

High-accuracy Ab Initio Rotation-Vibration Transitions for Water

Oleg L. Polyansky,^{1*} Attila G. Császár,² Sergei V. Shirin,³
Nikolai F. Zobov,³ Paolo Barletta,¹ Jonathan Tennyson,^{1†}
David W. Schwenke,⁴ Peter J. Knowles⁵

¹ Department of Physics and Astronomy, University College London,
Gower Street, London WC1E 6BT, UK.

² Department of Theoretical Chemistry, Eötvös University,
H-1518 Budapest 112, P.O. Box 32, Hungary.

³ Institute of Applied Physics, Russian Academy of Sciences,
Uljanov Street 46, Nizhnii Novgorod, Russia 603950.

⁴ NASA Ames Research Center, Moffett Field, CA 94035-1000, USA.

⁵ School of Chemical Sciences, University of Birmingham,
Birmingham B15 2TT, UK.

* Permanent address: Institute of Applied Physics, Russian Academy of Sciences,
Uljanov Street 46, Nizhnii Novgorod, Russia 603950.

† To whom correspondence should be addressed; E-mail: j.tennyson@ucl.ac.uk.

The spectrum of water vapor is of fundamental importance for a variety of processes including the absorption and retention of sunlight in our atmosphere. Therefore, there has long been an urgent need for a robust and accurate predictive model for this spectrum. We report first-principles calculations on the high-resolution spectrum of water that approach experimental accuracy. To achieve this we have performed exceptionally large electronic structure calculations and considered a variety of effects, including quantum electrodynamics, which have routinely been neglected in studies of small many-electron molecules. The high accuracy of the resulting *ab initio* procedure is demonstrated for the main isotopomers of water.

The absorption and emission of light by water vapor is responsible for about 70% of the known absorption of sunlight and 97% of the greenhouse effect (1). Hot steam is also a major product of most combustion processes and a dominant atmospheric constituent of cool stars. Water spectra have therefore been the subject of immense scientific effort but some properties, such as the details of atmospherically important weak line absorptions or the spectrum of superheated water, are extremely difficult to determine experimentally. Similarly, despite significant progress in this area (2), there exists no complete and accurate theoretical model with which one can predict the high-resolution spectrum of water.

Here we report first-principles calculations that determine the vibration-rotation energy level structure of water to better than 1 cm^{-1} on average. To achieve this accuracy it is not only necessary to perform very elaborate non-relativistic electronic structure calculations but also to consider a variety of small physical effects arising from a fully relativistic treatment, breakdown of the Born–Oppenheimer approximation, and even quantum

electrodynamics. All these corrections are routinely neglected even in high-accuracy treatments of small molecules. Nevertheless, as demonstrated below, only the combination of extrapolated large basis set electronic structure calculations and exotic corrections enables one to approach experimental accuracy during prediction of the spectrum of water.

First-principles quantum mechanical treatments of molecular spectra are usually based upon the Born–Oppenheimer approximation. According to this approximation, the motions of electrons can be separated from those of atomic nuclei because the former are so much lighter than the latter. Approximate solutions to the electronic motion problem can thus be obtained for a series of calculations at fixed nuclear geometries. The points are then fitted or otherwise interpolated to give an electronic potential energy surface (PES), which is independent of the masses of the nuclei. Nuclear motion calculations performed on such surfaces yield vibration-rotation energy levels and hence high-resolution spectra.

Only for few-particle systems, such as H_3^+ (2 electrons, 3 nuclei), have first-principles calculations approached the accuracy of high-resolution spectroscopy experiments (3, 4). For systems with more than two electrons, the goal of achieving predictions of near-spectroscopic accuracy, that is better than 1 cm^{-1} , has so far proved elusive. Water (10 electrons, 3 nuclei) has long been a benchmark system. Partridge and Schwenke (PS) (5) performed the first *ab initio* calculations for water to be accurate enough to be useful for detailed spectroscopic analysis. Yet at higher energies their calculated vibrational band origins (VBOs) are still up to 20 cm^{-1} away from experiment. It is also difficult to accurately determine how much energy it takes to distort water from its usual bent structure to a linear geometry (6, 7). One reason why water is so challenging is that upon bending the oxygen valence electrons rehybridize from sp^3 to sp , causing significant changes even to the oxygen $1s$ orbital.

To obtain high accuracy for first-principles rotation-vibration spectra, both the elec-

tronic and the nuclear motion problems must be treated accurately. For triatomic molecules moving on a given PES, the coupled rotation-vibration problem can be considered solved. Methods based on the Variational Principle give very accurate nuclear motion energy levels (see, for example, (4)). Similarly, the independent particle model for the motion of the electrons, whereby each electron moves in the average field of all others, has been solved for water yielding the so-called Hartree–Fock limit (8). The main focus of the present work is the solution of the electronic structure problem and the coupling between electronic and nuclear motion.

To improve upon the independent particle model it is necessary to account for the correlated motion of the electrons. An exact solution is offered by the full configuration interaction (FCI) technique involving every possible electron configuration. FCI calculations are only feasible for small basis sets. Very accurate approximations to FCI have been developed, and the one we use is called multireference configuration interaction (MRCI) (9). The required computations were performed using the parallel version of the electronic structure package MOLPRO (10). Correlation energy only converges slowly with increasing basis set size (11). Dunning (12) has developed a series of correlation consistent (cc) Gaussian basis sets which can be used for systematic studies. Here we report a series of calculations using Dunning’s aug-cc-pV n Z basis sets with $n = 3, 4, 5$ and 6 (13) and denoted n Z below. In this notation n stands for the number of basis functions used per physical atomic orbital and aug specifies that this basis is augmented with diffuse, low-exponent Gaussian functions. As n increases, the basis set approaches completeness, both by increasing the flexibility for a given symmetry and by including higher angular momentum components. The biggest basis we consider has $n = 6$, and consists of s, p, d, f, g, h and i functions. We found full augmentation of the basis to be of particular importance. The oxygen $1s$ core electrons were kept frozen during these

calculations.

At the 6Z MRCI level a calculation required 2 hours and 64 processors of an Origin 3000 machine for a single nuclear geometry. A total of 346 points were computed with each basis set; the initial grid of about 200 points spanned geometries up to 30,000 cm^{-1} above the water minimum. Further geometries were chosen according to their importance in determining the fitted PES.

Use of the valence-only 6Z surface plus corrections given below still gives residual errors of up to 19 cm^{-1} in the VBOs and a standard deviation for all VBOs, σ , of 5.5 cm^{-1} , see Table 1. The slow convergence with respect to n is due to the form of the electronic wave function employed in all electronic structure packages. Nevertheless, the smooth behaviour of the energies as n is systematically increased has been exploited by extrapolating the surface to the complete basis set (CBS), that is $n = \infty$, limit (8, 14-16).

Even at the CBS limit, our MRCI model does not include the entire valence electron correlation energy. This can only be achieved by performing a FCI computation which cannot presently be performed for water with basis sets larger than about 2Z. Comparison of FCI and MRCI calculations at this level (17) showed a smooth variation in the difference which rose to only 1 cm^{-1} at 25,000 cm^{-1} above the PES minimum. Because this effect is small and probably cannot be modelled accurately at the 2Z level, it has not been allowed for in the results presented below.

The largest error in our extrapolated CBS MRCI surface is due to the contribution to electron correlation effects neglected by freezing the oxygen core electrons. PS computed a correction surface (5) which they denoted CV for core-valence although in practise it also allows for core-core correlation effects. Recalculation of several points on this surface using a variety of procedures and basis sets (18) showed that PS's surface gives a reliable estimate of core correlation effects and it is employed here unchanged. Augmentation of

the CBS MRCI PES with this correction surface reduces the errors of the predicted VBOs by almost a factor of three, see Table 1.

The next most important correction to the electronic energies comes from the finite speed of light, not taken into account in nonrelativistic treatments of the electron motions. The inclusion of special relativity, which can formally be done using the Dirac Hamiltonian, gives rise to several effects (19–21). The dominant effect is a one-electron contribution arising from the high velocity of the electrons. We compute the resulting energy correction using first-order perturbation theory via the mass velocity and Darwin terms (19) by extending the available (22) surface to higher energies. This one-electron mass-velocity and Darwin (MVD1) correction surface gives the majority of the relativistic energy correction for closed-shell molecules comprised of light elements. Comparisons with full Dirac Hamiltonian calculations of Quiney *et al.* (23), who used small basis sets and less accurate wavefunctions, showed differences in the magnitude of this correction but excellent agreement in its variation with geometry, which is the key property for spectroscopy. The minuscule, two-electron contribution to the Darwin term (D2), which serves to correct point-like charge distributions, was also included (24). The Coulomb interaction, due to its instantaneous character, is not consistent with special relativity and needs to be supplemented by the Breit interaction (20). The Breit term was calculated (24) using the Dirac Hamiltonian and the program BERTHA (25).

One further effect was included in our PES, the one-electron Lamb shift. This lowest-order quantum electrodynamic effect was modelled using the prescription of Pyykkö *et al.* (26). We know of no case where the use of quantum electrodynamics (QED) has had any impact on molecular physics. However, our calculations are of such accuracy that incorporating this effect leads to a clear, systematic improvement in the results.

Thus far our treatment remained within the framework of the Born–Oppenheimer

(BO) approximation, which decouples electronic and nuclear motions. If high accuracy is required for prediction of molecular spectra, coupling between these motions must be considered. Corrections to the BO approximation can be computed efficiently by means of the second-order contact transformation method of Bunker and Moss (27). This introduces two terms: (a) the simple first-order Born–Oppenheimer diagonal correction (BODC), which allows for the action of the nuclear motion kinetic energy operator on the ground-state electronic wavefunction and gives rise to a mass-dependent correction surface; and (b) the more difficult second-order or non-adiabatic correction, which introduces coupling between electronic states and primarily results in corrections to the kinetic energy operator.

For this work an isotopomer-dependent BODC surface was computed using an MRCI wavefunction (28). Two methods of including non-adiabatic corrections to the vibrational motion were explored. The method of Schwenke (29) involving explicit coupling of the ground electronic state to electronically excited states was compared to a much simpler, two-term adjustment of the vibrational kinetic energy operator using parameters taken from Schwenke’s study (30). The results of the two methods agree to better than 0.1 cm^{-1} . These non-adiabatic treatments are only valid for energies up to 10,000 cm^{-1} and their effect is small, see Table 1. Vibrational non-adiabatic effects have therefore not been included in our final calculations. Rotational non-adiabatic corrections are, however, important for high values of the rotational quantum number, J . These are included in the results presented for H_2O isotopomers in Table 2. Again the full (29) and reduced (30) method of including this correction gave very similar results; the latter is used for the final calculations.

Except for the non-adiabatic corrections, energies for each of the above effects were generated on a grid of points. These energies were then fitted to a suitable functional

form for use in the nuclear motion calculations. Very accurate fits were obtained for all contributions to the PES except the ones due to the valence MRCI calculations. The 346 CBS MRCI data points were fitted using 113 parameters with an accuracy of 0.69 cm^{-1} (31).

The resulting compound surface gives the equilibrium structure of H_2^{16}O with an OH bond length of 0.95785 \AA and a bond angle of 104.501° , both with an uncertainty of about 5 units in the final digit. This accuracy is significantly better than that of any previous (*ab initio* or empirical) attempt (32) and is a reflection of the precision of our procedure for energies up to $15\,000 \text{ cm}^{-1}$. The barrier to linear HOH geometries is $11123 \pm 5 \text{ cm}^{-1}$ for an OH bond length of 0.93323 \AA .

Tables 1 and 2 summarize the results of our rotation-vibration energy level calculations. Nuclear motion calculations were performed using two independent codes (29, 33) which gave results which agreed to within 0.01 cm^{-1} for the same model.

There are 104 experimentally known VBOs for H_2^{16}O which arise from vibration-rotation spectra which uniquely extend into the ultraviolet (34). Similarly, analysis of hot ($T = 3000 \text{ K}$) water spectra (2) has yielded rotational energies also reaching $25,000 \text{ cm}^{-1}$ (34). Table 1 shows how the errors in the VBOs change as the model is systematically improved. In our final models the errors are dominated by states with high excitation in the ν_1 (symmetric stretching) modes.

Table 1 demonstrates the necessity of using both large basis sets, extrapolation to the CBS limit, and the four additional corrections to achieve subwavenumber accuracy. This accuracy is more than an order of magnitude better than that achieved by the previous most accurate *ab initio* study, see PS in Table 1.

Table 2 summarizes the results of extensive calculations on the five major isotopomers of water with rotational energy levels up to $J = 20$ for H_2^{16}O and $J = 12$ for other

isotopomers for which there are little data for $J > 12$. Table 2 shows that the overall standard deviation more than 17,000 levels is only 0.95 cm^{-1} and hence that our procedure, which results in a mass-dependent PES, is equally valid for all isotopomers.

We have studied the change in our VBOs as the MRCI calculation is increased from 5Z to 6Z and then extrapolated to the CBS limit. These comparisons show that the source of error for high ν_1 is caused by the limitations of the 6Z basis set. Increasing the basis set to the 7Z level (35), at an increased computational cost of about an order of magnitude, and extrapolating to a new CBS limit will affect the lower energy levels little but, according to our estimates, reduces the discrepancy of high ν_1 states to less than 1 cm^{-1} . Further improvement can be achieved by using a more dense grid of MRCI points. Nuclear motion calculations employing PESs fitted to 250, 300 and 346 MRCI points show a systematic improvement in our predictions. The use of a grid of at least 1000 MRCI points should reduce the fitting error and improve σ by 0.2 cm^{-1} . The final factor which should yield further improvement in the accuracy is the use of an accurate FCI correction. Our estimates show that including this correction could shift the highest VBOs by up to 1 cm^{-1} , giving an improvement of up to 0.2 cm^{-1} in σ .

The ultimate goal of theoretical rotation-vibration spectroscopy of water must be to model all water spectra, on the Sun, in the stars, in the atmosphere, in flames, and in the laboratory within a linewidth – or within the upper limit of the experimental accuracy – of 0.02 cm^{-1} or better. The present study improves the accuracy of the *ab initio* surface and energy levels of water by an order of magnitude in comparison with previous high accuracy work. Our calculations predict individual line positions, for which there is often significant cancellation of systematic errors, with a typical accuracy of 0.2 cm^{-1} . For example, analysis of the 571 transitions observed in sunspots and assigned to water with $J \leq 21$ (2) shows that the frequencies of these lines are reproduced by the present

calculations with a standard deviation of 0.204 cm^{-1} .

As well as accurate line positions and wavefunctions, very accurate dipole moment surfaces are required for reliable models of water spectra, including those used in atmospheric models to account for anomalous absorption and the greenhouse effect. The accurate experimental values of the Stark dipoles obtained recently (36) provide a stringent test of *ab initio* dipole moment surfaces. The electronic wavefunctions computed as part of the present work have been used to generate accurate dipole moment surfaces which are presently being tested. Finally, the benchmark results and procedures we described here can be used to predict the spectrum of other small polyatomic molecules to similar accuracy.

References and Notes

1. R. P. Wayne, *Chemistry of Atmospheres* (Oxford Univ. Press, Oxford, ed. 3, 2000), pp. 50-58.
2. O. L. Polyansky, N. F. Zobov, S. Viti, J. Tennyson, P. F. Bernath, L. Wallace, *Science* **277**, 346 (1997).
3. W. Cencek, J. Rychlewski, R. Jaquet, W. Kutzelnigg, *J. Chem. Phys.* **108**, 2831 (1998).
4. O. L. Polyansky, J. Tennyson, *J. Chem. Phys.* **110**, 5056 (1999).
5. H. Partridge, D. W. Schwenke, *J. Chem. Phys.* **106**, 4618 (1997).
6. A. G. Császár, W. D. Allen, H. F. Schaefer III, *J. Chem. Phys.* **108**, 9751 (1998).
7. E. F. Valeev, W. D. Allen, H. F. Schaefer III, A. G. Császár, *J. Chem. Phys.* **114**, 2875 (2001).

8. D. Feller, *J. Chem. Phys.*, **98**, 7059 (1993).
9. The internally contracted multi-reference CI (ICMRCI) model was used with orbitals determined using an MC-SCF scheme and the reference configurations of PS (5).
10. R. D. Amos *et al.* MOLPRO is a package of *ab initio* programs designed by H.-J. Werner and P. J. Knowles, version 2002.1 (2002).
11. W. Klopper, K. L. Bak, P. Jørgensen, J. Olsen, T. Helgaker, *J. Phys. B: At. Mol. Opt. Phys.* **32**, R103 (1999).
12. T. H. Dunning Jr., *J. Chem. Phys.* **90**, 1007 (1999).
13. The $n = 3$ and $n = 4$ basis sets are conventionally denoted as aug-cc-pVTZ and aug-cc-pVQZ, respectively.
14. T. Helgaker, W. Klopper, H. Koch, J. Noga, *J. Chem. Phys.* **106**, 9639 (1997).
15. A. Halkier, T. Helgaker, P. Jørgensen, W. Klopper, J. Olsen, *Chem. Phys. Lett.* **302**, 437 (1999).
16. MC-SCF energies were extrapolated using an exponential form (15) while a polynomial with $n = 3$ was used for the CI energies (14).
17. Full CI calculations were performed with a double zeta plus polarisation basis set.
18. Our tests included basis sets up to appropriately augmented 6Z, significantly larger than PS (5).
19. R. D. Cowan, D. C. Griffin, *J. Opt. Soc. Amer.* **66**, 1010 (1976).
20. P. Pyykkö, *Adv. Quantum Chem.* **11**, 353 (1978).

21. G. Tarczay, A. G. Császár, W. Klopper, H. M. Quiney, *Mol. Phys.* **99**, 1768 (2001).
22. A. G. Császár, J. S. Kain, O. L. Polyansky, N. F. Zobov, J. Tennyson, *Chem. Phys. Lett.*, **293**, 317 (1998); erratum **312**, 613 (1999).
23. H. M. Quiney, H. Skaane, I. P. Grant, *Chem. Phys. Lett.* **290**, 473 (1998).
24. H. M. Quiney, P. Barletta, G. Tarczay, A. G. Császár, O. L. Polyansky, J. Tennyson, *Chem. Phys. Lett.* **344**, 413 (2001).
25. I. P. Grant, H. M. Quiney, *Int. J. Quantum Chem.* **80**, 283 (2000).
26. P. Pyykkö, K. Dyall, A. G. Császár, G. Tarczay, O. L. Polyansky, J. Tennyson, *Phys. Rev. A* **63**, 024502 (2001).
27. P. R. Bunker, R. E. Moss, *J. Mol. Spectrosc.* **80**, 217 (1980).
28. The BODC surface was computed using the aug-cc-pVTZ basis set. This surface agrees well with the CASSCF surface (29) computed with the same basis at low energy but less well for the high stretching states. It remains unclear if this model is converged.
29. D. W. Schwenke, *J. Phys. Chem. A* **105**, 2352 (2001).
30. J. Tennyson, P. Barletta, M. A. Kostin, O. L. Polyansky, N. F. Zobov, *Spectrochim. Acta A* **58**, 663 (2002).
31. The functional form used to fit the valence MRCI PES followed that of PS (5) but adjustment of the parameters representing the behaviour at dissociation proved crucial for obtaining a satisfactory fit. Other surfaces were fitted with functional forms given previously (22, 24, 26, 29).

32. The most accurate empirical determinations of the equilibrium geometry of water rely on spectroscopically determined potentials (K. Kutchitsu, private communication). Our present *ab initio* geometry parameters agree, to within our quoted error bars, with the best of these (5).
33. J. Tennyson, J. R. Henderson, N. G. Fulton, *Computer Phys. Comms.* **86**, 175 (1995).
34. J. Tennyson, N. F. Zobov, R. Williamson, O. L. Polyansky, P. F. Bernath, *J. Phys. Chem. Ref. Data* **30**, 735 (2001).
35. E. F. Valeev, W. D. Allen, R. Hernandez, C. D. Sherrill, H. F. Schaefer III, *J. Chem. Phys.* submitted for publication.
36. A. Callegari *et al*, *Science* **297**, 993 (2002).
37. These calculations used the a 512 processor Origin3000 machine Green run by the CSAR Service at the University of Manchester; we thank the staff at CSAR for their help. We also thank the UK Engineering and Physical Science Research Council, the UK Natural Environment Research Council, the Royal Society, the British Council, the INTAS foundation, the Scientific Research Fund of Hungary (grant OTKA T033074), and the Russian Fund for Fundamental Studies who supported this project.

Table 1

Predicted vibrational band origins (VBOs) for various theoretical models. Results are presented as differences from the observed values in cm^{-1} . The standard deviation, σ , is for all experimentally known VBOs.

state	obs (34)	5Z	6Z	CBS	PS (5)	CBS+CV	+rel	+qed	+BODC	+non-ad
(010)	1594.74	-2.99	-2.29	-0.32	-2.79	0.48	-0.81	-0.75	-0.32	-0.27
(020)	3151.63	-4.22	-2.38	-0.78	-5.38	1.16	-1.57	-1.44	-0.56	-0.44
(030)	4666.78	-6.30	-3.24	-1.52	-7.91	2.05	-2.37	-2.16	-0.78	-0.60
(040)	6134.01	-9.81	-5.53	-2.74	-10.38	3.20	-3.30	-3.00	-1.06	-0.83
(050)	7542.43	-14.70	-9.18	-4.71	-12.90	4.82	-4.45	-4.02	-1.41	-1.14
(101)	7249.81	12.51	10.76	9.32	-4.78	-5.35	1.70	1.43	0.60	2.00
(201)	10613.35	18.72	16.46	13.97	-6.96	-7.47	2.98	2.57	1.23	
(301)	13830.93	25.72	22.81	18.74	-8.41	1.95	4.59	4.06	2.05	
(401)	16898.84	32.56	28.92	23.06	-9.47	-10.17	6.11	5.49	2.74	
(501)	19781.10	40.72	35.96	28.68	-9.31	-10.72	9.04	8.28	4.65	
(601)	22529.44	51.14	43.41	34.17	-7.61	-11.88	11.69	10.81	5.94	
(701)	25120.27	63.29	51.75	38.66	-5.49	-3.13	13.70	12.75	6.46	
all	σ	22.84	19.74	16.56	10.44	7.85	4.23	3.83	1.90	

Footnotes for Table 1.

5Z	aug-cc-pV5Z MRCI calculation
6Z	aug-cc-pV6Z MRCI calculation
CBS	MRCI calculation extrapolated to the complete basis set limit.
PS	partially augmented cc-pV5Z MRCI calculation plus core correlation due to Partridge and Schwenke (5).
CBS+CV	CBS with core correlation correction.
rel	CBS+CV with relativistic effects included.
QED	rel with one electron Lamb shift included.
BODC	QED with Born-Oppenheimer Diagonal Correction included.
non-ad	BODC with vibrational non-adiabatic effects included.

Table 2

Standard deviation (σ) with which our final (CBS+CV+rel+QED+BODC) potential reproduces the vibrational-rotation term values for various water isotopomers with rotational excitation $\leq J_{\max}$. $N(\text{levels})$ levels are considered in each case. Also given is the maximum error for each isotopomer, in cm^{-1} , and the associated value of J . The H_2O isotopomer calculations include rotational non-adiabatic effects.

Isotopomer	σ cm^{-1}	J_{\max}	$N(\text{levels})$	Maximum Deviation	
				obs-calc	J
H_2^{16}O	1.17	20	9426	6.5	7
H_2^{17}O	0.56	12	1083	1.4	12
H_2^{18}O	0.65	12	2460	2.3	6
D_2^{16}O	0.71	12	2807	3.0	7
HD^{16}O	0.47	12	2019	-1.2	11
All	0.95	20	17795		