

Improved estimates of the total correlation energy in the ground state of the water molecule

Arne Lüchow and James B. Anderson

Department of Chemistry, The Pennsylvania State University, University Park, Pennsylvania 16802

David Feller

Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, Washington 99352

(Received 1 October 1996; accepted 5 February 1997)

Two new calculations of the electronic energy of the ground state of the water molecule yield energies lower than those of any previously reported variational calculations. A fixed-node quantum Monte Carlo calculation gives $-76.420(1)$ hartrees and an analytic variational calculation gives -76.4274 hartrees. These values lie only 17 and 11 mhartrees, respectively, above the “experimental” value. © 1997 American Institute of Physics. [S0021-9606(97)00718-6]

We report here two new calculations of very high accuracy for the electronic energy of the ground state of the water molecule. The energy given by a fixed-node quantum Monte Carlo calculation is $-76.421(1)$ hartrees, a value lower than any previously reported variational energy and only 0.017 hartrees above the experimentally-derived nonrelativistic, infinite-mass value of $-76.438(3)$ hartrees. The energy given by an analytic variational calculation with a 341-function basis set approaching the complete-basis-set, full-configuration interaction limit is -76.4274 hartrees, a value even lower and only 0.011 hartrees above the experimentally derived value.

Fixed-node quantum Monte Carlo (FN-QMC) calculations^{1,2} are variational with respect to node locations. Like analytic variational calculations they provide an upper bound to the energy. For the water molecule and a number of other small molecules the lowest variational energies have been provided by FN-QMC calculations. If exact node locations could be specified the exact energy could be calculated, but exact node locations cannot be specified in advance except in the simplest cases. The best node locations give the lowest energies and these have, in general, been obtained in the form of single-determinant trial wavefunctions produced in analytic SCF calculations. Attempts to obtain improved node locations with multiconfiguration trial functions have not led to significantly lower energies for the water molecule. In the work reported here the node locations were obtained from a single-determinant function of near-Hartree–Fock-limit quality. The single-determinant function was multiplied by a Jastrow (Bijl) function to improve the electron-electron correlation without modifying the node locations.

The calculations were carried out for the water molecule in a single geometry corresponding to the experimentally determined equilibrium geometry. The same geometry has been used in earlier calculations of several types. It is given by: $r_{\text{OH}} = r_{\text{OH}'} = 0.9752 \text{ \AA}$, $\angle \text{HOH} = 104.52 \text{ deg}$.

The SCF function was determined for an STO basis set of the type $(5s, 4p, 2d/3s, 2p)$ derived from the STO basis for OH reported by Cade and Huo.³ The analytic SCF calculation was carried out after expanding the STO's into 10 (for 1s) to 6 (3d) Gaussians using the GAMESS⁴ program. The

resulting function was used without additional modification or simplification. The variational energy for this function alone was -76.064 hartrees, value only 0.004 hartrees above the estimated Hartree–Fock limit. The Jastrow function incorporated terms of the type e^U with

$$U_{aij} = \sum_k^{N_a} c_{ka} (\bar{r}_{ai}^{jka} \bar{r}_{aj}^{mka} + \bar{r}_{aj}^{jka} \bar{r}_{ai}^{mka}) \bar{r}_{ij}^{nka}, \quad (1)$$

where a and i, j refer to the nuclei and the electrons, respectively, and where \bar{r} is defined by $\bar{r} = br/(1+br)$.

Optimization of the coefficients of the Jastrow term was carried out by minimizing the variance in the local energy $E_{\text{loc}} = H\Psi/\Psi$ in a manner similar to that of a number of earlier works.⁵ The expectation value of the energy for the optimized trial function was determined in Metropolis Monte Carlo calculations to be $-76.371(1)$ hartrees. This corresponds to recovery of 81% of the correlation energy. The standard deviation in the local energy was 1.0 hartree.

The fixed-node QMC calculations were carried out using a diffusion algorithm along with importance sampling based on the trial wavefunction. The algorithm was essentially the same as that for earlier calculations in our laboratory⁶ but we added an acceptance/rejection step as described by Reynolds *et al.*⁷ to aid in reducing time-step error. The calculations were executed on parallel IBM SP2 machines located at the Maui High Performance Computer Center.

Calculations were performed with several different time-step sizes to allow extrapolation to a time-step of zero. The step sizes, the energies, and their statistical uncertainties were as follows: $[\Delta\tau \text{ (au)}, E \text{ (hartrees)}]$: 0.002, $-76.4210(7)$; 0.005, $-76.4224(7)$; 0.007, $-76.4250(7)$; 0.010, $-76.4281(7)$; 0.012, $-76.4306(7)$; 0.015, $-76.4384(7)$. A least-squares fit of a quadratic function gave for the intercept at zero step size the energy of $-76.421(1)$ hartrees.

The result is listed in Table I along with selected results from prior calculations of several types and with the accepted experimental value and that value adjusted to the non-relativistic, infinite-mass case. The calculated energy is 0.010 hartree lower than the previous low energy obtained

TABLE I. Selected results of calculations for the ground state of the water molecule.^a

	Total energy (hartree)	Type ^b	Correlation energy (percent)
Estimated Hartree–Fock limit ^c	−76.068(1)		0.0
Boys, Cook, Reeves, and Shavitt (1956) ^d			
CI/30 GTOs	−75.776	V	...
Reynolds, Ceperley, Alder, and Lester (1982) ^e			
FN-QMC	−76.377(7)	V	83.5
Rosenberg & Shavitt (1975) ^f			
CI-SD/39 STOs	−76.340	V	73.5
Bartlett <i>et al.</i> (1987) ^g			
MP4/39 STOs	−76.360		78.9
Feller, Boyle, & Davidson (1987) ^h			
MRCI-SD/140 CGTFs	−76.396	V	88.6
Kim, Lee, Lee, Mhin, & Kim (1995) ⁱ			
QCISD(T)	−76.406		91.4
MP4-SDTQ	−76.407		91.6
Huang & Cao (1996) ^j			
FN-QMC	−76.411(2)	V	92.7
This work, quantum Monte Carlo			
V-QMC	−76.371(1)	V	81.9
FN-QMC	−76.421(1)	V	95.4
This work, analytic variational			
aug-cc-pCV5Z iCAS-CI	−76.4274	V	97.1
estimated CBS/FCI limit	−76.431(3)	V	98.1
Experimental value (adjusted to nonrelativistic, infinite nuclear mass) ^c	−76.438(3)		100.0
Experimental value ^k	−76.480(1)		

^aAdditional values are given by Levine (Ref. 8). Uncertainties in the final digit are given in parentheses.

^bV = variational. Energy is an upper bound.

^cRevised value of Feller, Boyle, and Davidson (Ref. 9).

^dReference 10.

^eReference 7.

^fReference 11.

^gReference 12.

^hReference 9.

ⁱReference 13.

^jReference 14.

^kSee text.

by Huang and Cao¹⁴ and 0.017 hartree higher than the adjusted experimental value.

An adjusted experimental value of −76.440(3) hartrees was derived in 1987 by Feller, Boyle, and Davidson⁹ from available data and calculations. The value depends upon estimates of the exact, nonrelativistic, infinite-nuclear-mass energy of the oxygen atom. Using a more recent value¹⁵ for this energy gives a revised experimental value of −76.438(3) hartrees. The indicated uncertainty results primarily from an uncertainty of 0.0020 hartrees in the relativistic correction of −0.045 hartrees.

In our analytic variational calculations we obtain an alternative estimate of the total correlation energy by fitting the energies obtained from a systematic sequence of wavefunctions that approach the complete basis (CBS) set, full configuration interaction (FCI) limit. The core/valence correla-

tion consistent basis sets of Dunning and co-workers^{16–18} were used to help insure smooth convergence to the CBS limit. Additional diffuse functions, taken from the augmented correlation consistent family of basis sets, were also included. Thus, the final basis sets are denoted as aug-cc-pCVDZ, aug-cc-pCVTZ, etc. in accord with the usual labeling convention. The largest basis set used in the present work is the aug-cc-pCV5Z set which includes 341 functions on water and up through two sets of *h* functions on oxygen and *g* functions on hydrogen. The total energies from each basis set are fitted with a simple exponential function of the form $E(n) = E_{\text{CBS}} + B \exp(-Cn)$, where $n = 2$ (DZ), 3 (TZ), etc. A large number of molecular energies have been shown to be well described by this functional form.¹⁹

Convergence to the full CI limit is achieved with a sequence of multireference, single and double excitation (MR SD-CI) wave functions. Details of these wave functions are given elsewhere.²⁰ The wave functions are indexed by the sum of the squares of the reference configuration's CI coefficients, $\sum c_i^2$, which approaches unity as the CI approaches the FCI limit. This sequence of energies is described by another exponential form, with $\sum c_i^2$ as the independent parameter. In a relatively small number of cases where FCI energies were available, this procedure was shown to be capable of extrapolating to the correct limit with an error of less than 1 mhartree. Thus, our approach for estimating the CBS, FCI energy relies upon first estimating the FCI energy for each basis set and then extrapolating the results from the DZ through 5Z sets to the CBS limit. This approach has been used previously to estimate the valence space correlation limit in water.²¹ It is difficult to derive meaningful error bars for this procedure, since it employs simultaneously extrapolations in both the 1-particle and *n*-particle basis sets. Judging by the agreement of the fitting functions with the raw data, which is always better than ± 0.5 mhartree and experience with basis sets as large as sextuple zeta, we propose conservative error bars of ± 3 mhartree.

The MR SD-CI calculations with the DZ through QZ basis sets were performed with the MELDF-X program.²² Because this program cannot handle *h* functions, the aug-cc-pCV5Z calculations were performed with the MOLPRO-96 codes.²³ MOLPRO was unable to perform the same type of MR SD-CI calculations as MELDF-X. Therefore, we calibrated an internally contracted, complete active space CI (iCAS-CI) based on a (6a1/3b2/2b1) active space with the DZ through QZ basis sets. This wavefunction obtained 99.8% of the estimated FCI energy for the TZ and QZ basis sets. Therefore, we assume that the aug-cc-pCV5Z iCAS-CI correlation energy represents 99.8% of the full CI value.

The lowest variational energy, corresponding to the aug-cc-pCV5Z basis set iCAS-CI was −76.4274 hartrees. For comparison purposes, the corresponding Hartree–Fock (HF) energy was −76.0673 hartrees. At the optimal HF geometry (0.940 Å, 106.3 degrees) this basis set gave $E(\text{RHF}) = -76.0680$ hartrees. The estimated CBS, FCI energy is -76.431 ± 0.003 hartrees, which is in reasonable agreement with the “experimental” value of −76.438(3) hartrees.

We have also performed coupled-cluster singles and doubles calculations, with perturbative inclusion of triple excitations, CCSD(T), using the same correlation consistent basis sets. CCSD(T) energies are lower than our FCI estimates for the TZ and larger basis sets. Our lowest CCSD(T) energy, obtained with the aug-cc-pCV5Z basis, was -76.4323 hartrees. An exponential fit of the TZ, QZ, and 5Z CCSD(T) energies yields a CBS estimate of -76.436 hartrees, slightly higher than the estimate of -76.439 hartrees recently obtained from r12 calculations with even larger basis sets.²⁴ Since CCSD(T) probably overestimates the true correlation energy, the real value probably lies somewhere between -76.431 hartrees (our CI estimate) and -76.439 hartrees [the CCSD(T) estimate].

The research was supported in part by the National Science Foundation (Grant No. CHE-8714613) and by the Office of Naval Research (Grant No. N00014-92-J-1340). A.L. is indebted to the Alexander von Humboldt Foundation for financial support by the Feodor-Lynen program. We acknowledge a generous allotment of cpu time at the Maui High Performance Computer Center. D.F. was supported by the U.S. Department of Energy under Contract No. DE-AC06-76RLO 1830. The authors acknowledge the support of the Division of Chemical Sciences, Office of Basic Energy Sciences. Portions of this work were completed on the computer resources at the National Energy Research Supercomputer Center with a grant provided by the Scientific Computing Staff, Office of Energy Research, U.S. Department of Energy. The Pacific Northwest National Laboratory is a multiprogram national laboratory operated by Battelle Memorial Institute.

¹B. L. Hammond, W. A. Lester, Jr., and P. J. Reynolds, *Monte Carlo Methods in Ab Initio Quantum Chemistry* (World Scientific, Singapore, 1994).

²J. B. Anderson, *Int. Rev. Phys. Chem.* **14**, 85 (1995).

³P. E. Cade and W. M. Huo, *At. Data Nucl. Data Tabl.* **12**, 415 (1973).

⁴M. W. Schmidt, K. K. Baldrige, J. A. Boatz, S. T. Ebert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. J. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery, *J. Comput. Chem.* **14**, 1347 (1993).

⁵D. H. Weinstein, *Proc. Nat. Acad. Sci.* **20**, 529 (1934); A. A. Frost, *J. Chem. Phys.* **10**, 240 (1942); H. Conroy, *ibid.* **40**, 603 (1964).

⁶F. Mentch and J. B. Anderson, *J. Chem. Phys.* **74**, 6307 (1981).

⁷P. J. Reynolds, D. M. Ceperley, B. J. Alder, and W. A. Lester, *J. Chem. Phys.* **77**, 5593 (1982).

⁸I. N. Levine, *Quantum Chemistry*, 4th ed. (Prentice-Hall, Englewood Cliffs, New Jersey, 1991), pp. 474, 510.

⁹D. Feller, C. M. Boyle, and E. R. Davidson, *J. Chem. Phys.* **86**, 3424 (1987).

¹⁰S. F. Boys, G. B. Cook, C. M. Reeves, and I. Shavitt, *Nature* **178**, 1207 (1956).

¹¹B. J. Rosenberg and I. Shavitt, *J. Chem. Phys.* **63**, 2162 (1975).

¹²R. J. Bartlett, S. J. Cole, G. D. Purvis, W. C. Ermler, H. C. Hsieh, and I. Shavitt, *J. Chem. Phys.* **87**, 6579 (1987).

¹³K. S. Kim, J. Y. Lee, S. Lee, B. J. Mhin, and K. S. Kim, *J. Chem. Phys.* **102**, 310 (1995).

¹⁴H. Huang and Z. Cao, *J. Chem. Phys.* **104**, 200 (1996).

¹⁵E. R. Davidson, S. A. Hagstrom, S. J. Chakravorty, V. M. Umar, and C. F. Fischer, *Phys. Rev. A* **44**, 7071 (1991).

¹⁶T. H. Dunning, Jr., *J. Chem. Phys.* **90**, 1007 (1989).

¹⁷R. A. Kendall, T. H. Dunning, Jr., and R. J. Harrison, *J. Chem. Phys.* **96**, 6796 (1992).

¹⁸D. E. Woon and T. H. Dunning, Jr., *J. Chem. Phys.* **103**, 4572 (1995).

¹⁹See, for example, D. Feller, *J. Chem. Phys.* **96**, 6104 (1992); D. E. Woon and T. H. Dunning, Jr., *ibid.* **99**, 1914 (1993); K. A. Peterson, R. A. Kendall, and T. H. Dunning, Jr., *ibid.* **99**, 1930 (1993).

²⁰R. J. Cave, S. S. Xantheas, and D. Feller, *Theor. Chim. Acta* **83**, 31 (1992).

²¹D. Feller, *J. Chem. Phys.* **98**, 7059 (1993).

²²MELDF-X, L. McMurchie, S. Elbert, S. Langhoff, and E. R. Davidson, The University of Washington, Indiana University and the Molecular Science Research Center, (1993). The MELDF suite of programs was substantially modified by D. Feller, R. Cave, D. Rawlings, R. Frey, R. Daasch, L. Nitchie, P. Phillips, K. Iberle, C. Jackels, and E. R. Davidson.

²³MOLPRO is a package of *ab initio* programs written by H.-J. Werner and P. J. Knowles, with contributions from J. Almlof, R. D. Amos, M. J. O. Deegan, S. T. Elbert, C. Hampel, W. Meyer, K. A. Peterson, R. M. Pitzer, A. J. Stone, P. R. Taylor, and R. Lindh, Universität Bielefeld, Bielefeld, Germany, University of Sussex, Falmer, Brighton, England, 1996.

²⁴T. Helgaker, W. Klopper, H. Koch, and J. Nagel, *J. Chem. Phys.* (to be published).