

# The water-oxygen dimer: First-principles calculation of an extrapolated potential energy surface and second virial coefficients

Richard J. Wheatley<sup>a)</sup>

*School of Chemistry, The University of Nottingham, Nottingham NG7 2RD, United Kingdom*

Allan H. Harvey<sup>b)</sup>

*Physical and Chemical Properties Division, National Institute of Standards and Technology, Boulder, Colorado 80305*

(Received 5 April 2007; accepted 15 June 2007; published online 20 August 2007)

The systematic intermolecular potential extrapolation routine (SIMPER) is applied to the water-oxygen complex to obtain a five-dimensional potential energy surface. This is the first application of SIMPER to open-shell molecules, and it is the first use, in this context, of asymptotic dispersion energy coefficients calculated using the unrestricted time-dependent coupled-cluster method. The potential energy surface is extrapolated to the complete basis set limit, fitted as a function of intermolecular geometry, and used to calculate (mixed) second virial coefficients, which significantly extend the range of the available experimental data. © 2007 American Institute of Physics. [DOI: 10.1063/1.2756524]

## I. INTRODUCTION

The water-oxygen interaction is important in chemistry and biochemistry, in combustion processes, and in understanding the properties of humid air for metrology. The absorption of light by oxygen gas is affected by collisions between oxygen and water molecules, and this may have a significant effect on transmission of light through the earth's atmosphere.<sup>1</sup> To predict these properties of water-oxygen mixtures, particularly the thermodynamic properties that are the main focus of this work, a potential energy surface is needed, which describes the intermolecular interactions for all coordinates over a range of energies from the potential minimum at least up to dissociation of the complex. This paper describes the first calculation and analytical fit of a five-dimensional water-oxygen potential energy surface (assuming rigid monomers) and subsequent calculations of second virial coefficients for the water-oxygen system, which complete our study of the nonideal gas properties of mixtures of water with the main components of air.<sup>2,3</sup>

The weakly bound nature of the water-oxygen complex and the open-shell electronic structure make theoretical calculations of the intermolecular potential difficult. In fact, theoretical studies up to 1998 (Ref. 4) did not even include the planar, van der Waals bonded intermolecular geometry that is now believed to be the potential minimum.<sup>5-7</sup> This was first studied in 2002 by Kjaergaard *et al.*<sup>8</sup> using quadratic configuration interaction with single and double substitution calculations. However, several different equilibrium geometries were proposed in the following three years, including an out-of-plane geometry<sup>9</sup> and hydrogen-bonded geometries.<sup>10,11</sup> The limited amount of spectroscopic data on the water-oxygen dimer<sup>9,12</sup> has not yet enabled the equilibrium geometry to be established experimentally.

The water-oxygen intermolecular potential energy is calculated in this paper at the unrestricted second-order Møller-Plesset (UMP2) level, with extrapolation to the complete basis set limit. A correction for the approximate treatment of electron correlation is applied using the systematic intermolecular potential extrapolation routine (SIMPER),<sup>13</sup> and the resulting potential energy surface is tested for accuracy using restricted coupled cluster with single, double, and perturbative triple excitations [RCCSD(T)] calculations. The minimum-energy structure essentially agrees with that of Kjaergaard *et al.* Other previously reported minima are found to have higher energies, but there is a local hydrogen-bonded minimum in addition to the global minimum. The methods used to calculate the potential energy surface are described in Sec. II, and the results and second virial coefficients are presented in Sec. III and IV.

## II. METHODOLOGY

The intramolecular water and oxygen geometries are fixed at their ground-state expectation values (O–O bond length of  $2.288a_0$ ,<sup>14</sup> O–H bond length of  $1.8361a_0$ , and HOH bond angle of  $104.69^\circ$ ,<sup>15</sup> the Bohr radius is  $a_0 \approx 5.291\,772 \times 10^{-11}$  m). The five intermolecular coordinates are analogous to those previously used for the water-hydrogen<sup>16</sup> and water-nitrogen<sup>3</sup> complexes. The distance  $R$  and the spherical polar angles  $\theta$  and  $\phi$  define the position of the center of the O<sub>2</sub> molecule in a coordinate system with its origin at the water O atom,  $z$  axis from the origin directly away from the midpoint between the two H atoms, and  $x$  axis from one H atom to the other. The spherical polar angles  $\theta'$  and  $\phi'$  define the orientation of the O<sub>2</sub> molecule relative to the same axes.

A total of 3465 separate relative orientations ( $\theta, \phi, \theta', \phi'$ ) are used in the calculations. These consist of three randomly chosen sets of 1000 orientations each, which are used only for test calculations with small basis sets, and

<sup>a)</sup>Electronic mail: richard.wheatley@nottingham.ac.uk

<sup>b)</sup>Electronic mail: aharvey@boulder.nist.gov

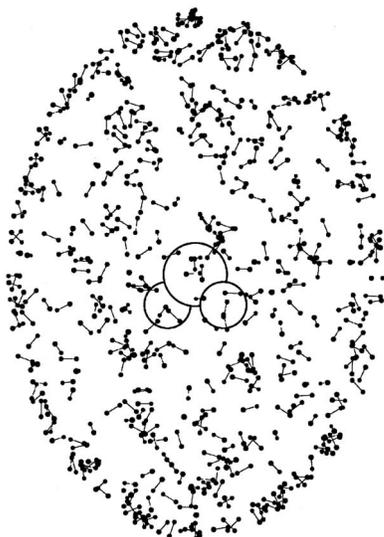


FIG. 1. Plane projection of the 465 angular geometries used in the calculation of the potential energy surface. The water molecule is shown as large open circles, and its plane is oriented at  $45^\circ$  to the plane of the paper. Oxygen molecules are placed on the surface of a quarter sphere, centered on the oxygen atom of the water molecule. The oxygen molecules on the right-hand edge of the diagram are centered in the plane of the water molecule, and those on the left-hand edge are equidistant from the hydrogen atoms. The bond length of the oxygen molecules is reduced by a factor of 5 for clarity. Oxygen molecules that are oriented at an angle to the plane of the paper have their upper atom shown larger.

one set of 465 orientations, which is used in the calculation of the final potential energy surface. (The number 465 is limited by available computing power.)

The set of 465 orientations is obtained as follows. Let  $\Omega_i$ ,  $i=1-465$ , denote the set of orientations, and let  $F_j(\Omega)$ ,  $j=1-465$ , denote a set of orthogonal angular functions, which are chosen here to be products of spherical harmonics  $Y_{lm}(\theta, \phi)Y_{l'm'}(\theta', \phi')$ , with  $l \leq 6$ ,  $l' \leq 6$ , and  $l+l' \leq 10$ . Define a matrix  $M$  such that  $M_{ij}=F_j(\Omega_i)$ . The coefficients  $c_j$  in an angular expansion of the energy can be obtained using  $c=M^{-1}E$ , where  $E$  is the vector of calculated energies,  $E_i=E(\Omega_i)$ . To make the matrix  $M$  as far from linear dependence as possible, the sum of squares of the elements of its inverse  $M^{-1}$  is minimized with respect to variations in the chosen orientations. (This is equivalent to minimizing  $\sum_i \lambda_i^{-2}$ , where  $\lambda_i$  are the eigenvalues of  $M$ .) This is done by repeatedly making random small changes to randomly chosen orientations and keeping the change if the sum of squares of  $M^{-1}$  decreases. When applied to smaller problems, for which the results can be explicitly checked, this procedure appears to be reasonable. In two dimensions  $(\theta, \phi)$ , with four angular functions  $Y_{lm}$ ,  $l \leq 1$ , the four resulting points are found to be at the vertices of a regular tetrahedron, and the nine points produced from the nine functions  $Y_{lm}$ ,  $l \leq 2$  form a regular tricapped trigonal prism. The 465 orientations used for the water-oxygen dimer are shown schematically in Fig. 1. Ten different water-oxygen distances  $R$  are used in the calculations:  $R/a_0=4.5, 5, 5.5, 6, 6.5, 7, 7.5, 8, 9, 10$ , so the final potential energy surface is calculated at 4650 separate points.

All calculations use the “SP” modified versions<sup>3,17</sup> of the Dunning basis sets,<sup>18</sup> where SP means “shifted polarization

functions.” For example, the SP-aug-cc-pVTZ basis set (abbreviated as SPTZ in what follows) is obtained from the aug-cc-pVTZ basis set by changing the exponents of the two oxygen  $f$  Gaussians to equal the most diffuse (smallest) exponents of the oxygen  $d$  Gaussians. The exponents of the  $d$  Gaussians on hydrogen are changed to be the same as the most diffuse hydrogen  $p$  Gaussians. Similarly, the SPQZ basis set is based on the aug-cc-pVQZ basis set, with the  $f$  and  $g$  Gaussian exponents on oxygen changed to equal the most diffuse  $d$  exponents, and the  $d$  and  $f$  Gaussian exponents on hydrogen changed to equal the most diffuse  $p$  exponents. The “SPDZ” basis set is identical to the aug-cc-pVDZ basis set, because there are no  $f$  polarization functions on O or H.

The oxygen and water molecules interact through weak van der Waals forces, so it is important to include nonlocal electron correlation in supermolecule calculations. The methods used to achieve this are MP2 perturbation theory, coupled-cluster theory with single and double excitations (CCSD), CCSD with perturbative triple excitations [CCSD(T)], and the SIMPER method.<sup>13</sup> Both restricted and unrestricted electron spins are considered. Supermolecule calculations are performed using the MOLPRO program,<sup>19,20</sup> apart from unrestricted CCSD (UCCSD), which uses the SIMPER program, and some UMP2 results are checked for consistency between the two programs. In the unrestricted-spin calculations, the expectation value of the total spin quantum number  $S^2$  is 2.048 or 2.049 for the oxygen monomer and for the water-oxygen dimer, so spin contamination of the wave function is small, and the use of unrestricted spins is justified. In preliminary studies, calculating an entire CCSD(T) potential energy surface was not feasible (it took over 8 h/point with the SPTZ basis set, and SPQZ calculations were not successful), and the CCSD method is found to be quite inaccurate for supermolecule calculations (see below), so the UMP2 method is chosen to produce the initial potential energy surface.

Although Møller-Plesset calculations usually offer reasonable accuracy for relatively low effort, they do not always produce reliable van der Waals energies; even for closed-shell atoms they can give errors of around 40% in the well depth.<sup>21</sup> Since CCSD(T) calculations are too demanding, the SIMPER method is used to improve the treatment of electron correlation.

The SIMPER method involves dividing the UMP2 supermolecule interaction energy  $E_{\text{int}}$  into a first-order Coulomb (electrostatic) interaction energy  $E_{\text{elec}}$ , a second-order induction interaction energy  $E_{\text{ind}}$ , a second-order dispersion interaction energy  $E_{\text{disp}}$ , and an “exchange-repulsion interaction energy”  $E_{\text{rep}}$  at each point on the potential energy surface,

$$E_{\text{int}}(\text{UMP2}) = E_{\text{elec}}(\text{UMP2}) + E_{\text{ind}}(\text{UMP2}) + E_{\text{disp}}(\text{UMP2}) + E_{\text{rep}}(\text{UMP2}). \quad (1)$$

The first three of these contributions are well-defined parts of the UMP2 supermolecule interaction energy. They depend on the molecular charge densities and polarization propagators, both of which can be calculated as the sum of the corre-

sponding unrestricted Hartree-Fock property plus the UMP2 correlation correction. The first-order Coulomb energy is the interaction between the unperturbed charge densities of the two molecules (excluding the interaction between the two UMP2 corrections), the induction energy is the interaction between the square of the charge density of water and the polarization propagator of oxygen and vice versa (excluding contributions involving products of two or three UMP2 corrections), and the dispersion energy is calculated using the uncoupled Hartree-Fock frequency-dependent polarization propagators (with no UMP2 correction). The exchange-repulsion energy is the difference between the sum of these three contributions and the supermolecule interaction energy  $E_{\text{int}}$ .

The different contributions to the supermolecule energy are then recalculated, or scaled, using a higher level of electron correlation. The UMP2 electrostatic interaction energy is simply replaced by the electrostatic interaction energy between monomer UCCSD charge densities. The UMP2 dispersion energy is expressed as a sum of multipolar contributions  $C_n R^{-n}$ ,  $n \geq 6$ , multiplied by damping functions  $f_n(bR)$ , where  $C_n$  is an orientation-dependent dispersion energy coefficient and  $R$  is the separation between the oxygen atom of water and the bond center of the  $O_2$  molecule. The damping length scale parameter  $b$  is determined uniquely at each point on the surface from the calculated UMP2 dispersion energy and the UMP2 dispersion energy coefficients. The UMP2 dispersion energy coefficients are replaced by unrestricted time-dependent coupled-cluster (TD-UCCSD) dispersion energy coefficients,<sup>22</sup> and the parameter  $b$  is multiplied at each point by the ratio of the TD-UCCSD to UMP2 values of the quantity  $\sqrt{C_6/C_8}$  (Ref. 23) at that orientation. This defines the SIMPER dispersion energy. The induction energy is a relatively small part of the total interaction energy and is not changed (but see below). Finally, the UMP2 exchange-repulsion energy is multiplied by the ratio of the UCCSD to UMP2 values of the quantity  $S_\rho$ , which is the overlap integral between the two monomer electron densities, to give the SIMPER exchange-repulsion energy. The total SIMPER interaction energy is then the sum of the UCCSD electrostatic energy, the UMP2 induction energy, and the SIMPER dispersion and exchange-repulsion energies,

$$E_{\text{int}}(\text{SIMPER}) = E_{\text{elec}}(\text{UCCSD}) + E_{\text{ind}}(\text{UMP2}) + E_{\text{disp}}(\text{SIMPER}) + E_{\text{rep}}(\text{UMP2}) \frac{S_\rho(\text{UCCSD})}{S_\rho(\text{UMP2})}. \quad (2)$$

This work is the first use of SIMPER for an open-shell dimer, and it is the first time that it has been implemented with time-dependent coupled-cluster dispersion energy coefficients. Note that the computational requirements per point on the potential energy surface are essentially the same for SIMPER and UMP2. The SIMPER method requires additional calculations of the UCCSD charge densities and TD-UCCSD frequency-dependent polarizabilities, but these calculations are only done once per monomer, not once per point on the surface.

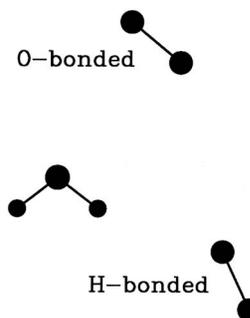


FIG. 2. The planar O-bonded and H-bonded equilibrium geometries of the water-oxygen dimer. The O-bonded dimer consists of the water molecule and the upper oxygen molecule. The H-bonded dimer consists of the water molecule and the lower oxygen molecule.

### III. RESULTS

Using UMP2/SPQZ supermolecule calculations, the equilibrium geometry of the water-oxygen dimer having the lowest energy is found to be planar, with the  $O_2$  molecule located toward the oxygen end of the water molecule and oriented roughly parallel to the nearest OH bond (“O bonded”). In terms of the coordinates given in Sec. II,  $R \approx 6.0a_0$ ,  $\theta \approx 36^\circ$ , and  $\theta' \approx 130^\circ$ . A second planar minimum is also found, with the oxygen molecule “H bonded” to a hydrogen atom of water. The coordinates of this structure are  $R \approx 7.4a_0$ ,  $\theta \approx 120^\circ$ , and  $\theta' \approx 155^\circ$ . These two structures are shown in Fig. 2; the lowest-energy pathway from the O-bonded to the H-bonded structure involves clockwise rotation of the oxygen molecule on the figure by more than  $180^\circ$ . The symmetry of the water and oxygen molecules means that there are four equivalent versions of each equilibrium structure. In general, the water-oxygen complex shows more “van der Waals” character than the water-nitrogen complex: water-nitrogen is more strongly bound at the equilibrium geometry, and its potential energy surface is more anisotropic, as a result of the larger nitrogen quadrupole. Some water-nitrogen orientations are repulsive for all intermolecular separations, which is not the case for water-oxygen. Furthermore, enlarging the basis set increases the water-oxygen binding energy for most of the geometries considered, including all the geometries with  $R < 8a_0$ ; this is also typical of a van der Waals complex.

The UMP2 and SIMPER interaction energies are shown in Table I for two geometries close to the O-bonded and H-bonded equilibrium structures described above. The binding energy changes by about 60% from the SPDZ to SPTZ basis set but only by about 4% from SPTZ to SPQZ. Basis set extrapolation using the two-point formula of Martin<sup>24</sup> increases the SPQZ binding energy by a further 2% or so.

The error associated with the incompleteness of the SPQZ basis set and/or with basis set extrapolation is almost certainly less than the error associated with the treatment of electron correlation. Results obtained with the SPTZ basis set, using the MP2, CCSD, and CCSD(T) methods, and with restricted and unrestricted electron spins are also shown in Table I. For the O-bonded global minimum, the difference between the largest and smallest energies is 29%. Even the RMP2 and UMP2 results differ by 20%. The SIMPER bind-

TABLE I. Intermolecular potential (in  $\text{cm}^{-1}$ ,  $1 \text{ cm}^{-1} \approx 1.98643 \times 10^{-23} \text{ J}$ ) near the O-bonded global minimum and H-bonded local minimum of the  $\text{H}_2\text{O}-\text{O}_2$  complex, obtained with different theoretical methods and basis sets. R and U denote restricted and unrestricted spins, respectively. CBS denotes extrapolation of SPTZ and SPQZ results to the complete basis set limit.

	O-bonded				H-bonded			
	SPDZ	SPTZ	SPQZ	CBS	SPDZ	SPTZ	SPQZ	CBS
UMP2	-139	-200	-207	-211	-73	-121	-126	-128
SIMPER	-110	-180	-184	-186	-77	-137	-146	-152
SPTZ	MP2	CCSD	CCSD(T)	SIMPER	MP2	CCSD	CCSD(T)	SIMPER
R spin	-167	-155	-192		-97	-86	-123	
U spin	-200	-155		-180	-121	-87		-137

ing energy is around the middle of the range. It is 10% shallower than the UMP2 potential, which is used as the starting point for the SIMPER method, and is 6% shallower than the RCCSD(T) potential, which is probably the most reliable supermolecule result in the table. Although the CCSD method usually gives reliable monomer properties (such as those used in the SIMPER method), it does not describe this weak intermolecular bond particularly well; it is considerably more computationally expensive, and apparently less accurate, than the UMP2 method. The agreement between the different methods improves rapidly as the intermolecular separation increases, as shown in Fig. 3.

For the H-bonded local minimum, the difference between the largest and smallest energies is 59%, or 42% for the supermolecule methods (excluding SIMPER). As with the O-bonded geometry, the RCCSD, UCCSD, and restricted MP2 (RMP2) methods give significantly smaller binding energies than the UMP2, RCCSD(T), and SIMPER methods. The difference between the RMP2 and UMP2 binding energies is 25% for the H-bonded geometry, whereas the RCCSD and UCCSD binding energies are again very similar, presumably because the single excitations in the CCSD cluster operator can interconvert the restricted and unrestricted SCF

orbitals. The SIMPER binding energy for the H-bonded geometry is 11% more than the RCCSD(T) binding energy.

Table II shows the SIMPER corrections to the separate components of the energy (electrostatic, induction, dispersion, and exchange repulsion). In the O-bonded geometry, the electrostatic and dispersion energies both become less negative when the relevant MP2 quantities are replaced by their CCSD counterparts. In the H-bonded geometry, where the RCCSD(T) result suggests that the SIMPER method may be overestimating the binding energy, the dispersion energy becomes more negative, and this roughly cancels the change in the electrostatic energy, but the largest effect is a reduction in exchange-repulsion energy arising from the reduced overlap of the CCSD ground-state electron densities relative to the MP2 densities. The effect of this difference between RCCSD(T) and SIMPER calculations on the calculated thermodynamic properties is considered in more detail in Sec. IV.

To obtain a complete potential energy surface, the calculated points are fitted to a function of the coordinates. First, the original choice of 465 angular points for the potential energy surface is investigated (see Sec. II). Supermolecule calculations are performed at MP2/SPDZ level for the 465 points plus three sets of 1000 randomly chosen angular points, with an intermolecular separation of  $5a_0$ . This intermolecular separation is chosen to sample the repulsive wall, where fitting is most difficult; the calculated interaction energies range from 150 to 20 000  $\text{cm}^{-1}$ . The results are fitted to a set of 195 angular functions, with  $l \leq 6$  for the water molecule and  $l \leq 4$  for  $\text{O}_2$ . When the original 465 points are fitted, and the resulting (unweighted) fit is compared to the 3465 calculated energies at the same separation, the rms error is 277  $\text{cm}^{-1}$ . Fitting the three sets of 1000 randomly chosen points gives rms errors, compared to the complete set of 3465 points, of 281, 292, and 270  $\text{cm}^{-1}$ . The best possible rms error with this set of angular functions, obtained by fitting all 3465 points, is 250  $\text{cm}^{-1}$ . Similar results are obtained for the other intermolecular separations, and it is concluded that the chosen set of 465 angular points gives a good representation of the angular behavior of the potential, and that the method used to select these points is about twice as efficient as random selection, in terms of the number of points

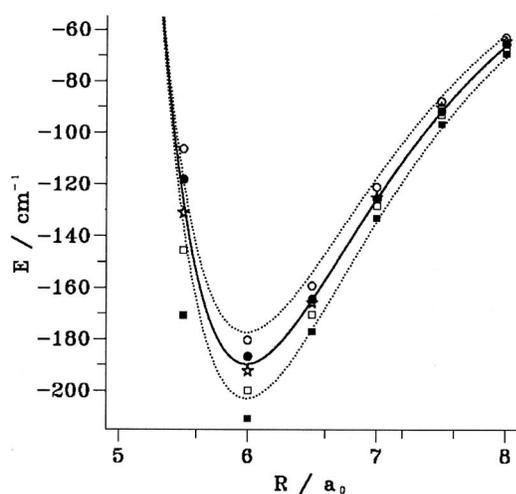


FIG. 3. The radial dependence of the water-oxygen intermolecular potential at the O-bonded geometry. Open squares, UMP2/SPTZ; open circles, SIMPER/SPTZ; stars, RCCSD(T)/SPTZ; filled squares, UMP2/CBS; filled circles, SIMPER/CBS; solid line, fit to SIMPER/CBS; and dotted lines, 7% confidence limits.

TABLE II. Contributions to the intermolecular potential ( $\text{cm}^{-1}$ ) near the O-bonded global minimum and H-bonded local minimum of the  $\text{H}_2\text{O}-\text{O}_2$  complex, obtained with different basis sets and extrapolation to the complete basis set limit (CBS). The first number is the UMP2 value, and the number in parentheses is the SIMPER correction to it.

	O-bonded			
	SPDZ	SPTZ	SPQZ	CBS
$E_{\text{elec}}$	-160(+8)	-147(+11)	-151(+13)	-154(+14)
$E_{\text{ind}}$	-32(0)	-35(0)	-43(0)	-49(0)
$E_{\text{disp}}$	-218(+18)	-274(+11)	-291(+14)	-301(+16)
$E_{\text{rep}}$	271(+3)	257(-3)	280(-5)	293(-6)
			H-bonded	
$E_{\text{elec}}$	-69(+12)	-92(+13)	-89(+14)	-87(+14)
$E_{\text{ind}}$	-56(0)	-80(0)	-90(0)	-95(0)
$E_{\text{disp}}$	-215(-1)	-249(-8)	-258(-12)	-263(-14)
$E_{\text{rep}}$	267(-15)	299(-20)	311(-23)	317(-24)

required to obtain a representation of equal quality of the entire surface. The fitted potential in the O-bonded geometry is shown in Fig. 3.

For the complete potential energy surface, a number of analytical fits have been investigated, and the most economical in terms of minimizing rms error and number of parameters is found to be a scheme that uses fitting functions expressed in terms of the  $\text{O} \rightarrow B$ ,  $\text{H} \rightarrow B$ , and  $\text{H}' \rightarrow B$  vectors, where O, H, and H' are atoms of the water molecule and B is the bond center of the oxygen molecule. A local coordinate system is used on each atom. The local  $x$  axis of the water oxygen is parallel to the  $\text{H} \rightarrow \text{H}'$  vector, and the local  $z$  axis points along the symmetry axis away from the hydrogen atoms. The local  $z$  axis of each hydrogen atom points away from the water oxygen, and the local  $x$  axis is in the molecular plane and has a component away from the other hydrogen atom. The local  $z$  axis of  $\text{O}_2$  points along the bond. The fitting functions are chosen to be the product of inverse powers of distance and  $S$  functions,<sup>25</sup>

$$E_{\text{fit}} = \sum_{i=\text{O,H}} \sum_{n=3}^6 \sum_{l_1, l_2, l, k_1, k_2} C_{i,n,l_1, l_2, l, k_1, k_2} R_{i,B}^{-2n} S_{l_1, l_2, l}^{k_1, k_2}(\Omega_i), \quad (3)$$

where  $R_{i,B}$  is the intersite distance and  $\Omega_i$  is the relative orientation of the local axes and the intersite vector. The maximum angular momenta are  $l=2$  for O, H, and H' and 6 for B, giving 316 independent parameters in the fit. The weighting used in the fit for most of the calculated energies  $E$  is  $w = \exp(-E/E^*)$ , where  $E^* = 658 \text{ cm}^{-1}$ , but this is found to give too little weighting to points high on the repulsive wall, resulting in unacceptably large discrepancies in this region, so for calculated energies  $E > E^*$  the larger of  $w$  and  $E^*/(3E)$  is used. The sum of the weights of all 4650 calculated points is usually about 2150 times greater than the largest individual weight, and the weighted rms error is less than  $3 \text{ cm}^{-1}$ . The parameters for the fitted potential-energy function and a FORTRAN subroutine to evaluate it are available as supplementary material via EPAPS.<sup>26</sup>

## IV. SECOND VIRIAL COEFFICIENTS

### A. Calculation

The cross second virial coefficient  $B_{12}$  for the water-oxygen mixture is calculated from the potential function fitted to the complete basis set extrapolation of the SIMPER results. The calculation procedure, which has been described previously,<sup>16</sup> includes translational and rotational quantum effects to first order; higher-order quantum effects are expected to be negligible above 100 K.<sup>16</sup> The calculated second virial coefficient is fitted as a function of temperature,

$$B_{12}(T) = \sum_{i=1}^4 c_i (T^*)^{d_i}, \quad (4)$$

where  $T^* = T/(100 \text{ K})$ ,  $B_{12}$  and the  $c_i$  have units of  $\text{cm}^3 \text{ mol}^{-1}$ , and the values of  $c_i$  and  $d_i$  are given in Table III. Equation (4) is valid for temperatures from 100 to 3000 K. Table IV shows calculated values of  $B_{12}$  at different temperatures.

### B. Estimate of uncertainty

The significant dependence of the intermolecular potential on the methodology, as shown in Sec. III, produces some uncertainty in the calculated second virial coefficients. To investigate this further, with the aim of estimating reasonable confidence limits, the second virial coefficients are also calculated using a fit to the calculated UMP2 potential energy surface, extrapolated to the complete basis set limit. Overall, the SIMPER potential energy surface produces somewhat more negative virial coefficients; it is necessary to make the

TABLE III. Coefficients for Eq. (4) for  $B_{12}(T)$  for the  $\text{H}_2\text{O}/\text{O}_2$  pair. The  $c_i$  are in  $\text{cm}^3 \text{ mol}^{-1}$  and the  $d_i$  are dimensionless.

$i$	$c_i$	$d_i$
1	124.605	-0.33
2	-214.421	-0.73
3	-102.818	-2.03
4	-22.360	-4.07

SIMPER potential about 7%–8% shallower (depending on the temperature) to give agreement with the UMP2 calculations. It is, therefore, important to determine whether this overall deepening of the potential by the SIMPER procedure is realistic or just an artifact of the method. To this end, 17 sets of  $(\theta, \phi, \theta', \phi')$  orientations are chosen in a similar way to the 465 orientations used for the main calculation, and UMP2, RCCSD(T), and SIMPER calculations are performed at these orientations. Intermolecular separations  $R=6.5a_0$  and  $7a_0$  are chosen because they have the most negative average energies, and they are the separations where the difference between the UMP2 and SIMPER potential energy surfaces has the greatest effect on the virial coefficients. The SPTZ basis set is used for these calculations. The average interaction energies for these 17 points (in  $\text{cm}^{-1}$ ) at  $R=6.5a_0$  are  $-14.3$  (UMP2),  $-30.7$  [RCCSD(T)], and  $-37.0$  (SIMPER). At  $R=7a_0$  they are  $-83.6$  (UMP2),  $-95.2$  [RCCSD(T)], and  $-96.4$  (SIMPER). Similar calculations on the repulsive wall at  $R=5.5a_0$  give average energies of  $+1245$  (UMP2),  $+1224$  [RCCSD(T)], and  $+1189$  (SIMPER). Based on these results, the deepening of the potential well by SIMPER appears to be qualitatively correct, although perhaps a little exaggerated.

The main difference between the SIMPER and RCCSD(T) calculations at these 17 orientations occurs near the H-bonded geometry shown in Sec. III, where the SIMPER well depth is greater. The reason for the discrepancy at this geometry is not clear. One possibility may be the treatment of exchange-repulsion by SIMPER; the SIMPER exchange-repulsion energy is based on the semiempirical assumption that the exchange repulsion is proportional to the electron density overlap, and it is probably the least well founded of the different SIMPER contributions. However, the analogous H-bonded geometry of  $\text{H}_2\text{O}-\text{N}_2$  showed very good agreement between SIMPER and CCSD(T),<sup>3</sup> and there is no obvious reason why  $\text{H}_2\text{O}-\text{O}_2$  should be different. Another possibility may be the use of the unmodified UMP2 induction energy in the SIMPER method. Two methods for

TABLE IV. Second virial coefficients  $B_{12}$  calculated using Eq. (4) and their estimated uncertainties  $\Delta B_{12}$ .

$T$ (K)	$B_{12}$ ( $\text{cm}^3 \text{mol}^{-1}$ )	$\Delta B_{12}$ ( $\text{cm}^3 \text{mol}^{-1}$ )
100	-214.99	29.53
150	-99.92	13.75
200	-56.65	8.74
250	-34.29	6.40
300	-20.75	5.06
350	-11.73	4.20
400	-5.32	3.61
450	-0.57	3.17
500	3.09	2.83
600	8.29	2.36
700	11.77	2.03
800	14.23	1.80
900	16.03	1.63
1000	17.39	1.49
1500	20.86	1.10
2000	22.06	0.91

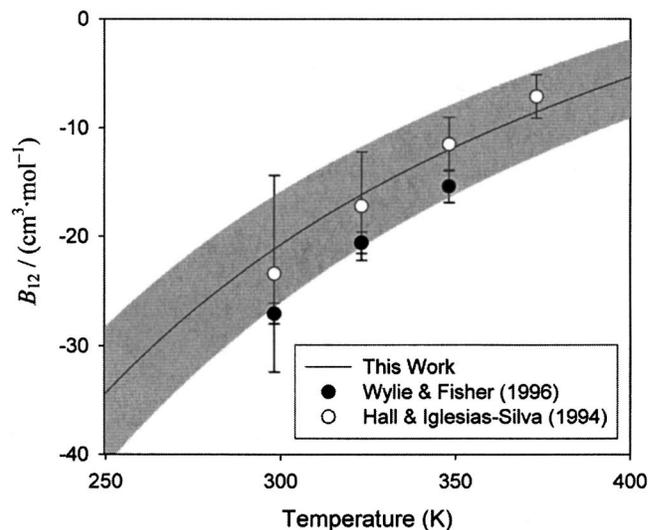


FIG. 4. Second virial coefficients for the water/oxygen mixture from the literature (shown as points with error bars), calculated second virial coefficients (full curve), and estimated uncertainties in the calculations (shaded area).

modifying the induction energy are therefore considered. Firstly, the UCCSD long-range multipolar induction energy coefficients are used instead of their UMP2 counterparts, in a similar fashion to the dispersion energy, and secondly, the UMP2 induction energy is replaced by the UMP2 response of oxygen to the UCCSD charge density of water and vice versa. However, neither modification affects the induction energy enough to account for the discrepancy with the RCCSD(T) energy. Of course, it should also be noted that the RCCSD(T) calculations themselves have errors, arising from the incomplete treatment of electron correlation.

Based on the foregoing discussion, a pessimistic estimate of the uncertainty in the virial coefficients is obtained by multiplying the negative interaction energies by 1.07 and dividing the positive interaction energies by 1.07 to give a lower bound on the virial coefficients, and vice versa, to give an upper bound. Table IV shows the uncertainties obtained in this manner at different temperatures.

### C. Comparison with experimental data

The only experimental  $B_{12}$  (mixed second virial coefficient) data for the  $\text{H}_2\text{O}/\text{O}_2$  mixture come from the study of Wylie and Fisher,<sup>27</sup> who measured the solubility of liquid water in gaseous oxygen at approximately 298, 323, and 348 K. Their reported virial coefficients are used directly for comparison with the results calculated in this paper; recalculation of  $B_{12}$  from the original solubility data of Wylie and Fisher with the procedures used previously<sup>2</sup> yields slightly different values, but the differences are well within their reported uncertainties for  $B_{12}$ . Hall and Iglesias-Silva<sup>28</sup> indirectly estimated  $B_{12}$  for the  $\text{H}_2\text{O}/\text{O}_2$  pair based on a correlation of  $B_{12}$  for water with air<sup>29</sup> and on literature  $B_{12}$  data for water with nitrogen.<sup>30</sup> Their results are somewhat less negative than those of Wylie and Fisher.<sup>27</sup>

In Fig. 4, the calculated values of  $B_{12}$  from this work are shown, along with results from the literature. The shaded area represents the uncertainty in the calculated results, as

discussed above. This uncertainty, and those of the literature points, may be taken as standard uncertainties with coverage factor  $k=2$  (approximately a 95% confidence interval). The calculated values for  $B_{12}(T)$  are systematically less negative than the experimental data of Wylie and Fisher,<sup>27</sup> although the uncertainties are large enough that the results cannot definitively be said to be inconsistent. The gas saturation technique employed by Wylie and Fisher requires lengthy and laborious experimental runs, so it is not unthinkable that additional systematic errors could be present in their data. Definitive resolution of the apparent discrepancy between these data and the present calculations would require either confirming experiments or a reduction in the uncertainty of the pair potential. On the other hand, the calculations agree well with the values obtained indirectly from water-air and water-nitrogen data by Hall and Iglesias-Silva.<sup>28</sup> This agreement is consistent with recent work by Harvey and Huang,<sup>31</sup> which shows that combining the water-oxygen results reported here with previous calculations for water-argon<sup>2</sup> and water-nitrogen<sup>3</sup> produces excellent agreement with available high-quality data for  $B_{12}$  for water with air.

Information about the function  $B_{12}(T)$  may also be derived from vapor-phase enthalpy-of-mixing data, which, when extrapolated to low pressure and combined with  $B(T)$  for the pure components, yield the function  $\phi_{12}=B_{12}-TdB_{12}/dT$ . While no such data have been published for the  $H_2O/O_2$  mixture, unpublished measurements by Wormald<sup>32</sup> over the temperature range from 383 to 493 K show that at atmospheric pressure the enthalpy of mixing for an equimolar mixture is approximately  $(1\pm 0.5)$  J mol<sup>-1</sup> more positive than it is for the  $H_2O/N_2$  mixture.

The quantity  $\phi_{12}=B_{12}-TdB_{12}/dT$  can also be calculated from Eq. (4). When this is combined with pure-component values of  $\phi$  for  $H_2O$  (Ref. 33) and  $O_2$ ,<sup>34</sup> the low-pressure limit of the quantity  $H^E/p$  for a mixture of  $H_2O$  and  $O_2$  is obtained,<sup>2</sup> where  $H^E$  is the excess enthalpy of mixing and  $p$  is the pressure. The resulting values of  $H^E$  for an equimolar  $H_2O/O_2$  mixture at atmospheric pressure between 383 and 493 K are more positive than those obtained for  $H_2O/N_2$  using a calculated  $H_2O/N_2$   $B_{12}(T)$  function<sup>3</sup> by amounts ranging from 1.4 J mol<sup>-1</sup> at the low end of this temperature range to 1.0 J mol<sup>-1</sup> at the high end. This result is in good agreement with the measurements done by Wormald.<sup>32</sup>

Estimation of the uncertainties in the calculated values of  $\phi_{12}$  is not trivial, and in previous works on other systems<sup>2,3,16</sup> these uncertainties have probably been underestimated. A more rigorous treatment of the uncertainty in  $\phi_{12}$  for water with oxygen and with the other components of air is provided by Harvey and Huang.<sup>31</sup>

## V. DISCUSSION

A new potential energy surface has been calculated for the interaction of rigid water and oxygen molecules and fitted to an analytical five-dimensional function. Water and oxygen molecules attract one another primarily by dispersion forces, but their electrostatic interaction is also important. The potential well is about half as deep as the water-nitrogen potential, and nearly twice as deep as the argon-oxygen

potential,<sup>35</sup> which reflects the differing influence of electrostatic interactions on the binding energies of these dimers. Two symmetry-distinct minima are found on the water-oxygen potential energy surface.

Second virial coefficients for the binary water/oxygen mixture have been calculated using the fitted potential energy surface. The results have rather wider confidence limits than analogous calculations for water/nitrogen and water/hydrogen. The greater uncertainty reflects a larger discrepancy between the SIMPER method used to calculate the potential energy surface and the benchmark CCSD(T) supermolecule method; the discrepancy occurs mainly in the region of the secondary (H-bonded) minimum. The reason for the discrepancy is not clear, but it may reflect the difficulty of treating unpaired electrons in van der Waals complexes.

Test calculations of the spherically averaged intermolecular potential suggest that the estimated confidence limits are quite generous, as the upper bound of the second virial coefficients is close to the second virial coefficients obtained from the UMP2 potential, whereas the RCCSD(T) potential is significantly deeper than the UMP2 potential and is closer to the SIMPER potential than it is to the UMP2 potential. (It is noted that better agreement with the limited experimental second virial coefficient data<sup>27</sup> on the water/oxygen mixture would require an even deeper potential than SIMPER.)

Other uncertainties in the second virial coefficients, which will arise from basis set incompleteness, from ignoring higher quantum corrections, and from assuming rigid monomers, are expected to be less significant than uncertainties arising from the potential energy surface. Basis set effects have been studied explicitly in this paper. The higher quantum corrections will be even less important than for water/hydrogen, and the effects of nonrigid monomers have been discussed in a previous work.<sup>3</sup>

Calculating thermodynamic properties requires averaging over a complete potential energy surface, and it is desirable to avoid methods that systematically under- or overestimate the interaction energy for most geometries. Based on the RCCSD(T) method as a benchmark, SIMPER appears to suffer less from this problem than do the UMP2 and CCSD methods, even though the computational cost of supermolecule CCSD scales much more steeply with system size than SIMPER. Thus, the deeper H-bonded minimum predicted by the SIMPER method, relative to CCSD(T), is offset to some extent by the shallower O-bonded minimum. The SIMPER method also seems to be reliable for smaller basis sets; no significant "counterpoise" correction is required, even though SIMPER uses monomer properties calculated in the monomer, not dimer, basis set. SIMPER also seems to be reliable at short intermolecular separations. For example, there do not appear to be any significant errors from the use of perturbation theory to obtain the electrostatic, induction, and dispersion energies.

## ACKNOWLEDGMENTS

The authors thank P. H. Huang for discussions concerning the experimental data of Ref. 27 and C. J. Wormald for

discussion of his enthalpy-of-mixing measurements. This work was supported by the Engineering and Physical Sciences Research Council and by a Study Abroad Fellowship for one of the authors (R.J.W.) from the Leverhulme Trust.

- <sup>1</sup>R. H. Tipping, A. Brown, Q. Ma, and C. Boulet, *J. Mol. Spectrosc.* **209**, 88 (2001).
- <sup>2</sup>M. P. Hodges, R. J. Wheatley, and A. H. Harvey, *J. Chem. Phys.* **117**, 7169 (2002).
- <sup>3</sup>A. S. Tulegenov, R. J. Wheatley, M. P. Hodges, and A. H. Harvey, *J. Chem. Phys.* **126**, 094305 (2007).
- <sup>4</sup>I. M. Svishchev and R. J. Boyd, *J. Phys. Chem. A* **102**, 7294 (1998).
- <sup>5</sup>T. W. Robinson and H. G. Kjaergaard, *J. Chem. Phys.* **119**, 3717 (2003).
- <sup>6</sup>A. Sabu, S. Kondo, N. Miura, and K. Hashimoto, *Chem. Phys. Lett.* **391**, 101 (2004).
- <sup>7</sup>A. Sabu, S. Kondo, R. Saito, Y. Kasai, and K. Hashimoto, *J. Phys. Chem. A* **109**, 1836 (2005).
- <sup>8</sup>H. G. Kjaergaard, G. R. Low, T. W. Robinson, and D. L. Howard, *J. Phys. Chem. A* **106**, 8955 (2002).
- <sup>9</sup>J. A. G. Gomes, J. L. Gossage, H. Balu, M. Kesmez, F. Bowen, R. S. Lumpkin, and D. L. Cocke, *Spectrochim. Acta, Part A* **61**, 3082 (2005).
- <sup>10</sup>A. J. Bell and T. G. Wright, *Phys. Chem. Chem. Phys.* **6**, 4385 (2004).
- <sup>11</sup>D. M. Upadhyay and P. C. Mishra, *J. Mol. Struct.: THEOCHEM* **624**, 201 (2003).
- <sup>12</sup>P. D. Cooper, H. G. Kjaergaard, V. S. Langford, A. J. McKinley, T. I. Quickenden, T. W. Robinson, and D. P. Schofield, *J. Phys. Chem. A* **109**, 4274 (2005).
- <sup>13</sup>E. N. Bichoutskaia, A. S. Tulegenov, and R. J. Wheatley, *Mol. Phys.* **102**, 567 (2004).
- <sup>14</sup>Y. Endo and M. Mizushima, *Jpn. J. Appl. Phys., Part 2* **21**, L379 (1982).
- <sup>15</sup>E. M. Mas and K. Szalewicz, *J. Chem. Phys.* **104**, 7606 (1996).
- <sup>16</sup>M. P. Hodges, R. J. Wheatley, G. K. Schenter, and A. H. Harvey, *J. Chem. Phys.* **120**, 710 (2004).
- <sup>17</sup>E. N. Bichoutskaia, M. P. Hodges, and R. J. Wheatley, *J. Comput. Methods Sci. Eng.* **2**, 391 (2002).
- <sup>18</sup>T. H. Dunning, *J. Chem. Phys.* **90**, 1007 (1989).
- <sup>19</sup>H.-J. Werner, P. J. Knowles, R. Lindh *et al.*, MOLPRO. Version 2002.6. A package of *ab initio* programs, 2003, see <http://www.molpro.net>.
- <sup>20</sup>Certain commercial products are identified in this paper but only in order to adequately specify the procedure. Such identification neither constitutes nor implies recommendation or endorsement by either the U.S. government or the National Institute of Standards and Technology.
- <sup>21</sup>R. J. Wheatley, A. S. Tulegenov, and E. N. Bichoutskaia, *Int. Rev. Phys. Chem.* **23**, 151 (2004).
- <sup>22</sup>R. J. Wheatley, *J. Comput. Chem.*, in press.
- <sup>23</sup>R. R. Fuchs, F. R. W. McCourt, A. J. Thakkar, and F. Grein, *J. Phys. Chem.* **88**, 2036 (1984).
- <sup>24</sup>J. M. L. Martin, *Chem. Phys. Lett.* **259**, 669 (1996).
- <sup>25</sup>S. L. Price, A. J. Stone, and M. Alderton, *Mol. Phys.* **52**, 987 (1984).
- <sup>26</sup>See EPAPS Document No. E-JCPSA6-127-004729 for the source code and compilation instructions for the FORTRAN subroutine described in the journal article. 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>).
- <sup>27</sup>R. G. Wylie and R. S. Fisher, *J. Chem. Eng. Data* **41**, 175 (1996).
- <sup>28</sup>K. R. Hall and G. A. Iglesias-Silva, *J. Chem. Eng. Data* **39**, 873 (1994).
- <sup>29</sup>R. W. Hyland, *J. Res. Natl. Bur. Stand., Sect. A* **79A**, 551 (1975).
- <sup>30</sup>M. Rigby and J. M. Prausnitz, *J. Phys. Chem.* **72**, 330 (1968).
- <sup>31</sup>A. H. Harvey and P. H. Huang, *Int. J. Thermophys.* **28**, 556 (2007).
- <sup>32</sup>C. J. Wormald, *J. Chem. Thermodyn.* (submitted).
- <sup>33</sup>A. H. Harvey and E. W. Lemmon, *J. Phys. Chem. Ref. Data* **33**, 369 (2004).
- <sup>34</sup>W. Wagner, J. Ewers, and R. Schmidt, *Cryogenics* **24**, 37 (1984).
- <sup>35</sup>F. Y. Naumkin and F. R. W. McCourt, *Mol. Phys.* **102**, 37 (2004).