.95	121.5	−5.029
VRT(ASP-W)[h]	2.924	[0.9572]	−2.06	131.5	−4.91
VRT(ASP-W)II[i]	2.952	[0.9572]	−2.40	131.5	−4.85
Experiment	2.91 ± 0.005[j]		−1 ± 10[k]	123 ± 10[k]	−5.00 ± 0.7[j]

[a]Energy calculated *ab initio* using Eq. (3). The fit at the minimum predicts −5.097 kcal/mol.
[b]Tschumper *et al.*, Ref. 41, CP-uncorrected optimization using CCSD(T) method in TZ2P(f,d)+dif basis set. Energy computed at the CCSD(T) level with CBS-extrapolated MBPT2 component and with a correction for core effects.
[c]As in footnote b plus $\delta E_{BD(TQ)}$ correction.
[d]Klopper *et al.*, Ref. 46, two-dimensional optimization using CBS-extrapolated CCSD(T).
[e]Huang *et al.*, Ref. 80, fit to CCSD(T) results in aug-cc-pVTZ basis. Values of $r_{O_1O_2}$ and $\theta_{O_2O_1H_2}$ computed by us from data in Table II of Ref. 80.
[f]Mas *et al.*, Ref. 74, SAPT-5s fit to 2510 energies computed using SAPT method.
[g]Reference 76, SAPT-5s fit with a subset of parameters tuned to reproduce one transition in the experimental VRT spectrum of the dimer.
[h]Reference 69, ASP-W potential with a subset of parameters tuned to reproduce the experimental VRT spectrum of the dimer.
[i]Reference 70, refit of VRT(ASP-W) potential.
[j]Mixed experimental-theoretical estimate from Ref. 74.
[k]The values extracted by Odutola and Dyke (Ref. 120) from rotational constants, i.e., averaged over the ground rovibrational state.

E_{int}(extr) results computed at 2510 geometries according to Eq. (3). During this nonlinear fit, the site charges q_a and the coefficients C_n^{ab}, obtained independently in the previous stages, were kept constant, which ensures that the asymptotic behavior of the potential is not altered by the final fitting. A variant of the Powell minimization routine was used to explore the space of nonlinear parameters α_{ab}, β_{ab}, and δ_n^{ab}, $n = 1, 6, 8, 10$. For each set of these parameters, the coefficients a_m^{ab}, with $m = 1, 2, 3$, were obtained by a linear least-square fit. Each data point was weighted[74] with a factor $w = 1/(E_{int} + 5.5 \text{ kcal/mol})^2$, favoring the low-energy configurations. The unweighted root-mean-square deviation (rmsd) of the resulting fit calculated for points with E_{int}(extr) < 0 was 0.09 kcal/mol, much smaller than the overall rmsd of 0.42 kcal/mol. Thus, the errors introduced by the fitting procedure in the negative energy region are of the same order of magnitude as the uncertainty of the *ab initio* calculations. The rmsd of our fit is similar to that of the overlap-model fit of Torheyden and Jansen[78] amounting to 0.29 kcal/mol for the 500 points considered there. On the other hand, Huang *et al.*[80] fitted their data more accurately as the rmsd for 9257 flexible-monomer grid points with energies up to 10 000 cm^{-1} above the minimum of the 12-dimensional potential was only 10.5 cm^{-1} (0.03 kcal/mol). This is due to the fact that the functional form of the fit used by Huang *et al.*, after the internal degrees of freedom are frozen, contains 108 free parameters compared to our 45 parameters (not counting the precomputed asymptotic parameters) and is thus more flexible. We could have made our fitting form

more flexible, but it would make MC/MD calculations too time consuming.

To emphasize the "polarizable" character of the fit obtained as described above and the fact that the *ab initio* data came from a coupled-cluster calculation, we shall refer to this potential surface as CC-pol. All the fit parameters as well as a FORTRAN routine computing the CC-pol PES are available in the Supporting Information.[104]

IV. STATIONARY POINTS

The CC-pol potential energy surface was explored using an implementation of the eigenvector following procedure[119] with derivatives calculated numerically. The list of stationary points found coincides with the set of structures numbers 1 through 10 considered in Refs. 40 and 41. We refer the reader to Fig. 1 of Ref. 41 for a graphical representation of these structures and the atom labeling used in the tables below.

In accordance with multiple previous studies, the global minimum is the nonplanar open structure number 1 with C_s symmetry and one hydrogen bond. The parameters of this structure as obtained from the CC-pol potential are shown in Table I and compared to their counterparts resulting from other theoretical models as well as to the experimental estimates. The results of Tschumper *et al.*[41] are of special interest here as these have been obtained using basis sets and a theory level very similar to ours. A direct comparison is, however, rather difficult, since the study of Ref. 41 involved full optimizations, including the intramonomer degrees of

freedom. This results, in particular, in monomer geometries generally close to the equilibrium isolated monomer geometries, whereas we use vibrationally averaged geometries. Nevertheless, the agreement between the obtained intermolecular geometric parameters of Ref. 41 and those from CC-pol is remarkably good. In the unrestricted optimizations, the only intramonomer parameter which departs substantially from the isolated-monomer equilibrium value is the length of the OH bond involved in the hydrogen bond $r_{H_2O_1}$. As shown in Table I, this distance from the calculations of Refs. 41 and 46 becomes quite close to the vibrationally averaged value used in our calculations. Both sets of results, the ones from the CC-pol potential and the ones from Ref. 41, agree very well with the geometry of Klopper *et al.*,[46] obtained using the CP-corrected CBS-extrapolated CCSD(T) method, as well as with the experimental estimates within the error bars. (In Refs. 96, 73, and 74, the experimental value of the angle $\theta_{O_2O_1H_2}$, describing the nonlinearity of hydrogen bond and referred to as α, was incorrectly quoted as $5 \pm 10°$.) Also shown in Table I are geometries obtained from the SAPT-5s potential of Ref. 74 and its semiempirical counterpart SAPT-5st,[76] obtained by "tuning" SAPT-5s to the VRT spectrum. While the angular parameters from SAPT-5s agree well with those from CC-pol, the O–O distance $(r_{O_1O_2})$ is by 0.05 Å too large. In Ref. 74, it was suggested that this increase might be caused by using the vibrationally averaged, "r_0," monomer geometry in the *ab initio* calculations leading to the SAPT-5s surface. However, since the same monomer geometry is employed in CC-pol, it is now clear that the too large O–O separation in the SAPT-5s minimum configuration must be due to deficiencies in basis sets and theory levels employed in Ref. 74. It is interesting that the O–O distance obtained from the recent VRT(ASP-W)II empirical dimer potential,[70] fitted to spectral data, is very similar to the one predicted by the SAPT-5s potential, and therefore somewhat too large. On the other hand, the $r_{O_1O_2}$ prediction of the older version of the fit, VRT(ASP-W),[69] virtually identical to that from the SAPT-5st potential, agrees much better with the best *ab initio* estimates. Both VRT(ASP-W) and VRT(ASP-W)II potentials agree with experiment in predicting the two H atoms of the donor molecule being on the opposite sides of the O–O axis (negative $\theta_{O_2O_1H_2}$), while all *ab initio* methods considered and SAPT-5st predict the converse trend. On the other hand, the *ab initio* derived $\theta_{O_1O_2Bisec}$ angles (tilt of the acceptor molecule with respect to the O–O line) are closer to experiment than the ones from VRT(ASP-W) potentials. However, one should point out that the theoretical angles are for the equilibrium of the potential, whereas the experimental ones were obtained by fitting the measured dimer rotational constants.[120] Odutola and Dyke[120] estimated the uncertainties of the angles with respect to equilibrium values due to the effects of the ground-state vibrational motion to be less than 5°.

Comparison of the CC-pol interaction energy at the minimum to the literature values from *ab initio* optimizations is not straightforward, as the latter are computed as differences between the dimer energy obtained from all-coordinate optimization and the monomer energies in monomer's equilibrium geometry. However, the authors of Ref. 46 also per-

formed optimizations with monomers fixed at their equilibrium geometries and using the same theory level and basis sets as for the result quoted in Table I. The result of this optimization (Table V of Ref. 46) is −4.98 kcal/mol. If the equilibrium monomer geometry is replaced by the r_0 one (in both the dimer and the monomers), the (vertical) interaction energy becomes more negative by about 0.12 kcal/mol, as shown in Ref. 96, which leads to −5.10 kcal/mol, in almost perfect agreement with the CC-pol result. This agreement is, in fact, not surprising considering that both CC-pol and the calculations of Ref. 41 used very similar theory and basis set levels. As already pointed out in Ref. 74, the SAPT-5s minimum is too shallow by about 0.3 kcal/mol. Similarly as with the O–O distance, the depth of the SAPT-5s minimum is in agreement with that obtained from the VRT(ASP-W)II potential. The more accurate VRT(ASP-W)III fit,[70] as well as the original VRT(ASP-W),[69] both predict the well depth in better agreement with the best *ab initio* results [note that neither the dimer geometry nor the tunneling barriers were published for VRT(ASP-W)III (Ref. 70)]. Also the "tuning" of SAPT-5s leads to deepening of the well.

In addition to locating the minimum structure of the dimer, we performed a more thorough exploration of the potential surface to find the nine other stationary points. Parameters characterizing the intermolecular geometry of these structures, as defined in Ref. 41, are collected in Table II and compared to their counterparts obtained from the full CCSD(T) optimizations in the TZ2P(f,d)+dif basis set.[41] The energies of the stationary points relative to the global minimum (transition barriers) are presented in Table III.

Despite the difference in treatment of the intramonomer coordinates in both optimizations, the agreement between the intermolecular geometries obtained is generally very good. The $r_{O_2H_2}$ distances differ by no more than a few hundredths of angstroms and the angular parameters by a few degrees. Consistent with the conclusions of Ref. 41, the lowest-energy saddle point on the CC-pol surface is the structure number 2 of C_1 symmetry, while the flat C_s structure number 3 turns out to be a stationary point of index 2. There is quite a substantial difference in the $\tau_{H_4O_2O_1H_2}$ dihedral angle between the number 2 structures from the two methods, which is a manifestation of the flatness of the potential in this region. Another noticeable difference can be observed in the case of the triply H-bonded structure number 7, for which the CC-pol potential gives an $r_{H_2O_2}$ distance longer by 0.2 Å than the CCSD(T) *ab initio* optimization. Despite the similarity between the parameters of structure number 5 from this optimization and from CC-pol, the latter predicts this structure to be of index 1 rather than 2. For all the other structures, the numbers of negative Hessian eigenvalues are in agreement between the two methods.

Table II also contains the geometric parameters obtained from the SAPT-5s surface. We extended the explorations of this surface described in Ref. 74 to account for stationary points of index higher than 1. We were able to locate all such structures considered in Ref. 41 except for number 5. Also, as concluded before, SAPT-5s misses the C_1 saddle point number 2 and predicts the flat structure number 3 as a saddle

TABLE II. Intermolecular parameters of the stationary-point structures of water dimer. Distances in Å, angles in degrees. See Ref. 41 for definitions and atom labeling. For each structure, the first line corresponds to the CC-pol potential of this work, the second line to the SAPT-5s potential of Ref. 74, and the third line to the CCSD(T) optimizations of Ref. 41.

Structure	Index	$r_{H_2O_2}$	$\theta_{O_1H_2O_2}$	$\theta_{H_3O_2O_1}$	$\theta_{H_4O_2O_1}$	$\tau_{O_2H_2O_1H_1}$	$\tau_{H_3O_2O_1H_2}$	$\tau_{H_4O_2O_1H_2}$
No. 1 nonplanar open C_s	0	1.9484	170.30	109.25	109.25	180.00	123.01	−123.01
	0	1.9921	170.54	111.66	111.66	180.00	121.59	−121.59
	0	1.9485	172.92	110.50	110.50	180.00	122.37	−122.37
No. 2 open C_1	1	1.9619	168.84	112.48	140.48	155.33	166.32	7.45
	…	…	…	…	…	…	…	…
	1	1.9724	168.97	107.33	135.00	144.61	159.63	25.88
No. 3 planar open C_s	2	1.9633	167.82	112.26	143.06	180.00	180.00	0.00
	1	2.0086	167.40	111.95	143.36	180.00	180.00	0.00
	2	1.9813	167.59	109.96	145.00	180.00	180.00	0.00
No. 4 cyclic C_i	1	2.2631	114.30	127.71	47.30	−128.20	106.09	180.00
	1	2.3135	113.40	131.50	48.32	−133.14	109.54	180.00
	1	2.2796	114.84	132.32	47.15	−134.78	111.86	180.00
No. 5 cyclic C_2	1	2.2665	111.89	143.16	49.18	−150.09	−116.64	−170.20
	…	…	…	…	…	…	…	…
	2	2.2810	112.52	145.10	48.95	−153.28	−118.27	−167.68
No. 6 cyclic C_{2h}	3	2.2691	109.84	155.48	50.79	180.00	180.00	180.00
	3	2.3122	109.66	155.91	51.22	180.00	180.00	180.00
	3	2.2756	110.27	155.80	50.66	180.00	180.00	180.00
No. 7 triply H-bonded C_s	2	3.1998	69.21	54.40	54.40	180.00	−103.17	103.17
	2	3.1551	73.23	55.13	55.13	180.00	−105.22	105.22
	2	2.9997	77.32	55.16	55.16	180.00	−108.75	108.75
No. 8 doubly bifurcated C_{2h}	3	3.1673	91.56	71.52	71.52	−64.34	180.00	66.83
	3	3.1981	91.46	71.77	71.77	−64.73	180.00	67.08
	3	3.1140	92.24	70.84	70.84	−64.02	180.00	67.93
No. 9 nonplanar bifurcated C_{2v}	1	2.5610	110.18	127.65	127.65	0.00	−90.00	90.00
	1	2.5890	110.37	127.66	127.66	0.00	−90.00	90.00
	1	2.5154	112.03	127.79	127.79	0.00	−90.00	90.00
No. 10 planar bifurcated C_{2v}	2	2.7149	111.19	127.65	127.65	0.00	0.00	180.00
	2	2.7209	111.24	127.66	127.66	0.00	0.00	180.00
	2	2.6830	112.92	127.95	127.95	0.00	0.00	180.00

point of index 1. Geometries of the remaining SAPT-5s stationary points are quite consistent with the two other sets.

The stationary-point structures, especially the ones of index 1, are important from the point of view of the dimer dynamics, since these structures lie on the paths of tunneling between different equivalent minima. The tunneling gives rise to splittings and shifts of energy levels with magnitudes related to the barrier heights, i.e., the differences between the interaction energies of the saddle points and that of the minimum configuration. The barrier heights for all stationary-point structures resulting from the CC-pol potential are presented in Table III and compared to the predictions of the *ab initio* optimizations of Ref. 41 as well as of the SAPT-5s and SAPT-5st potentials,[74,75] the recent potentials of Refs. 78 and 80, and two empirical VRT potentials fitted to the dimer spectrum. In addition to the barriers resulting directly from the CC-pol fit, the corresponding *ab initio* values are also shown in column 3, obtained by applying Eq. (3) at the geometries found on the fitted surface. Discrepancies between

these two sets of CC-pol barriers are typically below the rmsd of the CC-pol fit (0.09 kcal/mol or 31 cm^{-1} for the subset of data with $E_{int} < 0$) and usually amount to less than 10% of the barrier itself, except for the C_i structure number 4, for which the *ab initio* barrier is 18% higher than the fit prediction. The lowest saddle point number 2, which according to the fit is practically isoenergetic with the C_s structure number 3, turns out to be 15 cm^{-1} lower than number 3 if *ab initio* energies are used. This energy gap between structure numbers 2 and 3 is almost the same as predicted by *ab initio* CCSD(T) optimizations of Ref. 41. In general, the agreement between the barriers from the optimizations of Ref. 41 and those from CC-pol is quite good, considering that the energies have been obtained for slightly different dimer geometries and more substantially different monomer geometries. However, it is also likely that the agreement with Ref. 41 would be improved if the accuracy of our fit in reproducing the *ab initio* points were higher. Limitations of the accuracy for the same fitting form as used for CC-pol

TABLE III. Interaction energies (in cm^{-1}) at the stationary-point structures of the water dimer relative to the minimum. See Fig. 1 in Ref. 41 for a graphical representation of all structures.

Structure	CC-pol[a]	E_{int}(extr)[b]	CCSD(T)[c]	SAPT-5s[d]	SAPT-5st[e]	TJ[f]	HBB[g]	VRT[h]	VRT[i]
No. 2 open C_1	188	199	181 ± 5			161	164	157	154
No. 3 planar open C_s	189	214	198 ± 6	156	222	263	197		
No. 4 cyclic C_i	211	257	245 ± 15	185	248	456	245	207	229
No. 5 cyclic C_2	331	344	333 ± 15			603	329		
No. 6 cyclic C_{2h}	352	365	348 ± 18	280		594	347		
No. 7 triply H-bonded C_s	747	702	634 ± 18	659		839	601		
No. 8 doubly bifurcated C_{2h}	1334	1297	1249 ± 21	1244		1422	1186		
No. 9 nonplanar bifurcated C_{2v}	731	703	625 ± 16	636	685	829	589	394	430
No. 10 planar bifurcated C_{2v}	1006	1023	948 ± 23	897		1123	897		

[a]Predictions of the CC-pol fit.
[b]*Ab initio* calculations according to Eq. (3) at the geometry predicted by the CC-pol fit.
[c]Reference 41.
[d]Reference 74.
[e]Reference 76.
[f]Torheyden and Jansen, overlap-model fit. Reference 78.
[g]CCSD(T)-PES, flexible monomers. Reference 80.
[h]VRT(ASP-W). Reference 69.
[i]VRT(ASP-W)II. Reference 70.

have been discussed in Ref. 79. Although the well depth of the fit of Huang *et al.*[80] is by 91 cm^{-1} too low, cf. Table I, the barriers given by this fit are in significantly better agreement with Ref. 41 than our results (discrepancies ranging from below 1% to 9% whereas those of CC-pol range from 1% to 18%). One reason is certainly that the stationary point search by Huang *et al.* is performed with flexible monomers, just as in Ref. 41, but the other reason could be the better reproduction of the *ab initio* points by their fit. In the case of the SAPT-5s results, the lowest three barriers (for structure numbers 3, 4, and 6) are significantly lower than their counterparts predicted by the other two methods. The small height of the number 3 barrier, primarily responsible for the acceptor tunneling,[75,76] may be the reason for the corresponding tunneling splitting calculated with the SAPT-5s surface being too large compared to experiment. It is worth noting in this context that "tuning" of the SAPT-5s potential to reproduce this splitting[76] resulted in both numbers 3 and 4 barriers very close to the current CC-pol values, as seen from the SAPT-5st column of Table III. On the other hand, the VRT(ASP-W) (Ref. 69) and VRT(ASP-W)II (Ref. 70) empirical potentials produce barrier number 2 similar to barrier number 3 of SAPT-5s. These two barriers can be compared to one another since the geometries of the corresponding saddle points are very similar. Barrier number 4 of the former potentials is in turn larger than that of SAPT-5s and fairly close to that of CC-pol. On the other hand, the bifurcation barrier number 9 is predicted by the empirical potentials to be about 200 cm^{-1} lower than those obtained from *ab initio* potentials, but it should be noted that the bifurcation shifts in the spectra that are sensitive to this barrier are very badly reproduced by the VRT(ASP-W)III potential.[121] Although the barriers given by the four discussed potentials are fairly different, the spectra are of comparable accuracy.[66] Clearly, the heights of the barriers are not the only factors responsible for the magnitudes of the splittings; the barrier widths and coupling between different coordinates also play major roles. It is also interesting to note that the barriers for structure numbers 7 and 10

computed from the SAPT-5s surface are in very good agreement with those of Ref. 41, while the corresponding CC-pol results are somewhat higher. The overlap-model fit of Torheyden and Jansen[78] gives somewhat larger discrepancies with the barriers of Ref. 41 than the other potentials, ranging from 11% to 86%. On the other hand, although the SAPT-5s potential produces smaller discrepancies (from below 1% to 24%), it misses two saddle points whereas the fit of Ref. 78 recovers all of them. To our knowledge, the barriers of the TTM2 potentials[63–65] have not been published. For their predecessor, the TTM potential,[122] the barriers were found[123] for structure numbers 4, 5, and 9 and are in 27%, 13%, and 44% error compared to the results of Ref. 41.

V. OVERALL COMPARISON TO SAPT-5s

One of the advantages of developing interaction potentials from *ab initio* calculations is that the increase of the level of theory, improvements of the quality of the basis set, and use of more flexible fitting functions should, in principle, lead to more accurate predictions. In order to analyze future directions for such developments, it is important to identify the most essential changes in the current step, i.e., to compare the CC-pol potential with the SAPT-5s potential of Ref. 74. Section IV already highlighted some important differences between CC-pol and SAPT-5s pertaining to the predicted characteristic points and energy barriers. Here we take a more global approach and compare the presently computed interaction energies E_{int}(extr) with their counterparts E_{int}(SAPT) obtained in Refs. 73 and 74 using the SAPT method. Such a comparison is presented in Fig. 2, where the difference between the two energies is plotted against the total interaction energy for a large subset of the 2510 geometries used in the fits. Analysis of the leftmost part of the plot (corresponding to the most attractive geometries) confirms the conjecture of Ref. 74 that the truncation of basis set and theory level resulted in the SAPT interaction energies being a few tenths of kcal/mol too shallow in the minimum region. It
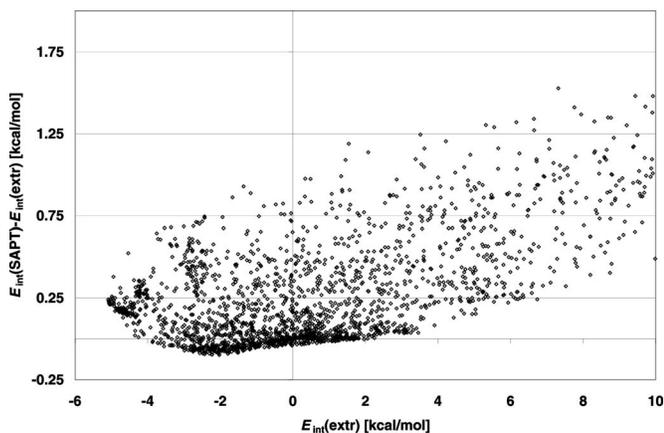
FIG. 2. Comparison of $E_{int}(extr)$ calculated in this work and $E_{int}(SAPT)$ of Ref. 74. Only those of the 2510 dimer geometries are included for which the interaction energy does not exceed 10 kcal/mol.

is clear from Fig. 2 that the $E_{int}(extr)$ energies are more attractive (positive values in Fig. 2) than the SAPT ones for a vast majority (80%) of all 2115 data points shown. In the negative-energy region, the differences are in a few cases as large as 1 kcal/mol, although most points in this region are within 0.3 kcal/mol. In the positive-energy (predominantly short-range) region, the discrepancies become overall more positive. The largest differences, up to about 1.5 kcal/mol, occur on the repulsive wall of the potential, where the radial derivative is large and negative. The onset of the repulsive wall on the $E_{int}(extr)$ surface occurs for slightly shorter intermolecular distances than in the case of the SAPT potential. This phenomenon is illustrated more explicitly in Fig. 3, where radial cross sections through the CC-pol and SAPT-5s fitted surfaces are plotted for three representative angular configurations of the dimer, corresponding to three of the stationary points found on the CC-pol surface (see Sec. IV). For all three configurations, the CC-pol potential is less repulsive than SAPT-5s, allowing the monomers to approach closer. In addition, for structure numbers 1 and 4, the radial minimum on the CC-pol potential is shifted to shorter distances. In the case of configuration number 9, SAPT-5s is

slightly more attractive in the region to the right of the minimum. Regions like this are the source of the small subset of points lying below the horizontal zero axis in Fig. 2. Overall, the difference between the CC-pol and SAPT-5s potentials is smaller for configuration number 9 than for numbers 1 and 4. Treating CC-pol as a benchmark, one concludes that the accuracy of SAPT-5s is uneven in different regions of the configurational space. In particular, the hydrogen-bonded configurations, similar to the global minimum, do not experience sufficient attraction compared to other configurations. Since the former type of configurations are important for the tetrahedral structure of liquid water, it is possible that this deficiency of SAPT-5s results in the problems this potential encounters in reproducing the liquid structure.[87]

Since the main difference between the CC-pol and SAPT-5s potentials is due to the level of theory and the basis set employed (the modification of the fitting function is only of secondary importance), one may ask which of the two factors was more relevant. Calculations of Torheyden and Jansen[78] indicate that the basis set incompleteness is the main source of errors of the SAPT-5s potential. These authors compared SAPT interaction energies computed using the aug-cc-pVTZ basis plus bond functions (a more complete basis set than that used in the development of SAPT-5s) with CBS (TQ)-extrapolated CCSD(T) interaction energies at 494 grid points and found the two sets of results to be very close to each other. They pointed out that this is due to a fortuitous but systematic cancellation of theory and basis set errors and SAPT interaction energies computed in the same basis set as CCSD(T) tend to be more negative. However, for the attractive part of the potential, the differences between the two methods at the CBS limit remain very small, of the order of 0.1 kcal/mol. Thus, in the medium- and long-range regions the CBS picture is in fact similar to that shown in Fig. 3, except that the very good agreement extends to smaller separations. On the other hand, on the repulsive wall, already in the negative interaction energy region, the SAPT results at the CBS level get below the CCSD(T) results, just opposite to the situation shown in Fig. 3, and the discrepancies can be a few tenths of kcal/mol. This trend is mainly due to the use of the S^2 approximation in the second-order exchange corrections of SAPT and to the fact that the exchange quenching of the dispersion energy is incomplete.[124]

VI. SUMMARY AND CONCLUSIONS

A six-dimensional CC-pol interaction potential energy surface for the water dimer has been developed based on *ab initio* calculations utilizing the supermolecular approach at the CCSD(T) level of theory with the MBPT2 component extrapolated from the aug-cc-pVTZ and aug-cc-pVQZ basis sets and terms beyond MBPT2 computed in the aug-cc-pVTZ basis. All basis sets have been supplemented by bond functions. The overall accuracy of the interaction energies obtained in this way is estimated to be better than about 0.07 kcal/mol in the region of the attractive well, a smaller uncertainty than that of any previous water dimer potential. The contribution from the CCSD(T) method beyond the MBPT2 level $\Delta E_{int}^{CCSD(T)}$ was found to be more important
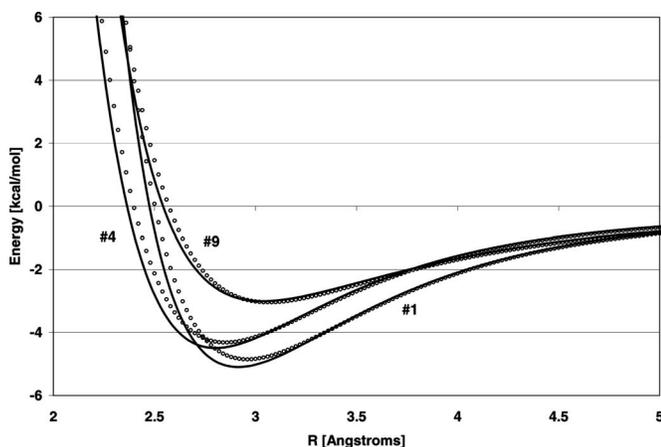

FIG. 3. Radial cross sections through the CC-pol (solid lines) and SAPT-5s (circles) potentials for angular configurations corresponding to three stationary points, as found on the CC-pol surface. R is the distance between the centers of mass of the monomers.

than previously thought. Whereas $\Delta E_{int}^{CCSD(T)}$ is almost negligible at the minimum and at other characteristic points, in large regions of configuration space this contribution exceeds 0.2 kcal/mol. The *ab initio* energies at 2510 geometries have been fitted using a site-site form with a one-center dipole polarization term added to represent the bulk of the induction energy. The latter term is readily generalized to approximately reproduce the nonadditive induction effects in larger clusters and condensed phases. The chosen form of the fitting function represents a compromise between the requirement of fitting the *ab initio* points as closely as possible and of reasonably fast calculations of the potential, critical for MC/MD simulations.

The CC-pol minimum angular configuration agrees with the experimental determination of Ref. 120, whereas the minimum separation and the depth of the potential agree very well with the best mixed experimental/theoretical estimates. The agreement between the current predictions for the minimum and the results of the most accurate literature *ab initio* calculations[41,46] is excellent. To some extent this is due to the use of nearly identical levels of theory and basis sets, but an analysis of the convergence patterns indicates that all these calculations are very close to limit values. In fact, purely theoretical parameters of the minimum configuration are now more accurate than the experimental or mixed parameters. Thus, the present work extends such an accurate treatment of the water dimer from a few characteristic points to the complete potential surface. The VRT(ASP-W) empirical potentials give parameters of the minimum somewhat farther from the benchmark values than the CC-pol ones.

The fitted CC-pol potential surface features the stationary points in very good agreement with the results of the *ab initio* investigations by Tschumper *et al.*[41] using the CCSD(T) method. The $r_{O_2H_2}$ distances differ by no more than a few hundredths of angstroms and the angular parameters by a few degrees. The *ab initio* barriers computed at CC-pol characteristic points differ by less than 12% from those of Ref. 41. This is a very reasonable agreement taking into account that the latter optimizations also vary the intramonomer coordinates. Compared to SAPT-5s, the discrepancies with respect to the barriers of Tschumper *et al.* were reduced by a factor of 2–5 for small barriers (those below 350 cm^{-1}), most relevant for the VRT spectrum. For larger barriers, SAPT-5s actually agrees better. The VRT(ASP-W) potentials show discrepancies in barrier heights exceeding 12% except for one case.

The CC-pol potential can be compared to a recent potential obtained by Huang *et al.*[80] using the CCSD(T) method and the aug-cc-pVTZ basis set. Since this basis set was smaller than that used by us at the MBPT2 level, contained no bond functions, and no CBS extrapolations were performed, our *ab initio* interaction energies should be more accurate than those of Ref. 80. The potential of Huang *et al.* was fitted to almost 20 000 interaction energies calculated at geometries including the variation of the monomer coordinates. The authors of Ref. 80 have not specified how many of these points would correspond to rigid monomers, but if the Euler angles determining the mutual orientation of the monomer and the intramonomer coordinates were sampled evenly,

this would correspond to a few hundred points for a rigid-monomer potential, less than the number used by us. The functional form of the fit used by Huang *et al.* contained relatively more adjustable parameters in the intermonomer coordinates than our form. Consequently, the rmsd of the fit of Ref. 80 was a few times smaller than ours. However, due to the larger number of terms and lack of proper asymptotic properties, the potential of Ref. 80 may be difficult to apply in simulations of liquid water. The barriers predicted by this potential are in most cases in a better agreement with the results of Tschumper *et al.*[41] than the CC-pol barriers. This is probably due both to the account for the monomer-flexibility effects and the better reproduction of the *ab initio* points by the fit of Ref. 80.

Comparison of the CC-pol surface with the SAPT-5s potential of Ref. 74 reveals that the accuracy of the latter is somewhat uneven for different angular configurations of the dimer. In hydrogen-bonded configurations characteristic for the tetrahedral structure, the well predicted by SAPT-5s is by about 0.3 kcal/mol too shallow and the onset of the repulsive wall occurs at intermolecular distances which are by 0.05 Å too large. On the other hand, at some other configurations, such as the stationary point number 9, the agreement between CC-pol and SAPT-5s is much better. The deficiencies of the SAPT-5s potential result mainly from the incompleteness of the basis set used to compute it.

Applications of the CC-pol potential to predictions of water dimer spectra, second virial coefficients, and properties of liquid water were recently presented[66] and more details of this work will be given in a subsequent paper. Comparison of these results to literature data shows that CC-pol (for liquid water, CC-pol was used together with the three-body potential from Ref. 81) simultaneously recovers the whole range of investigated properties with high accuracy. Although some properties are predicted slightly better by one or another *ab initio* potential, each of those potentials produces very large discrepancies with experiment on some other properties. Furthermore, all these properties are predicted by CC-pol with an accuracy comparable to that achieved by the empirical potentials fitted to a particular set of data. For example, the CC-pol predictions of liquid water properties are similar to those given by the empirical potentials fitted in MD simulations, whereas the predictions for the water dimer are similar to those given by the potentials fitted to the water dimer spectra. By contrast, the empirical liquid-optimized potentials perform very poorly in recovering the properties of the water dimer (dimer minimum geometry, spectra, and virial coefficients). The semiempirical potentials fitted to dimer spectra, SAPT-5st (Ref. 75) and VRT(ASP-W)III,[70] perform much better in such "across the board" comparisons, but still not as good as CC-pol. For example, CC-pol predicts dimer spectra somewhat better than the latter one and slightly worse than the former potential. On the other hand, CC-pol predicts the properties of liquid water better than either of the semiempirical potentials. Thus, it appears that CC-pol is the best current across the board representation of the water force field. The ability of *ab initio* methods to compute interaction potentials allowing predictions of observables with accuracies competitive to experiments should

lead to better understanding of properties of matter. Such potentials can be developed for molecules much larger than water.[125–127]

ACKNOWLEDGMENTS

[1] M. P. Allen and D. J. Tidesley, *Computer Simulation of Liquids* (Clarendon, Oxford, 1987).
[2] H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma, J. Phys. Chem. **91**, 6269 (1987).
[3] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, J. Chem. Phys. **79**, 926 (1983).
[4] M. W. Mahoney and W. L. Jorgensen, J. Chem. Phys. **112**, 8910 (2000).
[5] J. R. Reimers, R. O. Watts, and M. L. Klein, Chem. Phys. **64**, 95 (1982).
[6] I. M. Svishchev, P. G. Kusalik, J. Wang, and R. J. Boyd, J. Chem. Phys. **105**, 4742 (1996).
[7] P. Cieplak, P. Kollman, and T. Lybrand, J. Chem. Phys. **92**, 6755 (1990).
[8] J. Caldwell, L. X. Dang, and P. A. Kollman, J. Am. Chem. Soc. **112**, 9144 (1990).
[9] R. E. Kozack and P. C. Jordan, J. Chem. Phys. **96**, 3120 (1992).
[10] H. A. Stern, F. Rittner, B. J. Berne, and R. A. Friesner, J. Chem. Phys. **115**, 2237 (2001).
[11] R. A. Kuharski and P. J. Rossky, Chem. Phys. Lett. **103**, 357 (1984).
[12] R. A. Kuharski and P. J. Rossky, J. Chem. Phys. **103**, 5164 (1985).
[13] G. S. D. Buono, P. J. Rossky, and J. Schnitker, J. Chem. Phys. **95**, 3728 (1991).
[14] B. J. Berne and D. Thirumalai, Annu. Rev. Phys. Chem. **37**, 401 (1986).
[15] M. E. Tuckerman, B. J. Berne, G. J. Martyna, and M. L. Klein, J. Chem. Phys. **97**, 2635 (1992).
[16] H. A. Stern and B. J. Berne, J. Chem. Phys. **115**, 7622 (2001).
[17] J. S. Cao and G. A. Voth, J. Chem. Phys. **100**, 5093 (1994).
[18] J. S. Cao and G. A. Voth, J. Chem. Phys. **100**, 5106 (1994).
[19] J. S. Cao and G. A. Voth, J. Chem. Phys. **101**, 6157 (1994).
[20] J. S. Cao and G. A. Voth, J. Chem. Phys. **101**, 6168 (1994).
[21] G. A. Voth, Adv. Chem. Phys. **93**, 135 (1996).
[22] J. Lobaugh and G. A. Voth, J. Chem. Phys. **106**, 2400 (1997).
[23] L. H. de la Pena and P. G. Kusalik, J. Chem. Phys. **121**, 5992 (2004).
[24] L. H. de la Pena and P. G. Kusalik, J. Am. Chem. Soc. **127**, 5246 (2005).
[25] L. H. de la Pena and P. G. Kusalik, J. Phys. Chem. A **109**, 7236 (2005).
[26] F. Paesani, S. Iuchi, and G. A. Voth, J. Chem. Phys. **127**, 074506 (2007).
[27] G. S. Fanourgakis, G. K. Schenter, and S. S. Xantheas, J. Chem. Phys. **125**, 141102 (2006).
[28] R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985).
[29] R. N. Barnett and U. Landman, Phys. Rev. B **48**, 2081 (1993).
[30] D. Xenides, B. R. Randolf, and B. M. Rode, J. Mol. Liq. **123**, 2 (2006).
[31] M. E. Tuckerman, D. Marx, M. L. Klein, and M. Parrinello, J. Chem. Phys. **104**, 5579 (1996).
[32] K. Szalewicz, R. Bukowski, and B. Jeziorski, in *Theory and Applications of Computational Chemistry: The First* 40 *Years. A Volume of Technical and Historical Perspectives*, edited by C. E. Dykstra, G. Frenking, K. S. Kim, and G. E. Scuseria (Elsevier, Amsterdam, 2005), Chap. 33, pp. 919–962.
[33] E. Zwart, J. J. ter Meulen, and W. L. Meerts, J. Mol. Spectrosc. **173**, 115 (1990).
[34] E. Zwart, J. J. ter Meulen, W. L. Meerts, and L. H. Coudert, J. Mol. Spectrosc. **147**, 27 (1991).
[35] G. T. Fraser, Int. Rev. Phys. Chem. **10**, 189 (1991).
[36] E. N. Karyakin, G. T. Fraser, and R. D. Suenram, Mol. Phys. **78**, 1179 (1993).
[37] J. B. Paul, R. A. Provencal, and R. J. Saykally, J. Phys. Chem. A **102**, 3279 (1998).
[38] L. B. Braly, J. D. Cruzan, K. Liu, R. S. Fellers, and R. J. Saykally, J. Chem. Phys. **112**, 10293 (2000).
[39] A. H. Harvey and E. W. Lemmon, J. Phys. Chem. Ref. Data **33**, 369 (2004).
[40] B. J. Smith, D. J. Swanton, J. A. Pople, and H. F. Schaefer III, J. Chem. Phys. **92**, 1240 (1990).
[41] G. S. Tschumper, M. L. Leininger, B. C. Hoffman, E. F. Valeev, H. F. Schaefer III, and M. Quack, J. Chem. Phys. **116**, 690 (2002).
[42] N. C. Handy, J. A. Pople, M. Head–Gordon, K. Raghavachari, and G. W.

[43] K. Raghavachari, J. A. Pople, E. S. Replogle, and M. Head-Gordon, J. Phys. Chem. **94**, 5579 (1990).
[44] W. Klopper and H. P. Lüthi, Mol. Phys. **96**, 559 (1999).
[45] A. Halkier, W. Klopper, T. Helgaker, P. Jørgensen, and P. R. Taylor, J. Chem. Phys. **111**, 9157 (1999).
[46] W. Klopper, J. G. C. M. van Duijneveldt–van de Rijdt, and F. B. van Duijneveldt, Phys. Chem. Chem. Phys. **2**, 2227 (2000).
[47] O. Matsuoka, E. Clementi, and M. Yoshimine, J. Chem. Phys. **64**, 1351 (1976).
[48] E. Clementi and P. Habitz, J. Phys. Chem. **87**, 2815 (1983).
[49] V. Carravetta and E. Clementi, J. Chem. Phys. **81**, 2646 (1984).
[50] U. Niessar, G. Corongiu, M.-J. Huang, M. Dupuis, and E. Clementi, Int. J. Quantum Chem., Symp. **23**, 421 (1989).
[51] U. Niessar, G. Corongiu, E. Clementi, G. R. Kneller, and D. K. Bhattacharya, J. Phys. Chem. **94**, 7949 (1990).
[52] G. Corongiu and E. Clementi, J. Chem. Phys. **97**, 2030 (1992); **97**, 8818(E) (1992).
[53] C. Millot and A. J. Stone, Mol. Phys. **77**, 439 (1992).
[54] C. Millot, J. C. Soetens, M. T. C. M. Costa, M. P. Hodges, and A. J. Stone, J. Phys. Chem. **102**, 754 (1998).
[55] I. C. Hayes and A. J. Stone, Mol. Phys. **53**, 69 (1984).
[56] I. C. Hayes and A. J. Stone, Mol. Phys. **53**, 83 (1984).
[57] B. Jeziorski, R. Moszyński, and K. Szalewicz, Chem. Rev. (Washington, D.C.) **94**, 1887 (1994).
[58] B. Jeziorski and K. Szalewicz, in *Encyclopedia of Computational Chemistry*, edited by P. von Ragué Schleyer, N. L. Allinger, T. Clark, J. Gasteiger, P. A. Kollman, H. F. Schaefer III, and P. R. Schreiner (Wiley, Chichester, 1998), Vol. 2, pp. 1376–1398.
[59] K. Szalewicz, K. Patkowski, and B. Jeziorski, Struct. Bonding (Berlin) **116**, 43 (2005).
[60] W. Rijks and P. E. S. Wormer, J. Chem. Phys. **90**, 6507 (1989).
[61] W. Rijks and P. E. S. Wormer, J. Chem. Phys. 92, 5754 (1990).
[62] M. M. Szczesniak, R. J. Brenstein, S. M. Cybulski, and S. Scheiner, J. Phys. Chem. **94**, 1781 (1990).
[63] C. J. Burnham and S. S. Xantheas, J. Chem. Phys. **116**, 1500 (2002).
[64] C. J. Burnham and S. S. Xantheas, J. Chem. Phys. **116**, 5115 (2002).
[65] G. S. Fanourgakis and S. S. Xantheas, J. Phys. Chem. A **110**, 4100 (2006).
[66] R. Bukowski, K. Szalewicz, G. C. Groenenboom, and A. van der Avoird, Science **315**, 1249 (2007).
[67] R. S. Fellers, L. B. Braly, R. J. Saykally, and C. Leforestier, J. Chem. Phys. **110**, 6306 (1999).
[68] H. Chen, S. Liu, and J. C. Light, J. Chem. Phys. **110**, 168 (1999).
[69] R. S. Fellers, C. Leforestier, L. B. Braly, M. G. Brown, and R. J. Saykally, Science **284**, 945 (1999).
[70] N. Goldman, R. S. Fellers, M. G. Brown, L. B. Braly, C. J. Keoshian, C. Leforestier, and R. J. Saykally, J. Chem. Phys. **116**, 10148 (2002).
[71] N. Goldman and R. J. Saykally, J. Chem. Phys. **120**, 4777 (2004).
[72] N. Goldman, C. Leforestier, and R. J. Saykally, Philos. Trans. R. Soc. London, Ser. A **363**, 493 (2005).
[73] E. M. Mas, K. Szalewicz, R. Bukowski, and B. Jeziorski, J. Chem. Phys. **107**, 4207 (1997).
[74] E. M. Mas, R. Bukowski, K. Szalewicz, G. C. Groenenboom, P. E. S. Wormer, and A. van der Avoird, J. Chem. Phys. **113**, 6687 (2000).
[75] G. C. Groenenboom, E. M. Mas, R. Bukowski, K. Szalewicz, P. E. S. Wormer, and A. van der Avoird, Phys. Rev. Lett. **84**, 4072 (2000).
[76] G. C. Groenenboom, P. E. S. Wormer, A. van der Avoird, E. M. Mas, R. Bukowski, and K. Szalewicz, J. Chem. Phys. **113**, 6702 (2000).
[77] M. J. Smit, G. C. Groenenboom, P. E. S. Wormer, A. van der Avoird, R. Bukowski, and K. Szalewicz, J. Phys. Chem. A **105**, 6212 (2001).
[78] M. Torheyden and G. Jansen, Mol. Phys. **104**, 2101 (2006).
[79] K. Szalewicz, G. Murdachaew, R. Bukowski, O. Akin–Ojo, and C. Leforestier, in *Lecture Series on Computer and Computational Science: ICCMSE 2006*, edited by G. Maroulis and T. Simos (Brill Academic, Leiden, 2006), Vol. 6, pp. 482–491.
[80] X. Huang, B. J. Braams, and J. M. Bowman, J. Phys. Chem. A **110**, 445 (2006).
[81] E. M. Mas, R. Bukowski, and K. Szalewicz, J. Chem. Phys. **118**, 4386 (2003).
[82] V. F. Lotrich and K. Szalewicz, J. Chem. Phys. **106**, 9668 (1997).
[83] V. F. Lotrich and K. Szalewicz, Phys. Rev. Lett. **79**, 1301 (1997).
[84] V. F. Lotrich and K. Szalewicz, J. Chem. Phys. **112**, 112 (2000).
[85] V. F. Lotrich and K. Szalewicz, J. Phys. Chem. **106**, 9688 (1997).

094313-15   Polarizable interaction potential for water. I

J. Chem. Phys. **128**, 094313 (2008)

[86] V. F. Lotrich, P. Jankowski, and K. Szalewicz, J. Chem. Phys. **108**, 4725 (1998).

[87] E. M. Mas, R. Bukowski, and K. Szalewicz, J. Chem. Phys. **118**, 4404 (2003).

[88] A. K. Soper, Chem. Phys. **258**, 121 (2000).

[89] G. Hura, J. M. Sorenson, R. M. Glaeser, and T. Head–Gordon, J. Chem. Phys. **113**, 9140 (2000).

[90] R. Bukowski, K. Szalewicz, G. C. Groenenboom, and A. van der Avoird, J. Chem. Phys. **125**, 044301 (2006).

[91] A. J. Misquitta, R. Podeszwa, B. Jeziorski, and K. Szalewicz, J. Chem. Phys. **123**, 214103 (2005).

[92] A. J. Misquitta and K. Szalewicz, J. Chem. Phys. **122**, 214109 (2005).

[93] A. J. Misquitta and K. Szalewicz, Chem. Phys. Lett. **357**, 301 (2002).

[94] R. Bukowski, K. Szalewicz, G. C. Groenenboom, and A. van der Avoird, J. Chem. Phys. **128**, 094314 (2008).

[95] M. Jeziorska, P. Jankowski, K. Szalewicz, and B. Jeziorski, J. Chem. Phys. **113**, 2957 (2000).

[96] E. M. Mas and K. Szalewicz, J. Chem. Phys. **104**, 7606 (1996).

[97] S. F. Boys and F. Bernardi, Mol. Phys. **19**, 553 (1970).

[98] F. B. van Duijneveldt, J. G. C. M. van Duijneveldt–van de Rijdt, and J. H. van Lenthe, Chem. Rev. (Washington, D.C.) **94**, 1873 (1994).

[99] T. Korona, H. L. Williams, R. Bukowski, B. Jeziorski, and K. Szalewicz, J. Chem. Phys. **106**, 5109 (1997).

[100] K. Szalewicz and B. Jeziorski, J. Chem. Phys. **104**, 1198 (1998).

[101] R. A. Kendall, T. H. Dunning, and R. J. Harrison, J. Chem. Phys. **96**, 6796 (1992).

[102] D. E. Woon and T. H. Dunning, Jr., J. Chem. Phys. **98**, 1358 (1993).

[103] H.-J. Werner, P. J. Knowles, R. Lindh, M. Schütz *et al.*, MOLPRO, *version 2002.6, a package of ab initio programs* (2003), see http://www.molpro.net.

[104] See EPAPS Document No. E-JCPSA6-128-301805 for the list of computed interaction energies, fit parameters, and a FORTRAN subroutine calculating the fit. This document can be reached through a direct link in the online article's HTML reference section or via the EPAPS homepage (http://www.aip.org/pubservs/epaps.html).

[105] J. L. Prather, *Energy Levels in Crystals*, NBS Monogr., **19** (1961).

[106] A. J. Stone, *The Theory of Intermolecular Forces* (Clarendon, Oxford, 1996).

[107] P. E. S. Wormer and H. Hettema, POLCOR *package*, University of Nijmegen, 1992.

[108] P. E. S. Wormer and H. Hettema, J. Chem. Phys. **97**, 5592 (1992).

[109] S. A. Clough, Y. Beers, G. P. Klein, and L. S. Rothman, J. Chem. Phys. **59**, 2254 (1973).

[110] J. Verhoeven and A. Dymanus, J. Chem. Phys. **52**, 3222 (1970).

[111] K. T. Tang and J. P. Toennies, J. Chem. Phys. **80**, 3726 (1984).

[112] W. F. Murphy, J. Chem. Phys. **67**, 5877 (1977).

[113] G. Maroulis, Chem. Phys. Lett. **289**, 403 (1998).

[114] G. Avila, J. Chem. Phys. **122**, 144310 (2005).

[115] A. Russell and M. Spackman, Mol. Phys. **84**, 1239 (1995).

[116] C. Cuthbertson and M. Cuthbertson, Philos. Trans. R. Soc. London, Ser. A **213**, 1 (1913).

[117] K. B. Newbound, J. Opt. Soc. Am. **39**, 835 (1949).

[118] A. van der Avoird, P. E. S. Wormer, F. Mulder, and R. M. Berns, Top. Curr. Chem. **93**, 1 (1980).

[119] C. J. Cerjan and W. H. Miller, J. Chem. Phys. **75**, 2800 (1981).

[120] J. A. Odutola and T. R. Dyke, J. Chem. Phys. **72**, 5062 (1980).

[121] F. N. Keutsch, N. Goldman, H. A. Harker, C. Leforestier, and R. J. Saykally, Mol. Phys. **101**, 3477 (2003).

[122] C. J. Burnham, J. Li, S. S. Xantheas, and M. Leslie, J. Chem. Phys. **110**, 4566 (1999).

[123] C. J. Burnham and S. S. Xantheas, J. Chem. Phys. **116**, 1479 (2002).

[124] K. Patkowski, R. Podeszwa, and K. Szalewicz, J. Phys. Chem. A **111**, 12822 (2007).

[125] R. Bukowski, K. Szalewicz, and C. Chabalowski, J. Phys. Chem. A **103**, 7322 (1999).

[126] R. Podeszwa, R. Bukowski, and K. Szalewicz, J. Phys. Chem. A **110**, 10345 (2006).

[127] R. Podeszwa, R. Bukowski, B. M. Rice, and K. Szalewicz, Phys. Chem. Chem. Phys. **9**, 5561 (2007).