FCI quality energy for H_2O with a frozen core in cc-pV9Z: 8 electrons in 957 orbitals with micro-hartree precision

Nike Dattani[1]

[1]*University of Waterloo, Department of Electrical and Computer Engineering, Canada*

The total electronic energy for the ground electronic state of a water molecule, using the standard molecular Hamiltonian and a Dunning basis set with 957 spatial orbitals (1914 spin orbitals), is calculated with the two deepest core electrons frozen and the 8 remaining electrons fully correlated. The 8-electron energy is estimated to be accurate to within ± 2 micro-hartree of the true energy for this Hamiltonian and basis set. The calculations were done using a fully deterministic, composite coupled cluster approach.

I. INTRODUCTION

Empirical ionization energies of the water molecule have been determined to a precision of at best ± 0.003 kcal/mol (about 4.8 μE_h) in 2015 [1, 2], and empirical atomization energies of the water molecule have been determined to a precision of at best . Recently, computer spectroscopy has been able to determine energy differences with accuracies of around $\pm 1\,\mathrm{cm}^{-1}$ (2.2 μE_h) or better, for atoms and small molecules with a small number of electrons: If total energies can be calculated to with a precision of $\pm\sigma$ then energy differences involving them can be determined with a precision $\pm\sqrt{2}\sigma$ assuming gaussian?. Therefore calculating total (electronic?) energies of the water molecule to a precision of around ± 3 μE_h or better can be seen as a first step towards determining properties of the water molecule with unprecedented precision. The total electronic energy of a molecule can have many calculations contributing to it, and none of them can be calculated with a precision worse than about ± 3 μE_h if we wish to determine the total (electronic?) energy to no worse than ± 3 μE_h.

In this paper I calculate one contribution (the standard Hamiltonian...) to within ± 2 uH which is likely the best precision ever achieved for 8e- in 957 orbitals. Determining properties to this accuracy would require basis set extrapolations, which typically (in the most standard way) involve CBS for HF, which itself will need some work, but at least this paper eliminates the FCI portion of things as the bottleneck in this journey.

II. IMPLEMENTATION

A. Hardware

Despite the largest calculations in this work being state-of-the-art by a large margin, they were actually done on relatively old hardware (all calculations were done in 2019 on hardware from 2014). While this would have made the calculations significantly slower than they could have been (in terms of wall clock time), it meant that access to the hardware was more readily available than if using any of the newer (and more in-demand by other scientists) resources available to the author. This was especially important, since there was a lot of de-bugging involved, as the integrals, scf were done with a new code, which had not been tested on such basis sets before. Furthermore, the particular nodes on which the calculations were done, happened to have 55? T of node local storage, which was in fact necessary for the calculations using the biggest basis sets, and is not available on any of the other resources available to the author.

All calculations were done on Ontario's SHARCNET cluster called `wobbie`, on compute nodes containing two sockets, each with a 14-core Xeon E5-2697 v3 (from the Haswell-EP model, which used a 22nm process, and was first released in September 2014) processors (28 cores per node). [Check if I used 768G at any point].

B. Software

For the small basis sets (cc-pVXZ for $X = 2$ to 6), the following programs from the `CFOUR v2.00beta` package [3], which was released in March 2014, were used: `vmol` for the 1- and 2-electron integrals, `vscf` for the Hartree-Fock calculations, `vtran` for the AO to MO transformation, and `intprc` and `int` for processing and printing the integrals into `ASCII` format for the software package `MRCC` [4] to read and do the coupled cluster calculations. The July 2016 release of `MRCC` was used.

For the cc-pV7Z, cc-pV8Z, and cc-pV9Z basis sets, the aforementioned programs in `CFOUR v2.00beta` cannot treat the necessary k-type, l-type and m-type Gaussian-type orbitals involved, so a development versions of `vmol`, `vscf`, and `vtran` written by Devin Matthews, were de-bugged and tested as part of this work. `MRCC` was still used to do CCSD and CCSD(T) calculations for all three of these basis sets, but due to RAM limitations (? or size of fort.55?), CCSDT (which was only done for cc-pV7Z), was done using the program `ncc` from the `CFOUR v2.1` package [3], which was released in July 2019. CCSD and CCSD(T) calculations for all three of these basis sets were also done with `ncc`, `ecc`, and `vcc` of `CFOUR v2.1` in addition to the aforementioned calculations done in `MRCC`, and the difference among the different implementations of CCSD and

CCSD(T) was never more than 2 nE_h (nano-Hartrees), which gives us hope that our use of `ncc` for CCSDT with cc-pV7Z and `MRCC` for all smaller basis sets, was not problematic.

C. Data

The cc-pVXZ basis sets for $X = 2$ to 6, were obtained from the EMSL Basis Set Exchange before January 2019 (the input file containing the exact exponents and contraction coefficients used, and the file's version history, can be found in the file `GENBAS` on my GitHub page). For $X = 2$ to 5 the basis sets were originally published in [5], and for $X = 6$ the basis sets were originally published in [6].

For $X = 7$ to 9, the basis sets were obtained privately by David Feller, who built them for []. They were then converted to `CFOUR`'s input format and inserted into the same `GENBAS` file located in [].

III. METHODOLOGY

A. Hamiltonian

Standard Hamiltonian definition. NR-CPN.

B. Basis sets

The electronic wavefunction was modeled using a linear combination of atomic orbitals (LCAO) in which Gaussian-type basis sets were used for the atomic orbitals. A series of these basis sets was used, from the cc-pVXZ family, which was first proposed by Dunning [5]. The references for the specific cc-pVXZ basis sets used, were given in Section [].

The basis sets were contracted using a general contraction scheme (`CONTRACTION` = `GENERAL` in `CFOUR`), and the total number of spatial orbitals for each basis set

C. 1- and 2-electron integral calculations

The transformed integrals were calculated with the `CFOUR` keyword `XFORM_TOL` set to 18 as opposed to the default value of 11, meaning that only integrals with values smaller than 10^{-18} would be neglected as opposed to the default tolerance threshold of 10^{-11} (effectively, no integrals are ignored).

D. Hartree-Fock calculations

RHF (restricted Hartree-Fock) calculations were done with the spin multiplicity set to 1 (to obtain a singlet state). A core Hamiltonian initial guess was used, with no restriction on the occupation numbers in terms of spatial symmetry within the C_{2v} point group which was used, but at convergence the final occupation numbers were always 6 for the A_1 irrep (3 electrons with spin $+\frac{1}{2}$ and 3 with $-\frac{1}{2}$), 2 for the A_2 and 2 for the B_1 irreps (in each case, 1 electron with each spin direction), and 0 for B_2.

The tolerances were....

The output files, which include the input files, are given here, and show all the settings used.

E. Electron correlation calculations

F. Estimation of $\pm 2\mu E_h$ uncertainty for FCI energy with cc-pV9Z

IV. RESULTS

summary of results, maybe this table and a table showing the differences in CC values

V. DISCUSSION

A. Bottlenecks:

Bottlenecks:

- lack of method to get CBS limit of HF for triatomics.

- speed of CCSDTQ for 7Z (cfour will take 150 days? with 768G or would require 1.5T which doesn't allow more than 30 days, and NCC doesn't allow restarts ... not sure about MRCC... Hanrath program unavailable.)

- speed of CCSDTQP for 5Z (lack of hand-coded CCSDTQP, lack of parallel I/O in MRCC, lack of Hanrath code, speed of CC-QMC to get error down).

- ability to do CCSDTQ(P) for 5Z/6Z, or CCSDT(Q) for 7Z, 8Z, 9Z (MRCC can't restart perturbative calculations).

- getting higher than CCSDTQP is not necessary for any basis sets that it's not already possible for, and getting higher than CCSDTQ does not seem necessary for anything beyond 7Z?

[1] C. Lauzin, U. Jacovella, and F. Merkt, Molecular Physics **113**, 3918 (2015).

[2] D. Feller and E. R. Davidson, Journal of Chemical Physics **148**, 234308 (2018).

Table I.

	cc-pV2Z	cc-pV3Z	cc-pV4Z	cc-pV5Z	cc-pV6Z	cc-pV7Z	cc-pV8Z	cc-pV9Z
RHF	**-76.026 778 80**	**-76.057 139 16**	**-76.064 804**	**-76.067 060**	**-76.067 376**	**-76.067 426**	**-76.067 449**	**-76.067 455**
CCSD	**-76.237 998 10**	**-76.324 545 75**	**-76.350 801**	**-76.359 519**	**-76.362 323**	**-76.363 594**	**-76.364 296**	**-76.364 659**
CCSD(T)	**-76.241 033 59**	**-76.332 192 23**	**-76.359 794**	**-76.369 041**	**-76.372 019**	**-76.373 362**	**-76.374 100**	**-76.374 481**
CCSDT	**-76.241 195 72**	**-76.332 268 49**	**-76.359 807**	**-76.368 997**	**-76.371 941**	**-76.373 264**	**-76.373 990**	**-76.374 364**
CCSDT(Q)	**-76.241 679 91**	**-76.332 623 49**	**-76.360 217**	**-76.369 431**	**-76.372 384**	-76.373 706	-76.374 432	-76.374 806
CCSDTQ	**-76.241 650 58**	**-76.332 587 70**	**-76.360 183**	**-76.369 394**	**-76.372 345**	**-76.373 667**	-76.374 393	-76.374 768
CCSDTQ(P)	**-76.241 665 00**	**-76.332 602 77**	**-76.360 189**	-76.369 401	-76.372 352	-76.373 674	-76.374 400	-76.374 774
CCSDTQP	**-76.241 665 92**	**-76.332 606 22**	**-76.360 198**	-76.369 410	-76.372 361	-76.373 683	-76.374 409	-76.374 783
CCSDTQP(H)	**-76.241 667 98**	**-76.332 606 77**	-76.360 199	-76.369 410	-76.372 361	-76.373 683	-76.374 410	-76.374 784
CCSDTQPH	**-76.241 668 32**	**-76.332 606 94**	-76.360 199	-76.369 410	-76.372 361	-76.373 683	-76.374 410	-76.374 784
CCSDTQPH(S)	**-76.241 668 38**	-76.332 606 99	-76.360 199	-76.369 410	-76.372 361	-76.373 683	-76.374 410	-76.374 784
CCSDTQPHS	**-76.241 668 43**	-76.332 607 05	-76.360 199	-76.369 410	-76.372 362	-76.373 684	-76.374 410	-76.374 784
CCSDTQPHS(O)	**-76.241 668 43**	-76.332 607 05	-76.360 199	-76.369 410	-76.372 362	-76.373 684	-76.374 410	-76.374 784
FCI = CCSDTQPHSO	**-76.241 668 45**	-76.332 607 07	-76.360 199	-76.369 410	-76.372 362	-76.373 684	-76.374 410	-76.374 784

Table II.

	cc-pV2Z	cc-pV3Z	cc-pV4Z	cc-pV5Z	cc-pV6Z	cc-pV7Z	cc-pV8Z	cc-pV9Z
CCSD(T)-CCSD	-3035.	-7646.	-8993.	-9522.	-9696.	-9768.	-9805	-9822
CCSDT-CCSD(T)	-162.	-76.	-13.	44.	77.	98.	110	117
CCSDT(Q)-CCSDT	-484.	-355.	-410.	-434.	-442.			
CCSDTQ-CCSDT(Q)	29.3	35.8	34.4	36.7	38.4	**38.9(0.5)**		
CCSDTQ(P)-CCSDTQ	-14.4	-15.1	-6.5					
CCSDTQP-CCSDTQ(P)	-0.9	-3.5	-9.0					
CCSDTQP-CCSDTQ	-15.3	-18.5	**-15.5(3.0)**					
CCSDTQP(H)-CCSDTQP	-2.1	-0.5						
CCSDTQPH-CCSDTQP(H)	-0.3	-0.2						
CCSDTQPH-CCSDTQP	-2.4	**-0.7(1.7)**						
CCSDTQPH(S)-CCSDTQPH	-0.053							
CCSDTQPHS-CCSDTQPH(S)	-0.053							
CCSDTQPHS(O)-CCSDTQPHS	-0.003							
CCSDTQPHSO-CCSDTQPHS(O)	-0.019							
CCSDTQPHSO-CCSDTQP	**-0.128**							

Table III.

	cc-pV9Z
RHF	-76.067 455 2
CCSDT	-76.374 364 2
CCSDTQ	-76.373 961
CCSDTQP	-76.373 976
CCSDTQPH	-76.373 977
FCI = CCSDTQPHO	-76.373 977

[3] J. F. Stanton, J. Gauss, L. Cheng, M. E. Harding, D. A. Matthews, and P. G. Szalay, "CFOUR, Coupled-Cluster techniques for Computational Chemistry, a quantum-chemical program package,".

[4] M. Kállay, Z. Rolik, J. Csontos, I. Ladjánszki, L. Szegedy, B. Ladóczki, and G. Samu, "\texttt{MRCC}, a Quantum Chemical Program Suite,".

[5] T. H. Dunning, Jr, J. Chem. Phys. **90**, 1007 (1989).

[6] A. K. Wilson, T. van Mourik, and T. H. Dunning, Journal of Molecular Structure: THEOCHEM **388**, 339 (1996).