

# Basis-set convergence of correlated calculations on water

Cite as: J. Chem. Phys. **106**, 9639 (1997); <https://doi.org/10.1063/1.473863>

Submitted: 19 November 1996 • Accepted: 12 March 1997 • Published Online: 31 August 1998

Trygve Helgaker, Wim Klopper, Henrik Koch, et al.



View Online



Export Citation

## ARTICLES YOU MAY BE INTERESTED IN

[Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen](#)

The Journal of Chemical Physics **90**, 1007 (1989); <https://doi.org/10.1063/1.456153>

[A consistent and accurate ab initio parametrization of density functional dispersion correction \(DFT-D\) for the 94 elements H-Pu](#)

The Journal of Chemical Physics **132**, 154104 (2010); <https://doi.org/10.1063/1.3382344>

[Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions](#)

The Journal of Chemical Physics **96**, 6796 (1992); <https://doi.org/10.1063/1.462569>

**The Journal of Chemical Physics**

**Special Topics** Open for Submissions

Learn More

# Basis-set convergence of correlated calculations on water

Trygve Helgaker and Wim Klopper

*Department of Chemistry, University of Oslo, N-0315 Oslo, Norway*

Henrik Koch

*Department of Chemistry, Aarhus University, DK-8000 Aarhus C, Denmark*

Jozef Noga

*Institute of Inorganic Chemistry, Slovak Academy of Sciences, SK-84236 Bratislava, Slovakia*

(Received 19 November 1996; accepted 12 March 1997)

The basis-set convergence of the electronic correlation energy in the water molecule is investigated at the second-order Møller–Plesset level and at the coupled-cluster singles-and-doubles level with and without perturbative triples corrections applied. The basis-set limits of the correlation energy are established to within  $2 mE_h$  by means of (1) extrapolations from sequences of calculations using correlation-consistent basis sets and (2) from explicitly correlated calculations employing terms linear in the interelectronic distances  $r_{ij}$ . For the extrapolations to the basis-set limit of the correlation energies, fits of the form  $a + bX^{-3}$  (where  $X$  is two for double-zeta sets, three for triple-zeta sets, etc.) are found to be useful. CCSD(T) calculations involving as many as 492 atomic orbitals are reported. © 1997 American Institute of Physics. [S0021-9606(97)00823-4]

## I. INTRODUCTION

With the recent developments in computer technology and computational methods, it has become possible to explore in detail the basis-set convergence of correlated methods for molecular systems containing first-row atoms, allowing us to investigate the true convergence patterns and convergence rates towards the basis-set limit for the standard computational methods of molecular electronic ab initio theory. Information on basis-set convergence rates is crucial for any nonempirical computational scheme since it allows us to make extrapolations and estimates of error bars in calculations on large systems, where basis sets that would give results close to the basis-set limit cannot be employed.

In this paper, we present systematic calculations on the water molecule, investigating the basis-set convergence for Hartree-Fock (HF) theory, for perturbation theory at the second-order Møller–Plesset (MP2) level, and for coupled-cluster theory at two levels; the singles-and-doubles level (CCSD) and the singles-and-doubles level with perturbative corrections for triples [CCSD(T)]. At each level of theory, we carry out systematic calculations of energies using the correlation-consistent basis sets of Dunning and co-workers,<sup>1–4</sup> namely, the correlation-consistent polarized valence sets cc-pVXZ and the correlation-consistent polarized core-valence sets cc-pCVXZ.<sup>5</sup> These sets are well suited for our purposes, constituting hierarchical sequences of basis sets that allow us to approach the basis-set limits in controlled and systematic manner. However, even the largest correlation-consistent sets do not give results that are converged to the basis-set limit, in particular for the correlated methods. In order to establish these limits, additional calculations are required. For this purpose, we have employed explicitly correlated wave functions—that is, nonstandard wave functions that are particularly well suited for highly accurate calculations of dynamical correlation effects, containing terms linear in the interelectronic distances  $r_{ij}$ .

The extrapolation of energies towards the basis-set limit requires not only a systematic suite of basis sets such as the correlation-consistent sets, but also a simple model according to which the extrapolation can be carried out. For this purpose, Feller<sup>6</sup> has proposed the use of a three-parameter exponential model  $a + b \exp(-cX)$ , where  $X$  is the cardinal number of the basis set (i.e., two for cc-pVDZ and cc-pCVDZ, three for cc-pVTZ and cc-pCVTZ, etc.). However, although such a model may indeed give a reasonable fit to the calculated energies, it does not display the correct asymptotic behavior of the correlation energy and as such is inappropriate for extrapolations to the basis-set limit. Indeed, investigations of the asymptotic behavior of the second-order correlation energy—for the helium atom by Schwartz in 1962 using a  $1/Z$  perturbation expansion<sup>7</sup> and for  $n$ -electron atoms by Kutzelnigg and Morgan in 1992 using Møller–Plesset theory<sup>8</sup>—indicate that the energy contributions from atomic orbitals of angular-momentum quantum number  $l$  should converge asymptotically as  $(l + \frac{1}{2})^{-4}$ —that is, slower than the exponential convergence of Feller's model. For the correlation energies, we have therefore carried out extrapolations based on a two-parameter Schwartz model involving an inverse power of the cardinal number.

The extrapolation of electron correlation energies to the complete basis-set limit has been investigated by Petersson *et al.* within the framework of the complete-basis-set (CBS) model chemistry developed by these authors.<sup>9,10</sup> The CBS method uses the asymptotic forms of the pair-natural orbital expansions for the pair correlation energies and pair-coupling terms to extrapolate from finite basis sets to the basis-set limit. Furthermore, the higher-level correction (HLC) of the Gaussian-2 approach accounts for remaining basis-set deficiencies by adding empirical corrections (cf. Ref. 11), whereas the correlation energies obtained from finite basis sets are multiplied by an empirical parameter in the parametrized configuration-interaction (PCI-X) method of

TABLE I. Cartesian coordinates (in  $a_0$ ) for H<sub>2</sub>O.

Atom	$x$	$y$	$z$
O	0.000 000 00	0.000 000 00	0.000 000 00
H	1.430 432 09	0.000 000 00	1.107 159 59
H	-1.430 432 09	0.000 000 00	1.107 159 59

Siegbahn *et al.*<sup>12</sup> A variety of extrapolation schemes, including empirical corrections and inverse-power fits, have been studied by Martin.<sup>13</sup>

## II. RESULTS AND DISCUSSION

This section is organized as follows. After giving some technical details on the calculations in Sec. II A, we discuss in Sec. II B the explicitly correlated calculations. We then proceed to consider the conventional calculations, beginning with a discussion of basis sets in Sec. II C. Extrapolations of the Hartree–Fock energy using the exponential Feller model are discussed in Sec. II D and in Sec. II E we consider extrapolations of the correlated energies according to the Schwartz model. In Sec. II F we discuss calculations on the water molecule using basis sets larger than those available in the standard correlation-consistent sequences and consider in particular the convergence of the CCSD(T) triples corrections to the correlation energies.

### A. Computational details

In all calculations, the geometry of the water molecule was kept fixed at the experimental geometry adopted from Ref. 14, and all 10 electrons were correlated. The Cartesian coordinates are given in Table I, which correspond to internal coordinates of  $R_{\text{O-H}} = 1.808\,85\,a_0$  and  $\angle_{\text{H-O-H}} = 104.52^\circ$ . The explicitly correlated calculations at the MP2-R12, CCSD-R12, and CCSD(T)-R12 levels<sup>15–17</sup> were performed with the DIRCCR12-95 program<sup>18</sup> on IBM RS/6000 workstations (model 590 with 1 gigabyte of memory). The standard correlated calculations were carried out using the integral-direct techniques in Refs. 19–21 as implemented in the DALTON program.<sup>22</sup> These calculations were run on a Cray C92A supercomputer equipped with 1 gigabyte of memory.

### B. Explicitly correlated calculations

The MP2-R12, CCSD-R12, and CCSD(T)-R12 methods are similar to the standard methods MP2, CCSD, and CCSD(T) except that they employ explicitly correlated  $n$ -electron basis functions that depend linearly on the inter-electronic coordinates  $r_{ij}$ . These linear  $r_{ij}$  terms ensure a fast convergence of the correlation energies to the limiting value of a complete atomic-orbital (AO) basis set. We emphasize that, in the limit of a complete one-electron basis, the energies obtained using the MP2-R12, CCSD-R12, and CCSD(T)-R12 methods are identical to those obtained using the standard methods MP2, CCSD, and CCSD(T).

Four different AO basis sets were used for the correlated R12 calculations: The first basis (denoted as ‘I’ in Table II)

is the O:15s9p7d5f3g/H:9s7p5d3f subset of the O:15s9p7d5f3g1h/H:9s7p5d3f1g basis set (‘II’ in Table II) given in Ref. 14. The third basis is the cc-pV6Z basis of Dunning and Feller<sup>4</sup> in a fully uncontracted form, denoted in the present work as cc-pV6Z(uc) (the basis for O is available in the basis set library<sup>5</sup> under the name ‘pV6Z’). The fourth basis, cc-pV6Z(uc)+S, was obtained by adding steep functions to this uncontracted basis. On oxygen, two  $d$ -type functions with exponents 40.38 and 17.63, and one  $f$ -type function with exponent 14.44 were added; the basis on hydrogen was augmented with one  $p$ -type function with exponent 9.67. Augmentation of the ‘standard’ basis sets with steep functions improves the approximations inherent in the CCSD(T)-R12 method, based on inserting the approximate resolution of the identity in terms of the finite basis into certain three-, four-, and five-electron integrals.

The results obtained with these four basis sets are displayed in Table II. For the MP2 energies, two numbers are given since, at this level of theory, two different methods exist, designated as MP2-R12/A and MP2-R12/B. These two methods differ in the approximations used to compute some of the many-electron integrals, based either on the standard approximation A or on the standard approximation B—see Ref. 23 for details. It is important to note that the limiting values in a complete basis are the same for both approximations—that is, the complete-basis set MP2 value. The difference between MP2-R12/A and MP2-R12/B is only about  $1\,mE_h$ , indicating that these values are indeed already very close to the basis-set limit. It is our experience that for a molecule such as H<sub>2</sub>O the MP2-R12/A energy converges somewhat faster to the complete basis-set limit than does the energy obtained with approximation B. The MP2-R12/B energy usually converges from above to the limiting value, whereas the MP2-R12/A energy sometimes overshoots and, if so, approaches the limit from below. Overall, we estimate that the R12 values obtained with the cc-pV6Z(uc)+S basis set [i.e.,  $-76.428\,0\,E_h$  for MP2-R12/B,  $-76.427\,2\,E_h$  for CCSD-R12, and  $-76.437\,3\,E_h$  for CCSD(T)-R12] are accurate to within about  $2\,mE_h$ . Based on previous experience, we assume that these values represent upper bounds to the respective basis-set limits.

### C. Correlation-consistent basis sets

The standard calculations were carried out using the correlation-consistent basis sets of Dunning and co-workers—that is, the cc-pVXZ suite of basis sets designed to recover valence correlation energy, and the cc-pCVXZ suite designed to recover valence as well as core and core-valence correlation energies. These sets are particularly well suited for our purposes since they have been constructed to allow for a smooth convergence of the correlation energy towards the basis-set limit. Thus, with each extension of the basis set, new functions are added that make similar contributions to the correlation energy. For example, the 3s2p1d cc-pVDZ basis for the oxygen atom contains a set of atomic Hartree–Fock orbitals, augmented with one set of correlating functions for each occupied atomic orbital. Proceeding to the

TABLE II. Results for H<sub>2</sub>O.

Basis	<i>N</i> <sup>a</sup>	SCF	ΔMP2	ΔCCSD	ΔCCSD(T) <sup>b</sup>
cc-pVDZ	24	−76.026 799	−0.203 960	−0.213 284	−0.003 055
cc-pVTZ	58	−76.057 168	−0.275 076	−0.280 829	−0.007 770
cc-pVQZ	115	−76.064 835	−0.313 288	−0.317 004	−0.009 246
cc-pV5Z	201	−76.067 091	−0.328 772	−0.330 475	−0.009 852
cc-pV6Z	322	−76.067 401	−0.338 148	−0.338 802	−0.010 074
cc-pCVDZ	28	−76.027 204	−0.241 326	−0.251 750	−0.003 288
cc-pCVTZ	71	−76.057 358	−0.317 498	−0.324 186	−0.008 235
cc-pCVQZ	144	−76.064 948	−0.342 630	−0.346 496	−0.009 522
cc-pCV5Z	255	−76.067 105	−0.352 283	−0.353 947	−0.009 997
cc-pCV6Z <sup>c</sup>	412	−76.067 291 <sup>d</sup>	−0.356 330	−0.356 780	−0.010 161
cc-pCV5Z(uc)	272	−76.067 126	−0.352 393	−0.354 029	−0.010 000
cc-pCV5Z(uc)+ <i>sp</i>	304	−76.067 176	−0.352 434	−0.354 057	−0.010 001
cc-pCV5Z(uc)+ <i>spd</i>	334	−76.067 176	−0.352 579	−0.354 163	−0.010 003
cc-pCV5Z(uc)+ <i>spdf</i>	362	−76.067 176	−0.352 780	−0.354 324	−0.010 005
cc-pCV5Z(uc)+ <i>spdfg</i>	371	−76.067 176	−0.353 208	−0.354 642	−0.010 020
cc-pCV5Z(uc)+ <i>spdfghi</i>	435	−76.067 179	−0.355 020	−0.355 822	−0.010 074
cc-pCV5Z(uc)+ <i>spdfghi</i> +D	492	−76.067 401	−0.355 575	−0.356 175	−0.010 168
R12/B I	291	−76.067 345	−0.361 387	−0.359 841	−0.009 978
R12/A I			−0.362 149		
R12/B II	320	−76.067 350	−0.361 453	−0.359 945	−0.010 044
R12/A II			−0.362 005		
R12/B cc-pV6Z(uc)	351	−76.067 404	−0.359 889	−0.359 329	−0.010 072
R12/A cc-pV6Z(uc)			−0.361 062		
R12/B cc-pV6Z(uc)+S	374	−76.067 419	−0.360 595	−0.359 775	−0.010 116
R12/A cc-pV6Z(uc)+S			−0.361 486		
Reference 9			−0.360 47		
Reference 27			−0.356 36		

<sup>a</sup>Number of contracted Gaussian basis functions.<sup>b</sup>Triples contribution.<sup>c</sup>Constructed in the course of the present work.<sup>d</sup>One linear dependent function was removed from the basis.

4s3p2d1f cc-pVTZ basis, an additional set of correlating functions of each angular momentum is included as well as a set of higher angular-momentum functions. This process is continued for the larger basis sets, making sure that the correlating functions added at each level make similar contributions to the energy. The results obtained using the polarized valence basis sets cc-pVXZ with cardinal numbers X up to six are listed in Table II.

In the cc-pVXZ basis sets, only functions required for a correlated treatment of the valence electrons are included. For a balanced treatment of correlation effects of all electrons, the core-valence cc-pCVXZ basis sets have been developed. In these sets, additional steep (tight) correlating orbitals have been added for correlation in the core and core-valence regions. Thus, the 3s2p1d cc-pVDZ set is augmented with steep functions 1s1p to yield the cc-pCVDZ set of composition 4s3p1d, the 4s3p2d1f cc-pVTZ set is augmented with steep functions 2s2p1d to yield the 6s5p3d1f cc-pCVTZ set, and so on.

The cc-pCV6Z basis set is not available in the Extensible Computational Chemistry Environment Basis Set Database.<sup>5</sup> Since a basis set of this type was highly desired for our convergence study and basis-set limit extrapolation, we have attempted to construct a cc-pCV6Z-type set in the course of the present work. For oxygen, the cc-pCV6Z set was obtained from the cc-pV6Z basis in two steps. First, we added—without any modifications—the same set of

4s4p3d core-correlating orbitals that is added to the cc-pV5Z set in order to obtain the cc-pCV5Z set; next, the following sets of additional core-correlating orbitals were added: 1s (with exponent 390.0), 1p (500.0), 1d (300.0), 3f (116.0; 45.0; 17.354), 2g (86.9; 21.7), and 1h (60.0).

It is of some interest to note that, for the first-row atoms, the number of contracted functions in the correlation-consistent basis sets can be written in the following simple forms:

$$N_{\text{cc-pVXZ}} = \frac{1}{3} (X+1) \left( X + \frac{3}{2} \right) (X+2), \quad (1)$$

$$N_{\text{cc-pCVXZ}} = N_{\text{cc-pVXZ}} + N_{\text{cc-pV}(X-1)\text{Z}} - 1, \quad (2)$$

$$N_{\text{aug-cc-p(C)VXZ}} = N_{\text{cc-p(C)VXZ}} + (X+1)^2, \quad (3)$$

where we have also included the number of functions in the aug-cc-p(C)VXZ basis sets, obtained by augmenting the standard cc-p(C)VXZ sets with diffuse functions. Thus, the size of the correlation-consistent basis sets increases as the third power in the highest angular-momentum function.

## D. Hartree–Fock energies and extrapolations

When considering how the calculated energies converge with respect to the extension of the basis set, it is important to distinguish between the convergence of the Hartree–Fock energy and the convergence of the correlation

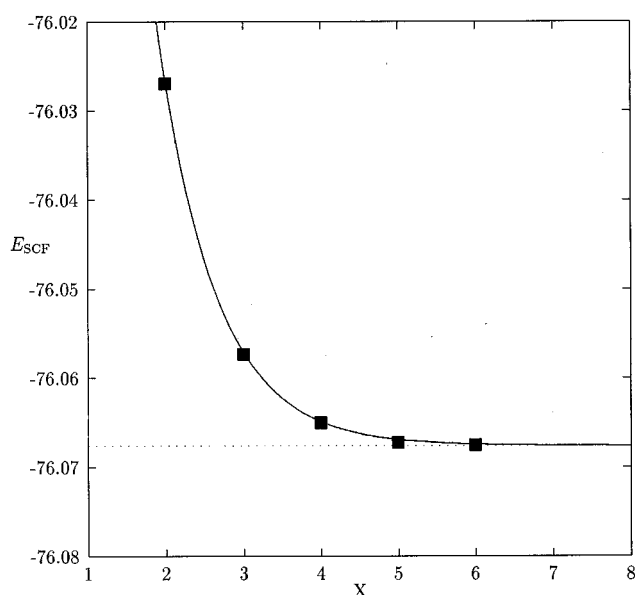


FIG. 1. Convergence of the SCF energy (in  $E_h$ ) for  $H_2O$  as a function of the cardinal number  $X=2,3,\dots,6$  in the series of basis sets cc-pVXZ. The solid line corresponds to the fit  $E_{SCF} = -76.067\,6 + 0.617\,161 \exp(-1.357\,71X)$ .

energy. The Hartree–Fock model allows for an exact treatment within a finite one-electron basis (with 5 orbitals sufficient for  $H_2O$ ), whereas the correlation energy cannot be treated exactly in any finite-dimensional basis. Consequently, it is reasonable to expect the convergence characteristics of these energy contributions to be different. The dominance of the Hartree–Fock energy over the correlation energy also makes a separate treatment of these contributions desirable.

In order to extrapolate the Hartree–Fock and correlation energies, we assume that simple analytical forms can be found that display the correct asymptotic behavior of the energies, and which allow for accurate fits of the calculated energies with only a few parameters (since only a few energies can be calculated). Dunning found that, in correlated calculations, the energy lowerings along the cc-pVXZ sequence decrease approximately geometrically. Indeed, this observation constitutes much of the justification for the correlation-consistent basis sets. Although the energy corrections in Hartree–Fock calculations are expected to be smaller than those obtained in correlated calculations, we may still at this uncorrelated level assume a geometrical progression for these basis sets. In Fig. 1, we have therefore plotted the cc-pVXZ energies for the water molecule together with an exponential fit of the form

$$E_{SCF} = a + b \exp(-cX), \quad (4)$$

as suggested by Feller.<sup>6</sup> This fit is seen to be an excellent one with errors of the order of  $0.1\,mE_h$  for all basis sets, indicating that the error in the SCF energy indeed decreases in a geometrical fashion for the correlation-consistent basis sets.

For the cc-pVXZ sets, the SCF error decreases by a factor of 3.9 with each increment in the cardinal number. A similar exercise for the core-valence sets cc-pCVXZ gives a plot indistinguishable from those of the valence set and also a similar fit, with basis-set limit of  $-76.067\,6\,E_h$  and a decrease by a factor of 3.9 with each increment in the cardinal number. As expected, there is nothing to be gained in the total energy by carrying out Hartree–Fock calculations in the cc-pCVXZ basis set, although molecular properties that depend critically on the electron distribution in the inner valence and core regions may be significantly affected.

## E. Correlation energies and extrapolations

Since the exponential fit was so successful for the Hartree–Fock energies and since, for the correlation energies, a geometrical progression has been observed by Dunning and co-workers, it would appear natural to use the same exponential fits for the correlation energies. Indeed, it appears that reasonable exponential fits can be made to the calculated correlation energies for the valence and core-valence calculations. At this point, we recall that there is no rigorous theoretical justification for the use of the exponential function for the extrapolation of the correlation-energy limit. On the contrary, the work of Schwarz indicates that the correlation energy should not converge exponentially, but as an inverse power in the highest angular-momentum function present in the basis set.<sup>7</sup> Investigating the second-order energy of the  $1/Z$  perturbation expansion for two-electron atoms, Schwarz showed that, asymptotically, the convergence with respect to the angular-momentum quantum number  $l$  contained in the one-particle basis goes as

$$\Delta E_l^{(2)} = -\frac{45}{256} \lambda^4 \left[ 1 - \frac{5}{4} \lambda^2 + \mathcal{O}(\lambda^4) \right], \quad (5)$$

where  $\lambda = (l + \frac{1}{2})^{-1}$ . A similar formula was found empirically by Carroll *et al.*<sup>24</sup> for configuration interaction calculations on the helium atom,

$$\Delta E_l = -0.074\lambda^4 - 0.031\lambda^5 + \mathcal{O}(\lambda^6), \quad (6)$$

later analyzed in detail by Hill.<sup>25</sup> Recently, atomic states other than the  $1S$  ground state have been studied by Kutzelnigg and Morgan,<sup>8</sup> who also considered the rates of convergence of  $n$ -electron atoms in second-order Møller–Plesset (MP2) perturbation-theory calculations.

Schwarz' work was based on a  $1/Z$  perturbation analysis of the wave function of the helium atom and applies to a situation where the basis sets at each level are completely saturated with respect to the radial part of the basis functions. We cannot apply these results directly to the cc-pVXZ and cc-pCVXZ calculations on the water molecule since these basis sets are constructed in a ‘‘correlation-consistent’’ manner—that is, convergence is obtained by adding functions of different angular momenta simultaneously if they give similar energy lowerings, rather than by saturating each angular-momentum space separately. Nevertheless, by analogy with the helium expansion, it is reasonable to expect that

TABLE III. Parameters of fits of the type  $E = a + bX^{-3}$  to the cc-p(C)VXZ results, based on data points for  $3 \leq X \leq 6$ .

Method	Basis	$a$	$b$
MP2	cc-pVXZ	-0.344 925	1.905 539
	cc-pCVXZ	-0.361 733	1.197 557
CCSD	cc-pVXZ	-0.345 448	1.756 453
	cc-pCVXZ	-0.361 902	1.013 900
CCSD(T)	cc-pVXZ	-0.355 847	1.827 728
	cc-pCVXZ	-0.372 359	1.073 822

the correlation energy in the water molecule should in the asymptotic limit converge at the best as an inverse power in the highest angular momentum.

For our extrapolations of the correlation energies, we have chosen to fit the calculated cc-p(C)VXZ energies at the triple-zeta and higher levels to an expression of the form

$$E_{\text{corr}} = a + bX^{-3}. \quad (7)$$

This functional form was chosen for its simplicity. It contains only two parameters and displays the same asymptotic behavior as the integrated Schwarz expansion for the helium atom,

$$\int_{L+1/2}^{\infty} (l + \frac{1}{2})^{-4} dl = \frac{1}{3}(L+1)^{-3}. \quad (8)$$

To arrive at the expression (7) for the correlation-consistent series, we have identified  $L+1$  in the Schwarz expansion (8) with the cardinal number of the cc-p(C)VXZ basis sets. Other forms similar to expression (7) such as

$$E_{\text{corr}} = a + b(X+c)^{-3+d}, \quad (9)$$

were investigated for fits to the correlation energies, without significant improvements. In particular, the form identical to expression (7) but with  $X+1$  rather than  $X$  [as suggested by the form (8)] gives a slightly less satisfactory fit than expression (7), with basis-set limit results too low relative to the R12 numbers. Note that, since the number of basis functions  $N$  grows with the third power of  $X$ , the inverse-power fit (7) is related to a rate of convergence with a leading term in  $1/N$ . A fit in powers of  $1/N$  was carried out in earlier work to the cc-pVXZ valence-shell second-order correlation energies.<sup>26</sup>

The Schwartz fits to the correlation-consistent energies are listed in Table III and (for the coupled-cluster calculations) plotted in Fig. 2. In these fits, the double-zeta energies have not been included since these basis sets are too small to give any information about the asymptotic convergence. The basis-set limits that may be extracted from the Schwartz fits to the cc-pCVXZ sets are close to the R12 results, the differences being only 1.1, 2.1, and 2.5  $mE_h$  for MP2, CCSD, and CCSD(T), respectively. Inclusion of the double-zeta sets in the fits would give results in worse agreement both with the calculated numbers and with the basis-set limits. Further justification for leaving out the double-zeta values may be obtained from CI calculations on the helium atom, which show that the expected asymptotic convergence with respect

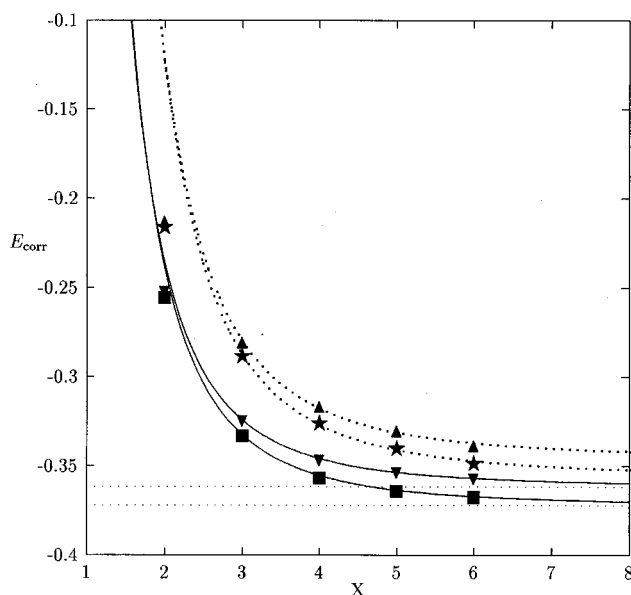


FIG. 2. Convergence of the total CCSD ( $\blacktriangle$ ;  $\blacktriangledown$ ) and CCSD(T) ( $\star$ ;  $\blacksquare$ ) correlation energies (in  $E_h$ ) for  $H_2O$  as a function of the cardinal number  $X = 2, 3, \dots, 6$  in the series of basis sets cc-pVXZ (dashed lines) and cc-pCVXZ (solid lines). The curves represent linear least squares fits of the type  $E_{\text{corr}} = a + bX^{-3}$ , with the double-zeta data point excluded from the fit.

to the maximum angular-momentum quantum number is not observed for the first few terms in the expansion. The optimum strategy is certainly to use the most extended basis sets (with the highest cardinal numbers) for the Schwartz fits.

In Table IV, we list the parameters of the exponential fits to the MP2, CCSD, and CCSD(T) correlation energies. In order to compare the fits with the Schwartz fits, we excluded the double-zeta values from the exponential fits also. For the cc-pCVXZ basis sets, the basis-set limits from the exponential fits are about 3–4  $mE_h$  higher than the limits obtained from the Schwartz fits. This is understandable, because the exponential fits decay more quickly than the Schwartz fits. Furthermore, the basis-set limits from the exponential fits are higher than the corresponding R12 values by roughly 2  $mE_h$ . If we assume that the R12 values represent upper bounds to the correlation energies, we must conclude that the basis-set limits obtained from the exponential fits are too high by at least 2  $mE_h$ . For this reason, we believe that the limits obtained from the Schwartz fits are closer to the true

TABLE IV. Parameters of fits of the type  $E = a + b \exp(-cX)$  to the cc-p(C)VXZ results based on data points for  $3 \leq X \leq 6$ .

Method	Basis	$a$	$b$	$c$
MP2	cc-pVXZ	-0.344 169	0.725 570	0.784 376
	cc-pCVXZ	-0.358 757	0.682 175	0.935 222
CCSD	cc-pVXZ	-0.343 160	0.778 586	0.842 282
	cc-pCVXZ	-0.358 078	0.834 846	1.068 110
CCSD(T)	cc-pVXZ	-0.353 388	0.816 685	0.845 285
	cc-pCVXZ	-0.368 338	0.876 152	1.064 860

TABLE V. Exponents of Gaussian basis functions for H<sub>2</sub>O.

O atom						
	<i>s</i>	<i>p</i>	<i>d</i>	<i>f</i>	<i>g</i>	<i>h</i>
1	492 600.00	2 042.452 3	1 075.749 0	935.059 70	48.578 000	2.319 000 0
2	164 200.00	934.293 90	379.742 20	247.567 00	9.438 600 0	0.823 000 0
3	24 590.000	427.380 00	134.050 00	65.546 000	3.350 000 0	
4	5 592.000 0	225.630 00	47.320 000	17.354 000	1.189 000 0	
5	1 582.000 0	195.500 00	16.703 000	4.016 000 0	0.422 000 0	
6	516.100 00	89.429 000	5.879 000 0	1.554 000 0		
7	187.200 00	46.160 000	2.307 000 0	0.601 000 0		
8	82.259 000	35.446 000	0.905 000 0	0.232 400 0		
9	73.930 000	14.580 000	0.355 000 0			
10	31.220 000	5.296 000 0	0.139 300 0			
11	13.810 000	2.094 000 0				
12	6.256 000 0	0.847 100 0				
13	2.776 000 0	0.336 800 0				
14	1.138 000 0	0.128 500 0				
15	0.460 000 0	0.049 000 0				
16	0.182 900 0					
17	0.072 700 0					

H atom					
	<i>s</i>	<i>p</i>	<i>d</i>	<i>f</i>	<i>g</i>
1	1 159.609 0	82.890 500	17.651 100	7.177 000 0	2.358 000 0
2	402.000 00	31.423 500	7.216 000 0	2.506 000 0	0.823 000 0
3	264.300 00	11.912 500	2.950 000 0	0.875 000 0	
4	60.240 000	4.516 000 0	1.206 000 0	0.305 500 0	
5	13.730 000	1.712 000 0	0.493 000 0		
6	3.905 000 0	0.649 000 0	0.201 500 0		
7	1.283 000 0	0.246 000 0			
8	0.465 500 0	0.093 200 0			
9	0.181 100 0				
10	0.072 790 0				
11	0.029 250 0				

basis-set limits than those obtained from the exponential fits. Finally, we note that the Schwartz fits correspond to linear least squares fits with only two parameters, while the exponential fits represent more complicated nonlinear fits with three parameters. The fact that the simple two-parameter linear fits perform so well indicates that the underlying model appears to be very useful.

F. Larger calculations and the convergence of the triples corrections

In order to investigate the convergence of the energy for basis sets larger than those present in the standard correlation-consistent sets, some additional calculations were carried out using special basis sets. First, the cc-pCV5Z basis was uncontracted and trimmed to remove linear dependencies within the primitive set. The resulting cc-pCV5Z(uc) basis was then used as a starting point for the construction of a sequence of even larger sets, designed to explore basis-set convergence more thoroughly. The exponents of these basis sets are given in Tables V and VI.

The first basis sets generated from the cc-pCV5Z(uc) set were intended to explore basis-set saturation in the inner regions. Thus, we generated the cc-pCV5Z(uc)+*sp* basis by adding steep *s* and *p* functions to both oxygen and hydrogen, with exponents obtained from a geometrical progression.

Next, the cc-pCV5Z(uc)+*spd* basis was obtained by adding two sets of steep *d* functions to both oxygen and hydrogen. Adding two sets of steep *f* functions to oxygen and one set of steep *f* functions to hydrogen, we then obtained the

TABLE VI. Construction of Gaussian basis sets for H<sub>2</sub>O.

O atom						
	<i>s</i>	<i>p</i>	<i>d</i>	<i>f</i>	<i>g</i>	<i>h</i>
cc-pCV5(uc)	2–16	4–14	3–9	3–7	1,3–4	1
+ <i>sp</i>	1	1–3				
+ <i>spd</i>			1–2			
+ <i>spdf</i>				1–2		
+ <i>spdfg</i>					2	
+ <i>spdfghi</i>						2
+D	17	15	10	8	5	1

H atom						
	<i>s</i>	<i>p</i>	<i>d</i>	<i>f</i>	<i>g</i>	<i>h</i>
cc-pCV5(uc)	2–10	4–7	3–5	2–3	1	
+ <i>sp</i>	1	1–3				
+ <i>spd</i>			1–2			
+ <i>spdf</i>				1		
+ <i>spdfg</i>						
+ <i>spdfghi</i>					2	1
+D	11	8	6	4		



TABLE VII. Estimated complete-basis-set energies (in  $E_h$ ) for  $H_2O$ .

Method	cc-pCV5Z(uc)+ <i>spdfghi</i> +D	cc-pCV6Z <sup>a</sup>	R12 result <sup>b</sup>	Best estimate <sup>c</sup>
SCF	−76.067 4	−76.067 3	−76.067 4 <sup>d</sup>	−76.067 6±0.000 2
MP2	−76.423 0	−76.423 6	−76.428 0 <sup>e</sup>	−76.429±0.002
CCSD	−76.423 6	−76.424 1	−76.427 2	−76.429±0.002
CCSD(T)	−76.433 7	−76.434 2	−76.437 3	−76.439±0.002
“Experimental”				−76.440 <sup>f</sup>

<sup>a</sup>Constructed in the course of the present work.<sup>b</sup>Obtained with the cc-pV6Z(uc)+S basis set.<sup>c</sup>Error bars estimated by the authors, based on comparisons of different fits and comparisons with R12 calculations.<sup>d</sup>SCF energy in the basis of the R12 calculation.<sup>e</sup>MP2-R12/B result.<sup>f</sup>Nonrelativistic total energy of  $H_2O$  estimated in Ref. 28.

cc-pCV5Z(uc)+*spdf* basis. In the cc-pCV5Z(uc)+*spdfg* basis, one set of *g* functions was added to oxygen only.

To explore the effects of atomic functions with angular momentum higher than that of those present in the original core-valence cc-pCV5Z basis, we constructed the cc-pCV5Z(uc)+*spdfghi* basis by introducing *h* and *i* functions on oxygen and *g* and *h* functions on hydrogen. The exponents for the oxygen *i* functions and for the hydrogen *h* functions were taken from the cc-pV6Z basis. Finally, the largest basis used in this study was obtained by augmenting the cc-pCV5Z(uc)+*spdfghi* set with sets of diffuse *s*, *p*, *d*, *f*, and *g* functions on oxygen and with diffuse *s*, *p*, *d*, and *f* functions on hydrogen, with exponents obtained by geometrical progression. The resulting basis set is designated cc-pCV5Z(uc)+*spdfghi*+D and contains 492 basis functions. No linear-dependency problems were encountered with this basis.

Interesting observations can be made from the results in Table II for the larger basis sets. First, we note that the CCSD(T) perturbative triples correction is converged to less than  $0.1 mE_h$  at the cc-pCV5Z level and thus nothing is gained by augmenting the basis with steep and diffuse functions. On the other hand, the MP2 and CCSD results are rather sensitive. By uncontracting the cc-pCV5Z basis, we observe changes in the correlation energies of  $-0.11 mE_h$  for MP2 and  $-0.08 mE_h$  for CCSD. Further augmentation of the basis with steep *spdf* functions changes the energy by  $-0.4 mE_h$  for MP2 and  $-0.3 mE_h$  for CCSD. A similar change is observed when a single set of *g* functions is added to the oxygen atom. The introduction of higher angular-momentum functions changes the MP2 energy by  $-1.8 mE_h$  and the CCSD energy by  $-1.1 mE_h$ . The final addition of diffuse functions also gives large changes, even in the Hartree–Fock energy, which is stable with respect to the other extensions of the basis set.

Noting that the CCSD(T) triples correction converges faster than the MP2 and CCSD energies to an error less than  $1 mE_h$ , for example, we are led to suggest an algorithm where the triples are calculated using a smaller basis set of virtual orbitals than the basis used for the CCSD calculation. The proposed algorithm consists of orthogonalizing the new (smaller) virtual space against the occupied space, diagonal-

izing the Fock matrix in the virtual space to obtain new canonical orbitals and orbital energies, and projecting the CCSD amplitudes in terms of the larger virtual basis onto the smaller space. The virtual space can, for example, be reduced by removing certain AOs (e.g., high angular-momentum functions) or natural orbitals with occupation numbers below some prescribed threshold. Some years ago, Adamowicz and Bartlett investigated effective reductions of the virtual space in the step from MP2 to CCSD theory.<sup>29,30</sup> Here, we propose to reduce the virtual space going from CCSD to CCSD(T). Work along these lines is in progress.

### III. CONCLUSIONS

We have presented large-scale calculations on the water molecule, establishing—by means of extrapolations and explicitly correlated calculations—the basis-set limits of the correlation energy at the MP2, CCSD, and CCSD(T) levels to within  $2 mE_h$ . A summary of the results is provided in Table VII. Our best estimates for the total energies are  $-76.068 E_h$  for the SCF energy,  $-76.429 E_h$  for the MP2 energy,  $-76.429 E_h$  for the CCSD energy, and  $-76.439 E_h$  for the CCSD(T) energy.

*Note added in proof.* Upon acceptance of this paper, we were made aware of a similar work carried out at the SCF and MP2 levels by D. Moncrieff and S. Wilson [J. Phys. B **29**, 6009 (1996)]. Their results are in good agreement with ours.

### ACKNOWLEDGMENTS

The authors would like to thank Poul Jørgensen and Jeppe Olsen for discussions. J.N. is obliged to the Slovak grant agency VEGA (project No. 2-501) for the partial support of this work. The research by W.K. was carried out in part at the ETH Zürich under Research Grant No. TH-19/94-1. A generous allocation of CPU time by the ETH Zürich Competence Center for Computational Chemistry is gratefully acknowledged. This work has further been partly supported by the Danish Natural Science Research Council (Grant No. 11-0924) and the Norwegian Research Council (Grant No. 111182/410).



- <sup>1</sup>T. H. Dunning, Jr., *J. Chem. Phys.* **90**, 1007 (1989).
- <sup>2</sup>R. A. Kendall, T. H. Dunning, Jr., and R. J. Harrison, *J. Chem. Phys.* **96**, 6796 (1992).
- <sup>3</sup>D. E. Woon and T. H. Dunning, Jr., *J. Chem. Phys.* **98**, 1358 (1993); **100**, 2975 (1994); **103**, 4572 (1995).
- <sup>4</sup>T. H. Dunning, Jr. and D. Feller (unpublished).
- <sup>5</sup>The cc-p(C) VXZ basis sets (X=D,T,Q,5,6) were obtained from the Extensible Computational Chemistry Environment Basis Set Database, Version 1.0, as developed and distributed by the Molecular Science Computing Facility, Environmental and Molecular Sciences Laboratory which is part of the Pacific Northwest Laboratory, P.O. Box 999, Richland, Washington 99352, USA, and funded by the U.S. Department of Energy. The Pacific Northwest Laboratory is a multiprogram laboratory operated by Battelle Memorial Institute for the U.S. Department of Energy under Contract No. DE-AC06-76RLO 1830. Contact David Feller, Karen Schuchardt, or Don Jones for further information.
- <sup>6</sup>D. Feller, *J. Chem. Phys.* **96**, 6104 (1992); **98**, 7059 (1993).
- <sup>7</sup>C. Schwartz, *Phys. Rev.* **126**, 1015 (1962); *Meth. Comp. Phys.* **2**, 241 (1963).
- <sup>8</sup>W. Kutzelnigg and J. D. Morgan III, *J. Chem. Phys.* **96**, 4484 (1992); **97**, 8821(E) (1992).
- <sup>9</sup>G. A. Petersson and M. Braunstein, *J. Chem. Phys.* **83**, 5129 (1985); G. A. Petersson, A. Bennett, T. G. Tensfeldt, M. A. Al-Laham, W. A. Shirley, and J. Mantzaris, *ibid.* **89**, 2193 (1988).
- <sup>10</sup>J. W. Ochterski, G. A. Petersson, and J. A. Montgomery, Jr., *J. Am. Chem. Soc.* **117**, 11 299 (1995).
- <sup>11</sup>K. Raghavachari and L. A. Curtiss, in *Modern Electronic Structure Theory, Part II*, edited by D. R. Yarkony (World Scientific, Singapore, 1995), p. 991, and references therein.
- <sup>12</sup>P. E. M. Siegbahn, M. R. A. Blomberg, and M. Svensson, *Chem. Phys. Lett.* **223**, 35 (1994); P. E. M. Siegbahn, M. Svensson, and P. J. E. Bousard, *J. Chem. Phys.* **102**, 5377 (1995).
- <sup>13</sup>J. M. L. Martin, *Chem. Phys. Lett.* **259**, 669 (1996).
- <sup>14</sup>W. Klopper, *J. Chem. Phys.* **102**, 6168 (1995).
- <sup>15</sup>J. Noga, W. Klopper, and W. Kutzelnigg, *Chem. Phys. Lett.* **199**, 497 (1992).
- <sup>16</sup>J. Noga and W. Kutzelnigg, *J. Chem. Phys.* **101**, 7738 (1994).
- <sup>17</sup>J. Noga, D. Tunega, W. Klopper, and W. Kutzelnigg, *J. Chem. Phys.* **103**, 309 (1995).
- <sup>18</sup>DIRCCR12-95 program, written by J. Noga and W. Klopper; see also, W. Klopper and J. Noga, *J. Chem. Phys.* **103**, 6127 (1995).
- <sup>19</sup>H. Koch, O. Christiansen, R. Kobayashi, P. Jørgensen, and T. Helgaker, *Chem. Phys. Lett.* **228**, 233 (1994).
- <sup>20</sup>H. Koch, A. Sanchez de Merás, T. Helgaker, and O. Christiansen, *J. Chem. Phys.* **104**, 4157 (1996).
- <sup>21</sup>H. Koch, P. Jørgensen, and T. Helgaker, *J. Chem. Phys.* **104**, 9528 (1996).
- <sup>22</sup>T. Helgaker, H. J. Aa. Jensen, P. Jørgensen, H. Koch, J. Olsen, H. Ågren, T. Andersen, K. L. Bak, V. Bakken, O. Christiansen, P. Dahle, E. K. Dalskov, T. Enevoldsen, A. Halkier, H. Heiberg, D. Jonsson, S. Kirpekar, R. Kobayashi, A. S. de Merás, K. V. Mikkelsen, P. Norman, M. J. Packer, K. Ruud, T. Saue, P. R. Taylor, and O. Vahtras. DALTON, an *ab initio* electronic structure program,
- <sup>23</sup>W. Kutzelnigg and W. Klopper, *J. Chem. Phys.* **94**, 1985 (1991).
- <sup>24</sup>D. P. Carroll, H. J. Silverstone, and R. M. Metzger, *J. Chem. Phys.* **71**, 4142 (1979).
- <sup>25</sup>R. N. Hill, *J. Chem. Phys.* **83**, 1173 (1985).
- <sup>26</sup>W. Klopper, M. Schütz, H. P. Lüthi, and S. Leutwyler, *J. Chem. Phys.* **103**, 1085 (1995).
- <sup>27</sup>R. Bukowski, B. Jeziorski, R. Rybak, and K. Szalewicz, *J. Chem. Phys.* **102**, 888 (1995).
- <sup>28</sup>D. Feller, C. M. Boyle, and E. R. Davidson, *J. Chem. Phys.* **86**, 3424 (1987).
- <sup>29</sup>L. Adamowicz and R. J. Bartlett, *J. Chem. Phys.* **86**, 6314 (1987).
- <sup>30</sup>L. Adamowicz and R. J. Bartlett, *Phys. Rev. A* **37**, 1 (1988).