

TOPICAL REVIEW

Highly accurate calculations of molecular electronic structure

Wim Klopper^{†¶}, Keld L Bak[‡], Poul Jørgensen[§], Jeppe Olsen[§] and Trygve Helgaker^{||}

[†] Theoretical Chemistry Group, Debye Institute, Utrecht University, Padualaan 14, NL-3584 CH Utrecht, The Netherlands

[‡] UNI-C, Olof Palmes Allé 38, DK-8200 Århus N, Denmark

[§] Department of Chemistry, Århus University, DK-8000 Århus C, Denmark

^{||} Department of Chemistry, University of Oslo, PO Box 1033, N-0315 Oslo, Norway

Received 1 March 1999

Abstract. The highly accurate calculation of molecular electronic structure requires the expansion of the molecular electronic wavefunction to be as nearly complete as possible both in one- and n -electron space. In this review, we consider the convergence behaviour of computed electronic energies, in particular electronic enthalpies of reaction, as a function of the one-electron space. Based on the convergence behaviour, extrapolations to the limit of a complete one-electron basis are possible and such extrapolations are compared with the direct computation of electronic energies near the basis-set limit by means of explicitly correlated methods. The most elaborate and accurate computations are put into perspective with respect to standard and—from a computational point of view—inexpensive density functional, complete basis set (CBS) and Gaussian-2 calculations. Using the explicitly correlated coupled-cluster method including singles, doubles and non-iterative triples replacements, it is possible to compute (the electronic part of) enthalpies of reaction accurate to within 1 kJ mol⁻¹. To achieve this level of accuracy with standard coupled-cluster methods, large basis sets or extrapolations to the basis-set limit are necessary to exploit fully the intrinsic accuracy of the coupled-cluster methods.

Abbreviations

ANO	Atomic natural orbital
AO	Atomic orbital
APNO	Atomic pair natural orbital
B3LYP	Becke3–Lee–Yang–Parr hybrid DFT functional
CBS	Complete basis set
CC	Coupled cluster
CCD	CC with double excitations
CCD-R12	CCD with linear r_{12} terms
CCSD(T)	CC with singles and doubles and non-iterative triple excitations
CCSD(T)-R12	CCSD(T) with linear r_{12} terms
cc-pVXZ	correlation-consistent polarized valence X -tuple zeta
cc-pCVXZ	correlation-consistent polarized core–valence X -tuple zeta
CI	Configuration interaction
DFT	Density functional theory
DT	Double-triple extrapolation

¶ Author to whom correspondence should be addressed.

FC	Frozen-core approximation
FCI	Full configuration interaction
Full	All electrons correlated
G1	Gaussian-1 model chemistry
G2	Gaussian-2 model chemistry
G3	Gaussian-3 model chemistry
HLC	Higher-level correction
MP2	Second-order Møller–Plesset perturbation theory
MP2-R12	MP2 with linear r_{12} terms
MP4	Fourth-order Møller–Plesset perturbation theory
NO	Natural orbital
QCI	Quadratic CI
QCISD(T)	QCI with singles and doubles and non-iterative triple excitations
SAPT	Symmetry-adapted perturbation theory
SCF	Self-consistent field
TQ	Triple-quadruple extrapolation
TZV	Triple-zeta valence basis set
VTZ	Valence triple-zeta basis set

1. Introduction

It is well known that the expansion of the electronic wavefunction in orbital products (that is, in Slater determinants) converges frustratingly slowly towards the limit of a complete basis set. This slow convergence is a serious bottleneck for highly accurate calculations of the electronic structure of molecules as the computational costs grow much faster than the rate at which the accuracy is improved. Often, the error arising from the truncation of the atomic orbital (AO) basis is severe and the calculations needed to provide the desired accuracy may not be tractable on present-day computers. As a rule of thumb, the computational costs grow as the fourth power in the reduction of the basis-set truncation error. Thus, to gain one more decimal place in accuracy—that is, to reduce the error of the computed energy by a factor of ten—the computational effort must be increased by a factor of 10 000.

It is obvious that there is a need for solutions to the slow convergence problem. Essentially, there are two approaches that one may take to solve this problem. The first approach is concerned with extrapolations. The standard technique for accelerating the convergence of an infinite series is to develop an extrapolation based on the asymptotic form of the series. In this review, we shall discuss the most important such techniques used today—that is, extrapolations based on the partial-wave expansion, the natural orbital expansion and the principal expansion.

In the second approach, one searches for (affordable) alternatives to the expansion in Slater determinants, using many-electron basis functions that depend explicitly on the coordinates of two electrons. Such ‘explicitly correlated’ methods are also discussed in this review, which is concluded with a comparison of the different approaches and computational techniques as applied to the computation of the total electronic ground-state energies of the 20 closed-shell molecules CH_2 ($\tilde{\alpha}^1\text{A}_1$ state), CH_4 , NH_3 , H_2O , HF , C_2H_2 , C_2H_4 , HCN , HNC , N_2 , N_2H_2 , CO , H_2CO , HNO , H_2O_2 , HOF , F_2 , CO_2 , O_3 and H_2 .

2. Errors in electronic structure calculations

Often, the truncation of the AO basis set is the most important source of error in electronic structure calculations. Nevertheless, we must also be concerned with other errors.

As well as the one-electron space, the n -electron space of all Slater determinants that constitute the full configuration-interaction (FCI) wavefunction is also truncated. Truncation errors occur when this FCI wavefunction is approximated by a low-order perturbation theory approach, a truncated coupled-cluster expansion, or a truncated configuration-interaction wavefunction.

In order to obtain highly accurate results, we must advance the electronic structure calculations as far as technically possible towards the limits of both a complete AO basis set *and* full configuration interaction. We may refer to this combined complete-basis-set/full-configuration-interaction limit as the focal point of electronic structure calculations (Császár *et al* 1998, Tarczay *et al* 1999).

Furthermore, most present-day electronic structure calculations are concerned with the non-relativistic Schrödinger equation within the framework of the Born–Oppenheimer approximation. Under certain circumstances, however, the restriction to a non-relativistic theory or to clamped nuclei may lead to noticeable errors, and in such cases, it becomes necessary to account for relativistic or non-Born–Oppenheimer corrections (Császár *et al* 1998, Tarczay *et al* 1999), or to switch to a fully relativistic or non-adiabatic treatment.

The focus of the present review is solely on the AO basis-set truncation error, which is illustrated by computations on small closed-shell molecules containing first-row atoms. For these calculations, the errors due to the non-relativistic and Born–Oppenheimer frameworks are of minor importance.

3. Partial-wave expansion for atoms

In the early 1960s (Schwartz 1962, 1963), it was found that, for the second-order energy of the $1/Z$ perturbation expansion of the ground state of two-electron atoms, the asymptotic formula (i.e. as ℓ and Z approach infinity)

$$\Delta E_\ell^{(2)} = -\frac{45}{256}(\ell + \frac{1}{2})^{-4} + \frac{225}{1024}(\ell + \frac{1}{2})^{-6} + \dots \quad (1)$$

represents the energy increments obtained by adding a saturated shell of AO basis functions of angular momentum ℓ to the AO basis set used to expand the first-order wavefunction (Kutzelnigg and Morgan 1992a, b). From these increments, we can compute the total basis-set truncation error due to the omission of all basis functions of $\ell > L$. In the limit where L approaches infinity, this error can be expressed as

$$\delta E_L^{(2)} = -\sum_{\ell=L+1}^{\infty} \Delta E_\ell^{(2)} \approx -\int_{\ell=L+\frac{1}{2}}^{\infty} \Delta E_\ell^{(2)} = \frac{15}{256}(L+1)^{-3} - \frac{45}{1024}(L+1)^{-5} + \dots \quad (2)$$

We note that, whereas the leading term to the *energy increment* is of the order ℓ^{-4} , the leading term to the *basis-set truncation error* is of the order L^{-3} . However, this asymptotic formula applies to the 1S ground state of the He isoelectronic series; for other electron pairs, increments of the order ℓ^{-6} (triplet pairs) or ℓ^{-8} (pairs with unnatural parity) occur (Kutzelnigg and Morgan 1992a, b). Another restriction is that (1) applies to the somewhat special $1/Z$ perturbation theory; that is, a perturbation theory based on hydrogenic zero-order Hamiltonians and wavefunctions. For more general electronic-structure methods such as Møller–Plesset perturbation theory, coupled-cluster theory, or configuration–interaction (CI) wavefunctions,

the odd terms contribute to the energy increments as well (Carroll 1979, Hill 1985). Thus, CI calculations of the He ground state converge as

$$\Delta E_\ell = -0.074\,226 \left(\ell + \frac{1}{2}\right)^{-4} - 0.030\,989 \left(\ell + \frac{1}{2}\right)^{-5} + \dots \quad (3)$$

$$\delta E_L = 0.024\,742 (L+1)^{-3} + 0.007\,747 (L+1)^{-4} + \dots \quad (4)$$

For atoms, it is common practice to extrapolate (using the known asymptotic formulae) the results obtained from finite partial-wave expansions to the limit of a complete AO basis (Byron and Joachain 1967, Sasaki and Yoshimine 1974, Lindgren and Salomonson 1980, Jankowski and Malinowski 1980, Termath *et al* 1991, Mårtensson-Pendrill *et al* 1991, Flores 1992, Flores and Redondo 1993).

For molecules, however, it seems that the partial-wave formulae cannot be applied. First, for molecules, the angular momentum is not a good quantum number. Secondly, molecular AO basis sets are usually not constructed in such a manner that function spaces of a given (atomic) angular-momentum quantum number are saturated before the next function space is added. In the next three sections, therefore, we discuss alternative expansions for molecules.

Nevertheless, the partial-wave formulae have been applied successfully to polyatomic systems as well (Martin 1996, 1997a–c, Martin and Taylor 1997, Wilson and Dunning 1997, Helgaker *et al* 1997b). When correlation-consistent basis sets of the type cc-pVXZ (Dunning 1989) are used, the highest angular momentum in the basis is $L = X - 1$ for H and He and $L = X$ for Li–Ar. Therefore, Martin (1996, 1997b) suggested to take the average value of $\bar{L} = X - \frac{1}{2}$; that is, to replace L by $X - \frac{1}{2}$ in (4) for calculations on molecules containing both hydrogen and non-hydrogen atoms, and to fit functions such as

$$E(X) \approx E(\infty) + a \left(X + \frac{1}{2}\right)^{-4} \quad (5)$$

$$E(X) \approx E(\infty) + b \left(X + \frac{1}{2}\right)^{-4} + c \left(X + \frac{1}{2}\right)^{-6} \quad (6)$$

$$E(X) \approx E(\infty) + d \left(X + \frac{1}{2}\right)^{-e} \quad (7)$$

to the molecular correlation energies obtained with cc-pVXZ basis sets. In (5) and (6), the X^{-3} and X^{-5} terms are omitted as it was found during the fitting that the X^{-4} term dominates the truncation error, at least for the small X used in the study. For example, for the trial function (7), optimal values were found in the range $3.5 < e < 4.5$ (Martin and Taylor 1997).

Recently, Halkier *et al* (1998) found that fits of the form

$$E(X) \approx E(\infty) + a(X + \delta)^{-\alpha} \quad (8)$$

all perform similarly for $\alpha \approx \delta + 3$. Of course, since a Taylor series expansion around $\delta = 0$ yields

$$(X + \delta)^{-k} = X^{-k} - k\delta X^{-(k+1)} + O(\delta^2) \quad (9)$$

the choice of δ becomes less important as more terms are included in the fitting procedure. A non-zero δ merely introduces higher-order terms. Thus, the simplest and most transparent approach is probably to set $\delta = 0$ and use a polynomial to prescribed order in $1/X$. In this spirit, Helgaker *et al* (1997b) advocated a simple two-point linear fit for the correlation energy based on the form

$$E(X) \approx E(\infty) + aX^{-3}. \quad (10)$$

With correlation energies available for the basis set $X = A$, we can easily perform a calculation with the next smaller set ($X = A - 1$) and carry out an extrapolation according to (Halkier *et al* 1998)

$$E(\infty) \approx \frac{E(A)A^3 - E(A-1)(A-1)^3}{A^3 - (A-1)^3}. \quad (11)$$

We shall return to the convergence behaviour of correlation-consistent basis sets and related extrapolation schemes in section 5.

4. Natural orbital expansion and complete basis-set (CBS) extrapolation

In recent years, the CBS extrapolation has become a standard technique for accurate computations of molecular electronic energies. Various versions of this scheme, for example, have been implemented in the commercial computer program Gaussian 94 (Frisch *et al* 1995) and thus become available to the majority of computational chemists. Comprehensive reviews of the CBS methods and assessments of their performance have been published recently (Ochterski *et al* 1995, 1996, Petersson *et al* 1998). The interested reader is referred to these publications for details; in this section, we shall examine only the key ideas of the CBS extrapolation scheme and how it is related to other methods. We shall restrict our discussion to the most accurate member of the family; namely, the CBS-QCI/APNO (complete-basis-set-quadratic configuration-interaction/atomic-pair-natural-orbitals) method, which starts from standard QCISD(T) calculations (table 1, Pople *et al* 1987, Raghavachari *et al* 1989). Being the most accurate member, it is also the most expensive one and can be applied only to molecules containing (a few) first-row atoms. CBS methods for larger molecules and heavier atoms exist (Ochterski *et al* 1996), but in view of our interest in highly accurate methods, we shall not consider these variants of CBS here.

In 1981, Nyden and Petersson realized that, when developing extrapolation schemes for calculations on molecules, explicit reference to angular momentum must be avoided. Therefore, they studied the convergence of natural orbital (NO) expansions, defined for molecules as for atoms.

For the ground state of the He atom, the AO set truncation error is inversely proportional to the number of NOs included in the full CI (FCI) wavefunction, when the orbitals are ordered according to monotonically decreasing occupation numbers (Nyden and Petersson 1981, Petersson and Nyden 1981). In particular, it was found that the FCI energy computed from the first N NOs is well represented by the formula

$$E(N) \approx E(\infty) + \left(\sum_{\mu=1}^N C_{\mu} \right)^2 \frac{25}{512} (N + \delta)^{-1} \quad (12)$$

where, to fit the energies for small N , an empirical parameter ($\delta = 0.363$) was introduced. Note that δ does not alter the limit as N approaches infinity. The C_{μ} are the coefficients of the NO expansion of the singlet ground state

$${}^1\Psi_N(1, 2) = {}^1\Theta(1, 2) \sum_{\mu=1}^N C_{\mu} \varphi_{\mu}(1) \varphi_{\mu}(2) \quad (13)$$

$${}^1\Theta(1, 2) = \frac{1}{\sqrt{2}} \{ \alpha(1) \beta(2) - \beta(1) \alpha(2) \} \quad (14)$$

where φ_{μ} is an NO and ${}^1\Theta(1, 2)$ the usual two-electron singlet spinfunction. It turns out, however, that (12) is correct only for certain ‘magic’ N ; namely, when shells of functions according to a principal quantum number n are filled: $N = 1$ for 1s, $N = 5$ for 1s2s2p, $N = 14$ for 1s2s2p3s3p3d shells of NOs, and so on. Thus, N must take the values

$$N = \frac{1}{3}n(n + \frac{1}{2})(n + 1) \quad n = 1, 2, 3, \dots \quad (15)$$

In other words, since N is proportional to n^3 , a possible interpretation of (12) is that it is essentially of the form

$$E(n) = E(\infty) + \left(\sum_{\mu=1}^N C_{\mu} \right)^2 (an^{-3} + bn^{-4} + \dots) \quad (16)$$

where n is the principal quantum number.

The ‘interference factor’ $(\sum_{\mu=1}^N C_{\mu})^2$, which can take on values between zero and one, is an important ingredient of CBS theory. Peterson and Nyden (1981) have interpreted this factor as being responsible for the fact that, when going from second-order ($1/Z$) perturbation theory to the infinite-order CI theory, the prefactor of the $(L+1)^{-3}$ terms decreases rapidly. For example, the prefactor in (4) is obtained by taking 42% of the prefactor in (2). In the modern CBS methods, a similar interference factor is computed from the coefficients of the first-order wavefunction of MP2 theory. More precisely, individual interference factors are computed for all the $\alpha\alpha$ and $\alpha\beta$ first-order pair functions. The interference factor, which implies that the MP2 truncation errors are *larger* in absolute terms than those at higher levels, provides a recipe for estimating the errors of high-level methods from the corresponding MP2 errors (Petersson and Licht 1981).

Compared with high-level methods such as coupled-cluster theory, MP2 calculations are inexpensive, and it appears promising to combine computations of the MP2 truncation error with high-level calculations in finite basis sets (Klopper *et al* 1994, Klopper and Lüthi 1996, 1999). Unfortunately, since the MP2 correction tends to overestimate the high-order errors (Martin 1997a), a direct addition of the full MP2 truncation correction is not optimal.

In the CBS-QCI/APNO method, each extrapolation for a single pair energy is scaled by an individual interference factor. However, an overall interference factor may be obtained by comparing the extrapolated but unscaled total MP2 correlation energy with the sum of the scaled extrapolations for the pair energies. In table 1, which contains details of the CBS-QCI/APNO calculations at the CCSD(T)(full)/cc-pCVQZ geometries, these overall interference factors are listed. (With the notation (full) we indicate that all electrons have been correlated, as opposed to the frozen-core (FC) approximation.)

We have compared the interference factors of table 1 with similar factors obtained from R12 calculations (cf section 7). The R12 methods, which employ explicitly correlated two-electron basis functions, yield MP2 and CCSD(T) energies (Purvis and Bartlett 1982, Raghavachari *et al* 1989, Bartlett 1995) very close to the basis-set limit. Thus, comparing MP2-R12 and CCSD(T)-R12 calculations with the finite basis-set MP2/APNO and CCSD(T)/APNO calculations, we can determine the ratio between the CCSD(T) and MP2 truncation errors. This has been done for both approximations currently used in MP2-R12 theory: the MP2-R12/A and MP2-R12/B approximations. The results for our 20 molecules are depicted in figure 1. The agreement between the CBS and R12 ratios is quite striking, in particular for the MP2-R12/A calculations.

Let us now consider the CBS/APNO extrapolation to the MP2 limit (Petersson *et al* 1985, 1988, 1991, Petersson and Braunstein 1985). For pairs of occupied α and β spin-orbitals ($\alpha\beta$ pairs) and for $\alpha\alpha$ -type second-order pair energies, the extrapolations are based on the expressions

$${}^{\alpha\beta}e_{ij}^{(2)}(N) = {}^{\alpha\beta}e_{ij}^{(2)}(\infty) + {}^{\alpha\beta}f_{ij} \frac{25}{512} (N + {}^{\alpha\beta}\delta_{ij})^{-1} \quad (17)$$

$${}^{\alpha\alpha}e_{ij}^{(2)}(N) = {}^{\alpha\alpha}e_{ij}^{(2)}(\infty) + {}^{\alpha\alpha}f_{ij} \frac{25}{512} (N + {}^{\alpha\alpha}\delta_{ij})^{-5/3} \quad (18)$$

Table 1. Individual contributions (in E_h) to the CBS-QCI/APNO total energy (E_{CBS}). The geometries have been optimized at the CCSD(T)(full)/cc-pCVQZ level.

Molecule	QCISD(T) ^a	$\delta E_{\text{SCF}}^{\text{b}}$	$\delta E_{\text{MP2}}^{\text{c}}$	Int. ^d	Emp. ^e	Core ^f	E_{CBS}
CH ₂	-39.0689	0.0	-0.0110	55.9	-0.0024	-0.0528	-39.1304
CH ₄	-40.4465	0.0	-0.0143	59.1	-0.0033	-0.0542	-40.5125
NH ₃	-56.4876	0.0	-0.0177	62.5	-0.0035	-0.0576	-56.5598
H ₂ O	-76.3542	-0.0001	-0.0212	66.8	-0.0039	-0.0597	-76.4320
HF	-100.3665	-0.0001	-0.0240	71.8	-0.0043	-0.0632	-100.4513
C ₂ H ₂	-77.2023	-0.0005	-0.0235	59.5	-0.0040	-0.1079	-77.3287
C ₂ H ₄	-78.4541	-0.0003	-0.0249	60.1	-0.0049	-0.1080	-78.5823
HCN	-93.2935	-0.0006	-0.0258	61.6	-0.0042	-0.1110	-93.4253
HNC	-93.2699	-0.0008	-0.0256	62.2	-0.0042	-0.1109	-93.4017
N ₂	-109.3963	-0.0009	-0.0283	64.1	-0.0044	-0.1136	-109.5333
N ₂ H ₂	-110.5023	-0.0006	-0.0311	63.8	-0.0053	-0.1139	-110.6418
CO	-113.1805	-0.0009	-0.0279	65.3	-0.0046	-0.1126	-113.3168
H ₂ CO	-114.3604	-0.0008	-0.0314	65.1	-0.0054	-0.1131	-114.5002
HNO	-130.3289	-0.0010	-0.0344	66.3	-0.0055	-0.1160	-130.4742
H ₂ O ₂	-151.3973	-0.0007	-0.0404	67.9	-0.0068	-0.1188	-151.5510
HOF	-175.3805	-0.0008	-0.0425	70.3	-0.0071	-0.1220	-175.5403
F ₂	-199.3494	-0.0009	-0.0444	72.4	-0.0075	-0.1253	-199.5151
CO ₂	-188.3711	-0.0018	-0.0484	67.2	-0.0077	-0.1728	-188.5859
O ₃	-225.1905	-0.0021	-0.0585	68.0	-0.0084	-0.1769	-225.4177
H ₂	-1.1731	0.0	-0.0023	47.6	-0.0008	0.0	-1.1750

^a [QCISD(T)-MP2](FC)/6-311++G(2df, p) + MP2(FC)/APNO. The APNO basis set is 6s6p3d2f for C, N, O, F and 4s2p1d for H.

^b Increment obtained from the CBS extrapolation to the Hartree–Fock limit.

^c Increment obtained from the CBS extrapolation to the MP2 limit. This increment is weighted with the overall interference factor given in the next column.

^d Overall interference factor, $(\sum_{\mu} C_{\mu})^2$, given in per cent.

^e An empirical correction.

^f Core correlation obtained from CBS extrapolated MP2 calculations using a 3s2p basis set for H and a 4s5p3d (Cartesian components) basis set for C, N, O, F.

where

$${}^{\alpha\beta} f_{ij} = |S|_{ij}^2 \quad {}^{\alpha\alpha} f_{ij} = 2|S|_{ij}^2 \left(\frac{1 - |S|_{ij}^2}{1 + |S|_{ij}^2} \right) \quad (19)$$

and

$$|S|_{ij} = \int |\varphi_i(\mathbf{r}) \varphi_j(\mathbf{r})| d\mathbf{r}. \quad (20)$$

We emphasize that the $\alpha\beta$ and $\alpha\alpha$ pairs of the spin-orbital CBS method do *not* represent spin-adapted singlet and triplet pairs; rather, the $\alpha\beta$ pair contributes to both singlet and triplet pairs. We also emphasize that $|S|_{ij}$ is the integral over the *absolute value* of the product of the occupied orbitals rather than the absolute value of the overlap integral. It requires a numerical integration. A formulation based on the overlap integral over the two charge distributions generated by the orbitals φ_i and φ_j , which is easily computed analytically, is perhaps an alternative worth investigating.

The purpose of the prefactors ${}^{\alpha\beta} f_{ij}$ and ${}^{\alpha\alpha} f_{ij}$ is to damp the extrapolated truncation errors for spatially distant pairs of (localized) orbitals. However, their introduction makes the method less suited for accurate calculations of weak interactions. For example, in the He dimer with $R_{\text{He-He}} = 5.6 a_0$, the CBS extrapolation does not contribute to the dispersion since

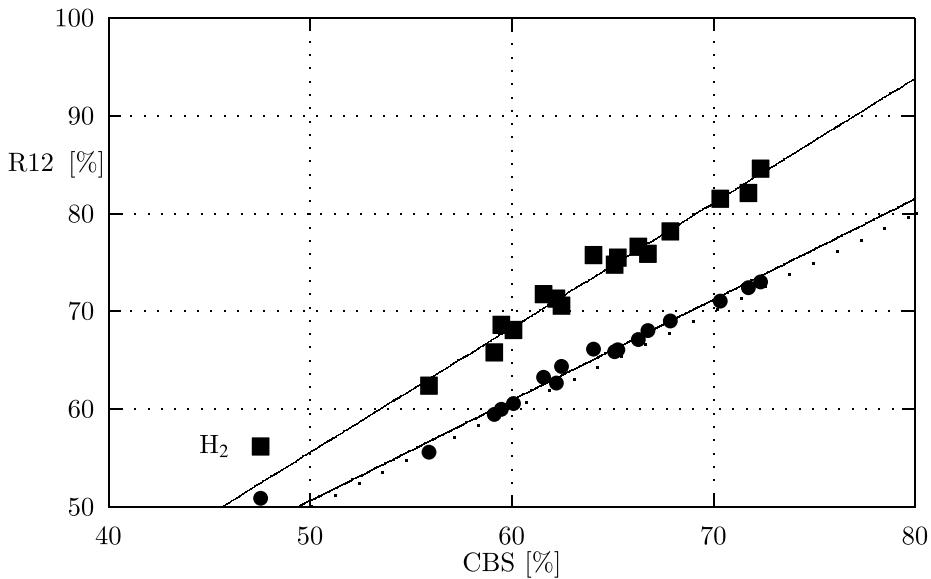


Figure 1. Ratio of CCSD(T) to MP2 basis-set truncation errors, $\delta E_{\text{CCSD(T)}}/\delta E_{\text{MP2}}$, of the APNO basis set. Comparison of the ratio (i.e. interference factor) computed by the CBS-QCI/APNO method with the ratio obtained from R12 calculations (cf section 7). At the R12 level, the CCSD(T) truncation error is $\delta E_{\text{CCSD(T)}} = \text{CCSD(T)} - \text{R12/B} - \text{CCSD(T)}/\text{APNO}$, while the MP2 truncation error is either $\delta E_{\text{MP2}} = \text{MP2} - \text{R12/A} - \text{MP2}/\text{APNO}$ (●) or $\delta E_{\text{MP2}} = \text{MP2} - \text{R12/B} - \text{MP2}/\text{APNO}$ (■). The dotted line is a straight line with slope one through the origin.

the f_{ij} factors tend to zero for the interatomic pairs. Although CBS extrapolations appear to be useful when bond functions are used (i.e. basis functions placed halfway between the atoms, cf Montgomery *et al* (1989)), we note that the standard CBS-QCI/APNO method—as implemented in the Gaussian 94 program—estimates the second-order basis-set truncation error for the He dimer to -5.534 mE_h , exactly twice that of the He atom. Still, Rablen *et al* (1998) have obtained useful CBS interaction energies for the hydrogen bond in the water dimer. Indeed, at a fixed dimer geometry (Klopper *et al* 1999), the two-body MP2(FC) interaction energy is extrapolated by CBS to $-20.6 \text{ kJ mol}^{-1}$, which compares well with the best estimate for the MP2(FC) basis-set limit of $-20.9 \text{ kJ mol}^{-1}$ (Klopper and Lüthi 1999, Klopper *et al* 1999). In contrast, the standard, unextrapolated MP2(FC)/APNO interaction is only $-19.7 \text{ kJ mol}^{-1}$.

In equations (17) and (18), ${}^{\alpha\beta}e_{ij}^{(2)}(\infty)$ and ${}^{\alpha\beta}\delta_{ij}$, which constitute the fitting parameters, are obtained from two-point fits. The first point is chosen as $N = 1$; that is, as the Hartree–Fock calculation, for which ${}^{\alpha\beta}e_{ij}^{(2)}(1) = 0$. A prescribed range of N values is then tried for the second point, each time computing the corresponding pair energy ${}^{\alpha\beta}e_{ij}^{(2)}(N)$. For each N , the two equations are solved for the two unknowns and the most negative ${}^{\alpha\beta}e_{ij}^{(2)}(\infty)$ is taken as the final, extrapolated second-order pair energy. The same procedure is followed for the $\alpha\alpha$ pairs. The extrapolated second-order corrections for the truncation error are shown in table 1 (δE_{MP2}), along with the Hartree–Fock corrections (δE_{SCF}) (Petersson and Brauenstein 1985, Petersson *et al* 1988). As discussed previously, to obtain the infinite-order corrections, the

second-order corrections are scaled by the interference factors

$$\delta e_{ij}^{(\infty)} = \left(\sum_{\mu_{ij}=1}^{N_{\text{virt}}+1} c_{\mu_{ij}}^{(1)} \right)^2 \delta e_{ij}^{(2)} \quad (21)$$

computed from the first-order wavefunction.

The CBS-QCI/APNO method also includes an empirical correction (Montgomery *et al* 1994),

$$\delta E_{\text{emp}} = -0.00174 \sum_{i=1}^{n_{\beta}} \left(\sum_{\mu_{ii}=1}^{N_{\text{virt}}+1} c_{\mu_{ii}}^{(1)} \right)^2 |S^{\alpha\beta}|_{ii}^2 \quad (22)$$

where $|S^{\alpha\beta}|_{ii}^2$ is the absolute overlap integral,

$$|S^{\alpha\beta}|_{ii}^2 = \int |\varphi_i^{\alpha}(r) \varphi_i^{\beta}(r)| dr \quad (23)$$

over the most similar α and β orbitals (Petersson and Al-Laham 1991). In equation (22), n_{β} is the number of β orbitals, which is taken to be smaller than the number of α orbitals. For closed-shell molecules, n_{β} is the number of doubly occupied orbitals and $|S^{\alpha\beta}|_{ii}^2 = 1$ for all φ_i . The empirical correction is also listed in table 1. Unlike the empirical corrections of the G1 and G2 methods discussed in section 6, the CBS correction is a smooth function of the geometry, thus contributing to the whole potential energy hypersurface of a molecule. Furthermore, a correction for core–core and core–valence correlation effects is computed at the MP2 level from CBS extrapolations with appropriate basis sets (table 1). Finally, there also exists a correction for spin contamination (Ochterski *et al* 1996), but this correction is unimportant for the present calculations on closed-shell molecules.

A comparison of the CBS-QCI/APNO method with other accurate methods is presented in section 8. For the purpose of this comparison, all CBS calculations presented here have been performed at the CCSD(T)(full)/cc-pCVQZ optimized geometries. This is not the way that CBS calculations are usually carried out, however. Most often, the CBS methods are employed within the framework of model chemistries, which include well defined procedures for obtaining geometries, vibrational (harmonic) frequencies and finite-temperature corrections.

5. The principal expansion and correlation-consistent basis sets

The successive addition of complete shells of functions that belong to a given angular-momentum quantum number ℓ is not the only systematic way to improve AO basis sets for correlated calculations. Quite the contrary, the addition of (nearly) complete ℓ -shells (that is, the partial-wave expansion) does not seem a practical approach for calculations of *molecular* electronic structure. Rather than using the quantum number ℓ for defining systematic sequences of AO basis sets, we may use the principal quantum number n . We shall refer to this approach as the ‘principal expansion’. For $n = 1$, there is only one AO (1s), for $n = 2$, there are five AOs (1s2s2p), for $n = 3$, there are 14 AOs (1s2s2p3s3p3d), and so on. We assume that the individual sets of 1, 5, 14, …, functions have been fully optimized, thereby defining a systematic expansion (figure 2). This principal expansion is closely related to the way that AO basis sets are usually applied in numerical quantum chemistry. In practice, the choice implies that those basis functions are added first that contribute most to the correlation energy.

The correlation-consistent basis sets of Dunning (1989) represent systematic sequences of AO sets that resemble closely the principal expansion. The basis sets are denoted cc-pVXZ

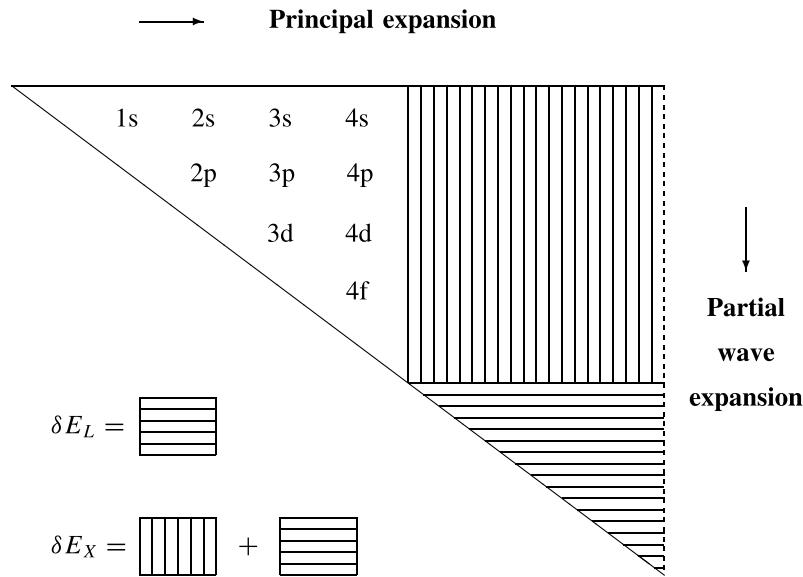


Figure 2. Principal and partial-wave expansions of the He ground state wavefunction.

for ‘correlation-consistent polarized valence X -tuple zeta’, where the ‘cardinal number’ X takes on the values $X = D, T, Q, 5, 6, \dots$. For H or He, a given cc-pVXZ basis set has the same composition as the principal basis set with $n = X$. For example, the cc-pVQZ basis set ($X = 4$) contains the functions: 4s3p2d1f. For first-row atoms, we identify n with $X + 1$.

The first investigations of the convergence of the correlation-consistent basis sets were carried out by Feller (1992, 1993), using the form

$$E(X) = E(\infty) + a \exp(-bX) \quad (24)$$

for fitting computed energies and extrapolations of the basis-set limit. Even though the exponential fit appears to be successful for relatively small X (Feller and Peterson 1998), the assumption of linear convergence (i.e. the exponential form) is not physically motivated and is too optimistic. For large X , it becomes important to account for the slow convergence of the computed correlation energy, which, as we shall shortly see, converges as X^{-3} .

For the ground state of the He atom, the energy contribution of an NO orbital is proportional to the inverse sixth power of its principal quantum number n (Bunge 1970). Indeed, for large n , Carroll *et al* (1979) found that the energy increments of the NOs are remarkably well described by

$$\Delta E_{n\ell m} \approx -A_\ell \left(n - \frac{1}{2}\right)^{-6} \quad (25)$$

with

$$A_\ell = \begin{cases} A_0 \approx 0.24 & \text{if } \ell = 0 \\ A' \approx 0.21 & \text{otherwise.} \end{cases} \quad (26)$$

The observation that the NO orbital increments are proportional to n^{-6} is a strong motivation for the development of systematic sequences of basis sets, where the level of hierarchy is determined by the principal quantum number. Examples of such systematic basis sets are the correlation-consistent (Dunning 1989, Kendall *et al* 1992, Woon and Dunning 1993, 1994,

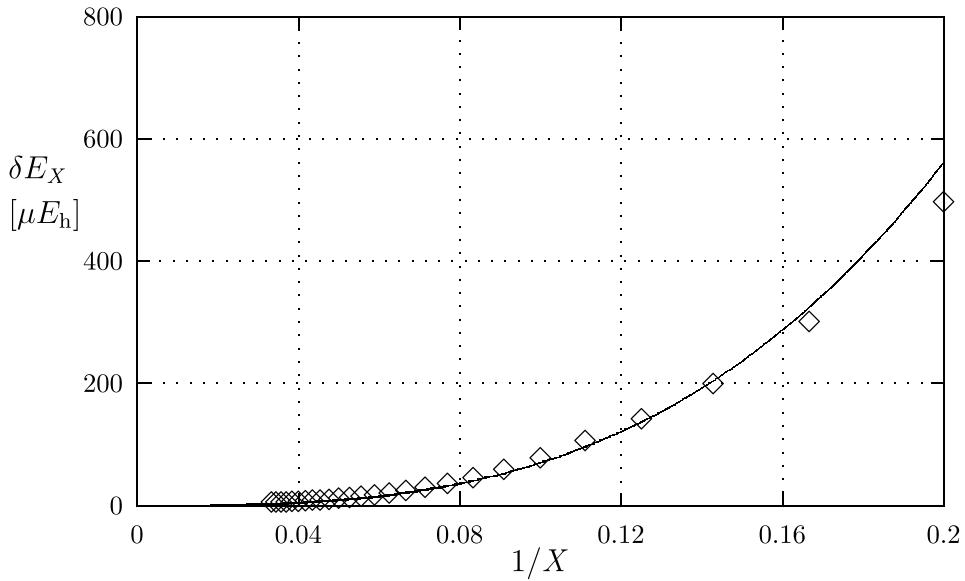


Figure 3. Basis set truncation error (δE_X in μE_h) of the He ground state energy as a function of $1/X$, where X is the cardinal number. The function $\delta E(X) = 0.07/X^3$ is shown as a full curve.

1995, Wilson *et al* 1996, Woon *et al* 1998) and atomic natural orbital (ANO) basis sets (Almlöf and Taylor 1987, 1992, Helgaker and Taylor 1995).

For a calculation that includes only shells with principal quantum number $n \leq X$, the truncation error δE_X is obtained by summing all increments $\Delta E_{n\ell m}$ for the combinations of quantum numbers n , ℓ , and m not included in the expansion:

$$\delta E_X \approx - \sum_{n=X+1}^{\infty} \sum_{\ell=0}^{n-1} \sum_{m=-\ell}^{\ell} \Delta E_{n\ell m}. \quad (27)$$

Replacing summation over n by integration, we obtain

$$\begin{aligned} \delta E_X &\approx A_0 \sum_{n=X+1}^{\infty} \left(n - \frac{1}{2}\right)^{-6} + A' \sum_{n=X+1}^{\infty} (n^2 - 1) \left(n - \frac{1}{2}\right)^{-6} \\ &\approx A_0 \int_{X+\frac{1}{2}}^{\infty} \left(n - \frac{1}{2}\right)^{-6} dn + A' \int_{X+\frac{1}{2}}^{\infty} (n^2 - 1) \left(n - \frac{1}{2}\right)^{-6} dn \\ &= \frac{1}{5} A_0 X^{-5} + \frac{1}{3} A' X^{-3} + \frac{1}{4} A' X^{-4} - \frac{3}{20} A' X^{-5}. \end{aligned} \quad (28)$$

We thus find that the truncation error of the principal expansion can be expressed as

$$\delta E_X \approx 0.07 X^{-3} + \dots \quad (29)$$

We have tested (29) numerically, retaining only the first term. Figure 3 shows the truncation error as a function of $1/X$. The plotted points have been computed from the increments in table VIII (the column marked ‘O-limit’) of Carroll *et al* (1979), where available, and from (25), where not. We note that the computed truncation errors are well represented by the asymptotic formula (29), which in figure 3 has been plotted with only the first term retained.

Moreover, from (25), we can compute the truncation error δE_L of a partial-wave expansion that includes all contributions up to $\ell \leq L$:

$$\begin{aligned}\delta E_L &\approx - \sum_{n=L+2}^{\infty} \sum_{\ell=L+1}^{n-1} \sum_{m=-\ell}^{\ell} \Delta E_{n\ell m} \\ &\approx A' \int_{L+\frac{3}{2}}^{\infty} [n^2 - (L+1)^2] (n - \frac{1}{2})^{-6} dn \\ &= \frac{2}{15} A' (L+1)^{-3} + \frac{1}{4} A' (L+1)^{-4} + \frac{1}{20} A' (L+1)^{-5} \\ &\approx 0.028 (L+1)^{-3} + \dots\end{aligned}\quad (30)$$

The prefactor of the $(L+1)^{-3}$ term is about 0.028, in reasonable agreement with the ‘exact’ value of 0.024 742 (Hill 1985). Also, we can compare (29) with the asymptotic formula derived by Nyden and Petersson (1981). Since N is proportional to the third power of X , Nyden and Petersson’s formula (12) can alternatively be written as

$$E(X) = E(\infty) + \left(\sum_{\mu=1}^N C_{\mu} \right)^2 \frac{75}{512} X^{-3} + O(X^{-4}). \quad (31)$$

Thus, the leading terms of (29) and (31) are similar if the interference factor $\left(\sum_{\mu} C_{\mu} \right)^2$ is close to 0.48. Indeed, for large X , the interference factor, which varies only slowly with X , is very close to this value (Petersson and Nyden 1981).

Obviously, since (25) is empirical, the above algebra does not constitute a rigorous derivation of the convergence rate of correlation-consistent basis sets. Furthermore, even though the correlation-consistent basis sets have the shell structure of the principal expansion, they are not based on (atomic or molecular) NOs. Rather, our analysis of the natural orbital contributions (25) should be regarded as *motivating* the use of extrapolations based on (29); that is, for fitting functions of the form

$$E(X) = E(\infty) + aX^{-3} + bX^{-4} + \dots \quad (32)$$

for energies obtained from cc-pVXZ basis sets. In short, our motivation for using (32) is that the convergence implied by this equation follows from the observation that the energy contribution of an NO orbital is proportional to the inverse sixth power of its principal quantum number.

6. Gaussian- N model chemistries

The Gaussian-1 (G1), Gaussian-2 (G2) and Gaussian-3 (G3) model chemistries were developed by Pople and co-workers with the aim of extracting from a series of *ab initio* calculations the most accurate estimate of the electronic energy of a molecule (Pople *et al* 1989, Curtiss *et al* 1990, 1991, 1992, 1993, 1998c). The models have been thoroughly studied, their performances are well documented (Curtiss *et al* 1997, 1998a, b, Petersson *et al* 1998), and comprehensive reviews exist (Raghavachari and Curtiss 1995, Curtiss and Raghavachari 1995). We can thus be brief in the present review, concentrating on the essential points of G2 single-point energy calculations.

The G2 model starts with QCISD(T)(FC)/6-311G(d, p) calculations (Pople *et al* 1987, Raghavachari *et al* 1989). This basis is relatively small and the effect of AO basis-set extensions is estimated by adding corrections computed at lower levels of theory. In particular, the difference between MP2(FC)/6-311+G(3df, 2p) and MP2(FC)/6-311G(d, p) is added as a second-order correction for the basis-set truncation error. In addition, two separate fourth-order corrections are added. The first one is concerned with the effect of diffuse AO functions

and is estimated by adding the difference between (MP4-MP2)(FC)/6-311+G(d, p) and (MP4-MP2)(FC)/6-311G(d, p) calculations. Analogously, the second fourth-order correction, which is concerned with the effect of additional polarization functions, is estimated from the difference between (MP4-MP2)(FC)/6-311G(2df, p) and (MP4-MP2)(FC)/6-311G(d, p) calculations. By using the notation (MP4-MP2)(FC), we indicate that only the valence-shell third- and fourth-order correlation energies are taken into account, calculated as the difference between the MP4 and MP2 total frozen-core energies.

Finally, the G2 model contains an empirical correction, denoted as the ‘higher-level correction’, of the form

$$\Delta_{\text{HLC}}/\text{m}E_{\text{h}} = -5.00 \min(n_{\alpha}, n_{\beta}) - 0.19|n_{\alpha} - n_{\beta}| \quad (33)$$

where n_{α} and n_{β} are the numbers of occupied α and β spin orbitals, respectively. Thus, an energy of $-5 \text{ m}E_{\text{h}}$ is added for each pair of electrons, and $-0.19 \text{ m}E_{\text{h}}$ for each unpaired electron. Note that the higher-level correction does not contribute to the reaction enthalpies computed in section 8, which involve only closed-shell molecules.

The G3 method has been developed only very recently (Curtiss *et al* 1998c). It is very similar to the G2 method, but it uses slightly different basis sets, and it includes a modified HLC correction and in addition to this a correction for core–core and core–valence correlation effects. The latter effects are neglected in the G2 model.

7. Explicitly correlated wavefunctions

Thus far, we have discussed the convergence of the correlation energy in the framework of basis-set expansions in Slater determinants constructed from one-electron functions. Insight into the convergence of such expansions can be exploited to develop extrapolation schemes based on systematic sequences of basis sets. However, a major drawback of the extrapolation techniques is that they are biased towards the (total) energy of the system. Thus, although the techniques may provide accurate estimates of the total energy, they do not provide accurate electronic wavefunctions or accurate electron densities. In principle, these extrapolations can also be employed to relative energies such as binding energies (Klopper and Helgaker 1998), reaction enthalpies, and excitation energies or to energy-related properties derived from the potential energy hypersurfaces such as geometries and force fields. However, such applications are often hampered by the fact that the one-electron basis-set hierarchies are often less systematic with respect to the property of interest than originally anticipated. As basis sets are enlarged, effects other than the systematic improvement of the description of electron correlation also take place. The extrapolation techniques become particularly difficult with respect to van der Waals interactions or electric and magnetic molecular properties, which require basis sets with special characteristics (Halkier *et al* 1997). In short, it is important to develop and search for methods that enable the computation of accurate molecular electronic wavefunctions; that is, the accurate determination of molecular electronic structure.

An alternative to using products of one-electron functions (i.e. Slater determinants) is to employ basis functions that depend *explicitly* on the coordinates of two electrons. Methods of this type are known as explicitly correlated methods or r_{12} -dependent wavefunctions. For a recent review, see Klopper (1998). These explicitly correlated methods can be divided into two broad families of methods: the Gaussian geminals methods and the linear r_{12} methods.

The first family of explicitly correlated methods employs two-electron functions (geminals) containing Gaussian functions of r_{12} , the distance between the electrons 1 and 2:

$$G_{\gamma}(12) = \exp(-\gamma r_{12}^2). \quad (34)$$

Such Gaussian geminals have a long tradition in numerical quantum chemistry, but modern computer implementations concentrate on two main applications: first, FCI calculations for relatively small molecular systems with up to six electrons (Cencek and Rychlewski 1993, 1995, Komasa and Rychlewski 1997). Secondly, applications to larger systems and to intermolecular interactions within the framework of MP2 and symmetry-adapted perturbation theory (SAPT) (Bukowski *et al* 1996, Williams *et al* 1996, Persson and Taylor 1996, Korona *et al* 1997).

The second family of explicitly correlated methods (the linear r_{12} methods) uses basis functions where the r_{12} factors enter the wavefunction linearly. These linear terms have been implemented into single-reference coupled-cluster methods (Kutzelnigg 1985, Klopper and Kutzelnigg 1987, 1991, Kutzelnigg and Klopper 1991, Klopper 1991, Klopper *et al* 1991, Noga *et al* 1992, 1995, 1997, Klopper and Almlöf 1993, Noga and Kutzelnigg 1994, Klopper and Noga 1995) and multi-reference CI methods (Gdanitz 1993, 1998a–d, 1999, Gdanitz and Röhse 1995, 1996). Currently, the linear r_{12} methods are the only explicitly correlated methods applicable to large molecules.

There exist recent reviews concerned with explicitly correlated wavefunctions (Noga *et al* 1997, Klopper 1998, Bukowski *et al* 1999). Therefore, the theory of the r_{12} methods need not be repeated here in full detail. Rather, we sketch the most important ideas and ingredients of the methods, restricting ourselves to the single-reference coupled-cluster linear r_{12} methods, which in section 8 are used for the comparison of various accurate computational techniques.

The CCD-R12 ansatz for the explicitly correlated coupled-cluster doubles (CCD) model is based on a doubles operator that consists of the standard doubles replacement operator augmented with contracted doubles replacements into a complete basis set $\{\varphi_\alpha\}$ of virtual orbitals (we assume summation over repeated indices),

$$T_2 = \frac{1}{8} c_{kl}^{ij} \bar{R}_{\alpha\beta}^{kl} a_{ij}^{\alpha\beta} + \frac{1}{4} t_{ab}^{ij} a_{ij}^{ab}. \quad (35)$$

The contraction coefficients are given by

$$\bar{R}_{\mu\nu}^{kl} = \langle \mu\nu | Q(12) r_{12} | kl \rangle = \bar{r}_{\mu\nu}^{kl} - \delta_\nu^q \bar{r}_{\mu q}^{kl} - \delta_\mu^p \bar{r}_{p\nu}^{kl} + \frac{1}{2} \Delta_{\mu\nu}^{pq} \bar{r}_{pq}^{kl} \quad (36)$$

with

$$\bar{r}_{\mu\nu}^{kl} = \langle \mu\nu | r_{12} | kl \rangle \quad (37)$$

$$\Delta_{\kappa\lambda}^{\mu\nu} = \delta_\kappa^\mu \delta_\lambda^\nu - \delta_\lambda^\mu \delta_\kappa^\nu \quad (38)$$

and

$$Q(12) = \{1 - P(1)\}\{1 - P(2)\} \quad P = \sum_p |\varphi_p\rangle\langle\varphi_p|. \quad (39)$$

Here we use a tensor notation for the replacement operators $a_{ij}^{\alpha\beta}$ and a_{ij}^{ab} , where the indices i, j, k, \dots are used for occupied spin-orbitals, a, b, c, \dots for unoccupied (virtual) spin-orbitals, and p, q, r, \dots for arbitrary spin-orbitals. For the spin-orbital space complementary to the occupied space, we use Greek symbols $\alpha, \beta, \gamma, \dots$, and the complete set of spin-orbitals is denoted by $\kappa, \lambda, \mu, \dots$:

$$\sum_i |\varphi_i\rangle\langle\varphi_i| + \sum_\alpha |\varphi_\alpha\rangle\langle\varphi_\alpha| = \sum_\mu |\varphi_\mu\rangle\langle\varphi_\mu| = 1 \quad (40)$$

$$\sum_i |\varphi_i\rangle\langle\varphi_i| + \sum_a |\varphi_a\rangle\langle\varphi_a| = \sum_p |\varphi_p\rangle\langle\varphi_p| = P. \quad (41)$$

The use of the operator \bar{R} ensures that the functions $\bar{R}_{\alpha\beta}^{kl} a_{ij}^{\alpha\beta} |\Phi\rangle$ are strongly orthogonal to all Slater determinants constructed from orbitals contained in the finite basis $\{\varphi_p\}$. In a first-quantization form (Klopper and Kutzelnigg 1991), the ansatz (35) and (36) reads

$$|\tau_{ij}\rangle = \sum_{k < l} c_{kl}^{ij} Q(12) r_{12} |kl\rangle + \sum_{a < b} t_{ab}^{ij} |ab\rangle \quad (42)$$

where the τ_{ij} are explicitly correlated pair functions. Models other than the CCD-R12 model are obtained by adding the standard singles and triples replacements operators, which contain no r_{12} -dependent terms.

At this point, we note that Gaussian geminals are easily introduced into the wavefunction by replacing r_{12} in the above equations by a Gaussian gernal or by a linear combination of geminals. Of course, more general functions of the type $f(r_{12})$ can also be chosen. The linear r_{12} correlation factors have the advantage that the electronic wavefunctions can be expanded such that they (nearly) fulfil the electron-electron cusp conditions. On the other hand, a short expansion in six or nine Gaussian geminals shows, with regard to accuracy, roughly the same performance as the linear r_{12} function itself (Persson and Taylor 1996), even though the gernal expansion cannot fulfil the cusp conditions. However, the chief difference between the linear r_{12} approach and the Gaussian geminals method is the way that many-electron integrals are computed—or better, avoided.

In closed-shell, single-reference explicitly correlated coupled-cluster methods, non-standard two-, three-, four- and even five-electron integrals occur in addition to the standard one- and two-electron integrals. Indeed, the occurrence of these many-electron integrals has hampered the development of explicitly correlated methods for a long time. However, the coupled-cluster energy functionals can be altered in a particular way (replacing the strongly orthogonal pair functions by ‘weakly orthogonal’ functions) to reduce the number of electrons in the many-electron integrals. Thus, in Gaussian geminals methods with weakly orthogonal pair functions, three-electron rather than four-electron integrals occur in MP2 theory, and four-electron rather than five-electron integrals occur in coupled-cluster theory. The linear r_{12} -type methods, on the other hand, follow the original suggestion made by Kutzelnigg in 1985, avoiding the many-electron integrals by insertion of the resolution of the identity, thereby retaining the accuracy of the methods. Then, all integrals can be reduced to two-electron integrals at the most, which has proven to be an important advantage of the linear r_{12} -type methods, making them efficient and practical.

To illustrate the insertion of the resolution of the identity, which has been denoted as the ‘standard approximation’ of linear r_{12} theory, we consider the three-electron integral

$$I_3 = \langle \varphi_i(1) \varphi_j(2) \varphi_k(3) | f(r_{12}) \frac{1}{r_{13}} | \varphi_l(1) \varphi_m(2) \varphi_n(3) \rangle. \quad (43)$$

This integral is approximated as

$$\begin{aligned} I_3 &\approx \langle \varphi_i(1) \varphi_j(2) \varphi_k(3) | f(r_{12}) P(1) \frac{1}{r_{13}} | \varphi_l(1) \varphi_m(2) \varphi_n(3) \rangle \\ &= \sum_p \langle \varphi_i(1) \varphi_j(2) | f(r_{12}) | \varphi_p(1) \varphi_m(2) \rangle \langle \varphi_p(1) \varphi_k(2) | \frac{1}{r_{12}} | \varphi_l(1) \varphi_n(2) \rangle. \end{aligned} \quad (44)$$

There is no unique way of inserting the resolution of the identity. Rather, two schemes are in use, known as the standard approximations A and B. Standard approximation B is used for all coupled-cluster and multi-reference CI methods, while both approximations A and B are used at the single-reference MP2 level. For details, see Noga *et al* (1997) and Klopper (1998).

An advantage of the Gaussian geminals is that the many-electron integrals over Gaussian geminals are much simpler to evaluate than the same integrals over linear r_{12} factors. However,

this advantage applies only if the many-electron integrals are evaluated without resorting to approximations. If, at any stage, the resolution of identity is invoked, the argument of the simpler Gaussian integrals cannot be given much weight.

8. Comparison of various methods

8.1. Computational details

An assessment of various quantum chemical calculations has been conducted by performing calculations on the 20 closed-shell molecules CH₂ ($\tilde{\alpha}^1A_1$ state), CH₄, NH₃, H₂O, HF, C₂H₂, C₂H₄, HCN, HNC, N₂, N₂H₂, CO, H₂CO, HNO, H₂O₂, HOF, F₂, CO₂, O₃ and H₂. All calculations have been performed at the CCSD(T)(full)/cc-pCVQZ optimized geometries.

The B3LYP (Becke 1993, Lee *et al* 1988), G2 and CBS-QCI/APNO calculations have been performed using standard basis sets and the Gaussian 94 suite of programs (Frisch *et al* 1995), the CCSD(T) calculations (Purvis and Bartlett 1982, Raghavachari *et al* 1989, Bartlett 1995) with the correlation-consistent basis sets have been performed partly with a development version of the DALTON program (Helgaker *et al* 1997a) and partly with the ACESII program (Stanton *et al*), and the explicitly correlated CCSD(T)-R12 calculations have been carried out with the DIRCCR12-95 program (Noga and Klopper 1995).

The resolution of the identity invoked in the R12 calculations requires basis sets optimized with respect to the standard approximations (A or B). Except for the fluorine atom, such basis sets have been optimized recently (Bakken 1999), derived from the TZV (Schäfer *et al* 1994), VTZ (Schäfer *et al* 1992) and aug-cc-pV5Z (Dunning 1989, Kendall *et al* 1992) basis sets. In table 2, we present these basis sets together with a similar basis set for the fluorine atom. The results (total energies) of the R12 calculations of the 20 molecules are collected in tables 3 and 4.

Table 2. Exponents of augmented functions. For C, N, O and F, the 11s6p primitive set of the TZV basis (Schäfer *et al* 1994) has been augmented with the 5d4f part of the aug-cc-pV5Z basis and with the functions given below. For H, the 5s primitive set of the VTZ basis (Schäfer *et al* 1992) has been augmented with the 5p4d part of the aug-cc-pV5Z basis and with the functions given below.

		C	N	O	F	H
Tight	s	140 000.0	210 000.0	270 000.0	360 000.0	230.0
	p	200.0	280.0	360.0	460.0	
	d	8.0	12.0	15.0	19.9	
	f	4.8	7.2	10.4	14.0	
Diffuse	s	0.040	0.055	0.074	0.094	0.032
	p	0.035	0.050	0.058	0.070	

In these calculations, the core correlation due to the 1s orbitals is treated differently by the various methods: in the coupled-cluster calculations, we have accounted for core correlation through CCSD(T) calculations with the cc-pCV5Z basis set, adding the difference between the CCSD(T)(full)/cc-pCV5Z and CCSD(T)(FC)/cc-pCV5Z energies to the CCSD(T)-R12 and standard CCSD(T) calculations with all other correlation-consistent basis sets (see table 4). The B3LYP calculations automatically account for electron correlation, CBS-QCI/APNO calculations include core-correlation corrections and G2 calculations neglect core correlation.

Table 3. Valence-only correlated electronic energies (in E_h) at various levels of MP2-R12 theory as obtained with the basis sets 13s8p6d5f for C, N, O, F and 7s5p4d for H. The geometries have been optimized at the CCSD(T)(full)/cc-pCVQZ level.

Molecule	SCF	MP2	MP2-R12/A	MP2-R12/B
CH ₂	-38.8953	-39.0424	-39.0516	-39.0502
CH ₄	-40.2161	-40.4246	-40.4354	-40.4339
NH ₃	-56.2238	-56.4739	-56.4888	-56.4869
H ₂ O	-76.0660	-76.3475	-76.3671	-76.3644
HF	-100.0692	-100.3642	-100.3889	-100.3853
C ₂ H ₂	-76.8543	-77.1802	-77.2008	-77.1974
C ₂ H ₄	-78.0693	-78.4225	-78.4425	-78.4393
HCN	-92.9143	-93.2788	-93.3023	-93.2987
HNC	-92.8989	-93.2505	-93.2739	-93.2702
N ₂	-108.9912	-109.3874	-109.4137	-109.4094
N ₂ H ₂	-110.0479	-110.4811	-110.5087	-110.5045
CO	-112.7888	-113.1670	-113.1941	-113.1899
H ₂ CO	-113.9212	-114.3417	-114.3707	-114.3662
HNO	-129.8475	-130.3122	-130.3445	-130.3395
H ₂ O ₂	-150.8494	-151.3822	-151.4204	-151.4147
HOF	-174.8194	-175.3660	-175.4090	-175.4023
F ₂	-198.7692	-199.3355	-199.3828	-199.3752
CO ₂	-187.7220	-188.3626	-188.4107	-188.4032
O ₃	-224.3617	-225.1874	-225.2426	-225.2332
H ₂	-1.1333	-1.1663	-1.1676	-1.1674

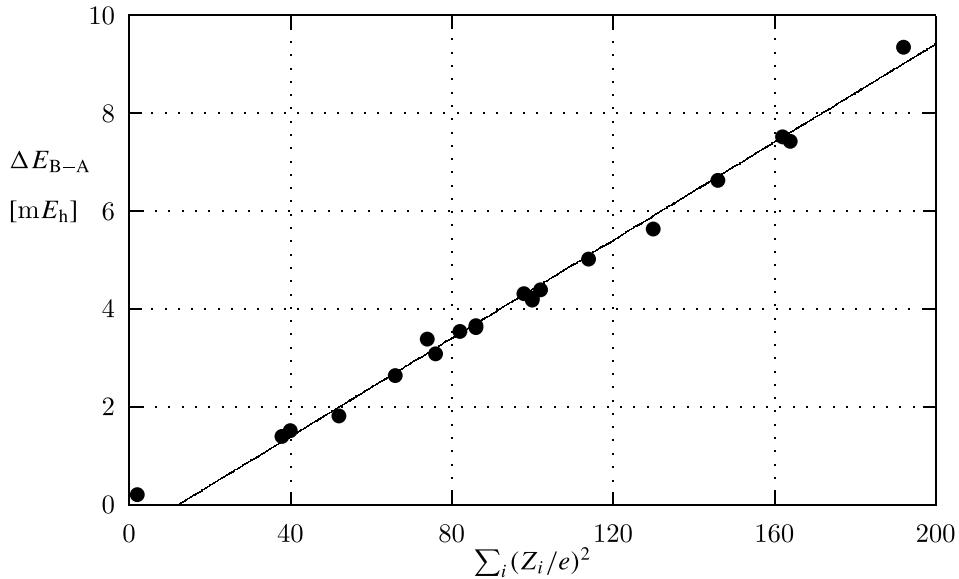


Figure 4. Energy difference ΔE_{B-A} between the MP2-R12/B and MP2-R12/A correlation energies obtained with the 13s8p6d5f (7s5p4d) basis set for C, N, O, F (H) as a function of the sum of the squares of the nuclear charges Z_i in the molecule.

8.2. Linear R12 results

In table 3, we compare the MP2-R12/A and MP2-R12/B energies with the standard SCF and MP2 total energies, computed with the same $[X/H] = [13s8p6d5f/7s5p4d]$ basis set,

Table 4. Valence-only correlated electronic energies (in E_h) at various levels of CC-R12 theory as obtained with the basis sets 13s8p6d5f for C, N, O, F and 7s5p4d for H. The geometries have been optimized at the CCSD(T)(full)/cc-pCVQZ level.

Molecule	CCSD-R12	CCSD[T]-R12	CCSD(T)-R12	Core ^a
CH ₂	-39.0704	-39.0758	-39.0757	-0.0548
CH ₄	-40.4484	-40.4555	-40.4552	-0.0561
NH ₃	-56.4922	-56.5012	-56.5009	-0.0590
H ₂ O	-76.3627	-76.3721	-76.3718	-0.0620
HF	-100.3808	-100.3891	-100.3887	-0.0653
C ₂ H ₂	-77.2000	-77.2181	-77.2175	-0.1122
C ₂ H ₄	-78.4539	-78.4705	-78.4700	-0.1117
HCN	-93.2929	-93.3129	-93.3120	-0.1148
HNC	-93.2695	-93.2890	-93.2881	-0.1145
N ₂	-109.3981	-109.4190	-109.4181	-0.1173
N ₂ H ₂	-110.5053	-110.5267	-110.5260	-0.1172
CO	-113.1835	-113.2033	-113.2019	-0.1172
H ₂ CO	-114.3657	-114.3850	-114.3840	-0.1177
HNO	-130.3341	-130.3570	-130.3560	-0.1201
H ₂ O ₂	-151.4097	-151.4316	-151.4309	-0.1235
HOF	-175.3967	-175.4186	-175.4176	-0.1266
F ₂	-199.3687	-199.3906	-199.3898	-0.1299
CO ₂	-188.3790	-188.4123	-188.4101	-0.1799
O ₃	-225.1849	-225.2388	-225.2352	-0.1844
H ₂	-1.1742	-1.1742	-1.1742	0.0

^a Core correlation computed as energy difference between CCSD(T)(full)/cc-pCV5Z and CCSD(T)(FC)/cc-pCV5Z energies.

where $X = \text{C}, \text{N}, \text{O}, \text{F}$. Whereas the difference between the MP2-R12 and MP2 energies is of the order of 10–40 m E_h , the difference between approximations A and B is of the order of a few m E_h . Since the latter difference is proportional to the sum of the squares of the nuclear charges in the molecule (figure 4), it makes little difference for the computation of the enthalpies of reaction whether the standard approximation A or B is used.

8.3. Standard CCSD(T)/cc-pVXZ calculations

The total electronic energies of the 20 molecules as obtained from standard CCSD(T)/cc-pVXZ calculations (for $X = 2, \dots, 6$) are displayed in table 5. These energies refer to fixed CCSD(T)(full)/cc-pCVQZ geometries and include only electron correlation effects due to the valence shell. For a comparison with the experimental enthalpies of reaction, a correction (see table 4) for core–core and core–valence correlation effects should be added to the total energies of table 5.

8.4. Experimental enthalpies of reaction

When comparing the computed (electronic) enthalpies of reaction, ΔH , with the values extracted from experimental data, we encounter the problem that some of the experimental error bars are substantially larger than our target accuracy of 1 kJ mol⁻¹. For example, the uncertainty of one of the reactions is as large as 4 kJ mol⁻¹ (cf. table 6). How should we treat the deviation of a few kJ mol⁻¹ (that is, a deviation well within the experimental error bars) from a very uncertain experimental mean value?

Table 5. Valence-only correlated electronic energies (in E_h) at the standard CCSD(T) level as obtained with various correlation-consistent cc-pVXZ basis sets. The geometries have been optimized at the CCSD(T)(full)/cc-pCVQZ level.

Molecule	cc-pVDZ	cc-pVTZ	cc-pVQZ	cc-pV5Z	cc-pV6Z	$X = \infty^a$
CH ₂	-39.0220	-39.0614	-39.0719	-39.0750	-39.0761	-39.0774
CH ₄	-40.3868	-40.4381	-40.4509	-40.4545	-40.4558	-40.4574
NH ₃	-56.4020	-56.4732	-56.4930	-56.4994	-56.5015	-56.5039
H ₂ O	-76.2410	-76.3322	-76.3598	-76.3690	-76.3720	-76.3757
HF	-100.2275	-100.3379	-100.3727	-100.3847	-100.3886	-100.3935
C ₂ H ₂	-77.1092	-77.1876	-77.2093	-77.2158	-77.2181	-77.2211
C ₂ H ₄	-78.3544	-78.4388	-78.4617	-78.4685	-78.4711	-78.4744
HCN	-93.1884	-93.2751	-93.3013	-93.3093	-93.3122	-93.3159
HNC	-93.1632	-93.2513	-93.2777	-93.2856	-93.2885	-93.2921
N ₂	-109.2753	-109.3739	-109.4044	-109.4142	-109.4178	-109.4223
N ₂ H ₂	-110.3670	-110.4780	-110.5116	-110.5225	-110.5263	-110.5309
CO	-113.0544	-113.1555	-113.1879	-113.1982	-113.2018	-113.2065
H ₂ CO	-114.2183	-114.3338	-114.3690	-114.3803	-114.3842	-114.3892
HNO	-130.1710	-130.2984	-130.3381	-130.3513	-130.3559	-130.3617
H ₂ O ₂	-151.1937	-151.3586	-151.4082	-151.4250	-151.4314	-151.4395
HOF	-175.1519	-175.3343	-175.3907	-175.4103	-175.4169	-175.4255
F ₂	-199.0975	-199.2961	-199.3589	-199.3810	-199.3886	-199.3983
CO ₂	-188.1475	-188.3271	-188.3845	-188.4031	-188.4096	-188.4180
O ₃	-224.9091	-225.1326	-225.2023	-225.2261	-225.2345	-225.2450
H ₂	-1.1634	-1.1723	-1.1738	-1.1742	-1.1743	-1.1745

^a Extrapolation to the limit of a complete basis.

We have here chosen to damp the deviation of the computed value with an exponential function depending on the experimental uncertainty σ (standard deviation). Thus, rather than investigating the directly obtained absolute error,

$$\delta(\Delta H) \equiv \delta = |\Delta H_{\text{computed}} - \Delta H_{\text{experimental}}| \quad (45)$$

where $\Delta H_{\text{computed}}$ is the computed enthalpy of reaction and $\Delta H_{\text{experimental}}$ is the experimental value, we study the damped error

$$\delta_{\text{damped}} = \delta \{1 - \exp(-\delta^2/2\sigma^2)\}. \quad (46)$$

This procedure ensures that only those errors significantly larger than the experimental uncertainty contribute to the statistical analysis. The damping function is depicted in figure 5.

8.5. Computed enthalpies of reaction

The computed electronic contributions to our selection of enthalpies of reaction are listed in tables 6 and 7. The reference values have been obtained from a combination of experimental standard heats of formation at 0 K (Chase 1998) with partly experimentally derived, partly computed zero-point vibrational energies. A detailed analysis of the reference values will be presented elsewhere (Helgaker *et al* 1999). Unfortunately, reliable experimental data for the molecules HNC and N₂H₂ were not available to the authors when preparing this review.

Let us first compare table 4 with table 5. We first observe that the CCSD(T)-R12 values as obtained with the [X/H] = [13s8p6d5f/7s5p4d] basis set are almost identical with the CCSD(T)/cc-pV6Z values. Actually, the latter are slightly lower. Thus, the two-point extrapolation (11) with $A = 6$ yields energies even further below the CCSD(T)-R12 values,

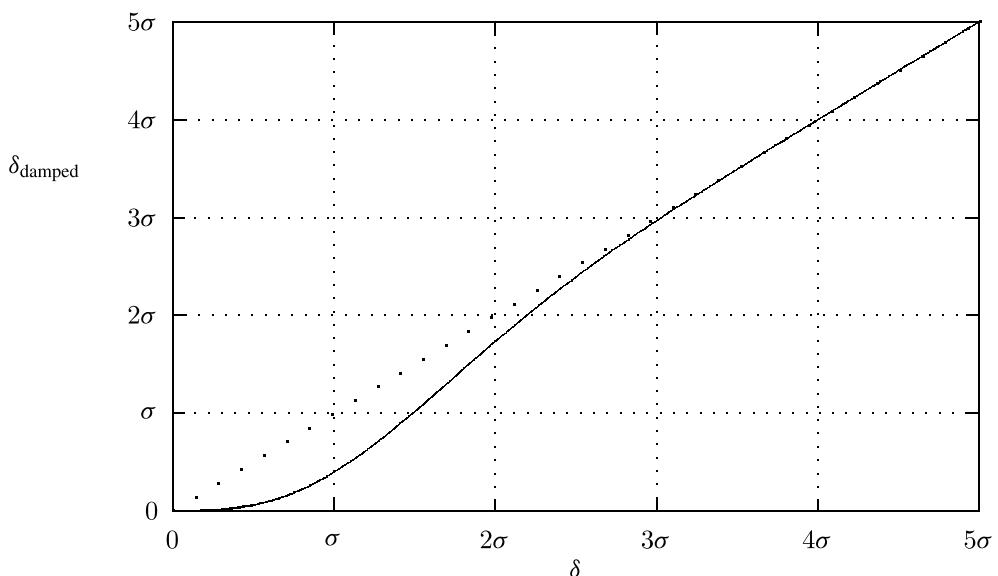


Figure 5. Illustration of the damping function used to reduce computed errors δ that are within the experimental standard deviation σ . When the error is equal to σ or smaller, only 39% or less of the error is taken. For errors larger than 2σ , 86% or more is taken into account.

see table 5 under $X = \infty$. We note, however, that the basis set chosen for the R12 calculations is of intermediate size and that the CCSD(T)-R12 numbers obtained with this set should be accurate to within a few mE_h . Although more accurate CCSD(T)-R12 calculations with larger basis sets are easily possible, we present here results that can be obtained routinely, using standardized R12-type Gaussian basis sets. The cost of these CCSD(T)-R12 calculations is relatively small, comparable with standard calculations using basis sets between cc-pVQZ and cc-pV5Z.

Let us now consider the computed enthalpies of reaction, as displayed in tables 6 and 7. The excellent performance and superiority of the CCSD(T) model is clearly demonstrated in table 6—both by the R12 and $X = \infty$ numbers. Large errors (i.e. of the order of 10 kJ mol^{-1}) are observed only for the reactions involving O_3 and HOF. Concerning O_3 , we must acknowledge that the single-reference CCSD(T) method cannot describe this molecule very well, its near degeneracies requiring a multi-reference treatment or the inclusion of higher-order connected excitations. This is obvious from the fact that the (T) triples correction alone contributes about 60 kJ mol^{-1} to the enthalpy of reaction. Moreover, the (T) and [T] triples corrections differ by as much as 7 kJ mol^{-1} . In such cases, one cannot expect convergence to within one or a few kJ mol^{-1} at the CCSD(T) level.

The HOF molecule, on the other hand, should be well described by the CCSD(T) method. The discrepancy of 11 kJ mol^{-1} cannot be explained by deficiencies in the wavefunction. We note that, for this reaction, the (T) and [T] triples corrections do not differ, contributing only 3 kJ mol^{-1} to the enthalpy of reaction. We suspect that the tabulated value of the JANAF tables is incorrect. Thus, we have excluded the two reactions involving the O_3 and HOF molecules from our analysis.

Table 6 shows that the CCSD(T) calculations are accurate to within 1 kJ mol^{-1} . In the form of CBS-QCI/APNO calculations, the CBS method achieves a root-mean-square (rms) error of

Table 6. Electronic contributions to reaction enthalpies (in kJ mol⁻¹) as obtained from various computational methods.

Reaction	Ref. ^a	R12 ^b	CBS ^c	G2 ^d	DFT ^e	X = ∞ ^f
CH ₂ ^g + H ₂ → CH ₄	-544(2)	-542	-544	-534	-543	-543
C ₂ H ₂ + H ₂ → C ₂ H ₄	-203(2)	-204	-207	-202	-208	-206
C ₂ H ₂ + 3 H ₂ → 2 CH ₄	-446(2)	-447	-450	-440	-450	-447
N ₂ H ₂ → N ₂ + H ₂		-174	-174	-177	-162	-173
CO + H ₂ → H ₂ CO	-21(1)	-22	-22	-17	-34	-23
N ₂ + 3 H ₂ → 2 NH ₃	-164(1)	-162	-161	-147	-166	-165
F ₂ + H ₂ → 2 HF	-563(1)	-562	-558	-564	-540	-564
O ₃ + 3 H ₂ → 3 H ₂ O	-933(2)	-943	-928	-912	-909	-946
H ₂ CO + 2 H ₂ → CH ₄ + H ₂ O	-251(1)	-250	-248	-235	-234	-250
H ₂ O ₂ + H ₂ → 2 H ₂ O	-365(2)	-365	-362	-360	-346	-362
CO + 3 H ₂ → CH ₄ + H ₂ O	-272(1)	-272	-270	-251	-268	-273
HCN + 3 H ₂ → CH ₄ + NH ₃	-320(3)	-320	-321	-305	-320	-321
HNO + 2 H ₂ → H ₂ O + NH ₃	-444(1)	-445	-440	-426	-429	-446
HNC → HCN		-64	-62	-63	-58	-63
H ₂ O + F ₂ → HO ₂ + HF	-129(4)	-118	-117	-123	-119	-118
CO ₂ + 4 H ₂ → CH ₄ + 2 H ₂ O	-244(1)	-242	-238	-216	-211	-244
2 CH ₂ ^g → C ₂ H ₄	-844(3)	-842	-844	-829	-845	-845
Mean error ^h		0	1	12	7	0
Mean absolute error ^h		0	2	12	10	1
Root-mean-square error ^h		0	3	14	15	1
Maximum absolute error ^h		1	6	28	34	3

^a Reference values for the electronic energies have been extracted from experimental data (Chase 1998, Helgaker *et al* 1999), corrected for zero-point vibrational energy. The values in parentheses indicate the uncertainty in the last digit. All molecular geometries used for the quantum mechanical calculations have been determined through CCSD(T)(full)/cc-pCVQZ optimizations.

^b CCSD(T)-R12 results obtained with the basis sets 1s8p6d5f for C, N, O, F and 7s5p4d for H. The difference between CCSD(T)(full)/cc-pCV5Z and CCSD(T)(FC)/cc-pCV5Z correlation energies has been added to account for core–core and core–valence correlation effects.

^c CBS-QCI/APNO method as implemented in the Gaussian 94 program (Frisch *et al* 1995). Core correlation has been included.

^d Gaussian-2 model as implemented in the Gaussian 94 program (Frisch *et al* 1995). Core correlation is *not* included in this model.

^e B3LYP/6-311++G(2df, p) results.

^f Extrapolation of CCSD(T)(full) results to the limit of a complete basis.

^g Singlet state, $\tilde{\alpha}^1 A_1$, of methylene.

^h Reactions with O₃ and HO₂ omitted.

3 kJ mol⁻¹—an excellent result in view of the modest cost of these calculations. In contrast, the G2 and DFT (B3LYP/6-311++G(2df, p)) schemes yield disappointing results, with rms errors of 14 and 15 kJ mol⁻¹, respectively. Adding corrections (obtained at the CCSD(T)/cc-pCV5Z level) for core–core and core–valence correlation effects to the G2 results would bring down its rms error only from 14 to 13 kJ mol⁻¹. In view of this small change, we do not expect the G3 method to perform substantially better than the G2 method.

Table 7 displays the basis-set convergence of standard CCSD(T)/cc-pVXZ calculations. An important and perhaps depressing observation is that the CCSD(T)/cc-pVDZ calculations represent the poorest calculations performed in the present study—with a maximum absolute error of 66 kJ mol⁻¹, the results are completely unacceptable. The triple-zeta basis performs considerably better, but its accuracy is only roughly the same as for the G2 and DFT calculations. The cc-pVQZ basis performs significantly better and excellent results are obtained with the cc-pV5Z basis.

Table 7. Electronic contributions to reaction enthalpies (in kJ mol^{-1}) as obtained from CCSD(T)/cc-pVXZ calculations.

Reaction	Ref. ^a	Cardinal number X					
		∞^b	6	5	4	3	2
$\text{CH}_2^{\text{g}} + \text{H}_2 \rightarrow \text{CH}_4$	-544(2)	-543	-543	-542	-542	-540	-532
$\text{C}_2\text{H}_2 + \text{H}_2 \rightarrow \text{C}_2\text{H}_4$	-203(2)	-206	-205	-205	-205	-206	-213
$\text{C}_2\text{H}_2 + 3\text{H}_2 \rightarrow 2\text{CH}_4$	-446(2)	-447	-448	-448	-449	-451	-457
$\text{N}_2\text{H}_2 \rightarrow \text{N}_2 + \text{H}_2$		-173	-173	-173	-175	-179	-189
$\text{CO} + \text{H}_2 \rightarrow \text{H}_2\text{CO}$	-21(1)	-23	-23	-22	-20	-17	-3
$\text{N}_2 + 3\text{H}_2 \rightarrow 2\text{NH}_3$	-164(1)	-165	-165	-165	-160	-148	-103
$\text{F}_2 + \text{H}_2 \rightarrow 2\text{HF}$	-563(1)	-565	-564	-564	-560	-546	-511
$\text{O}_3 + 3\text{H}_2 \rightarrow 3\text{H}_2\text{O}$	-933(2)	-946	-946	-945	-938	-916	-854
$\text{H}_2\text{CO} + 2\text{H}_2 \rightarrow \text{CH}_4 + \text{H}_2\text{O}$	-251(1)	-250	-250	-250	-248	-242	-218
$\text{H}_2\text{O}_2 + \text{H}_2 \rightarrow 2\text{H}_2\text{O}$	-365(2)	-362	-364	-366	-363	-352	-329
$\text{CO} + 3\text{H}_2 \rightarrow \text{CH}_4 + \text{H}_2\text{O}$	-272(1)	-273	-273	-272	-269	-259	-221
$\text{HCN} + 3\text{H}_2 \rightarrow \text{CH}_4 + \text{NH}_3$	-320(3)	-321	-322	-321	-319	-314	-290
$\text{HNO} + 2\text{H}_2 \rightarrow \text{H}_2\text{O} + \text{NH}_3$	-444(1)	-446	-446	-445	-441	-429	-384
$\text{HNC} \rightarrow \text{HCN}$		-63	-63	-63	-63	-63	-67
$\text{H}_2\text{O} + \text{F}_2 \rightarrow \text{HOF} + \text{HF}$	-129(4)	-118	-118	-118	-117	-115	-107
$\text{CO}_2 + 4\text{H}_2 \rightarrow \text{CH}_4 + 2\text{H}_2\text{O}$	-244(1)	-244	-245	-244	-239	-227	-178
$2\text{CH}_2^{\text{c}} \rightarrow \text{C}_2\text{H}_4$	-844(3)	-845	-843	-842	-840	-835	-820
Mean error ^d		0	-1	0	2	9	32
Mean absolute error ^d		1	1	0	2	10	36
Root-mean-square error ^d		1	1	1	3	11	41
Maximum absolute error ^d		3	2	1	5	17	66

^a Cf table 6.

^b Extrapolation to the limit of a complete basis.

^c Singlet state, $\tilde{\alpha}^1\text{A}_1$, of methylene.

^d Reactions with O_3 and HOF omitted.

In table 7, the reactions are all written such that they are exothermic and many of the reactions involve hydrogen addition. For such reactions, it appears that the CCSD(T) method exhibits a small systematic error, overestimating the exothermicity of the reactions. Therefore, the most favourable comparison with the reference values is obtained with the cc-pV5Z basis set rather than at the basis-set limit. Relative to experiment, the cc-pV6Z basis yields slightly inferior results. Consequently, the extrapolated values for $X = \infty$ are even further off, although these effects are indeed small. Since the R12 numbers are comparable with the cc-pV5Z results, they also compare very well with the reference values—even though, to some extent, this excellent agreement is fortuitous.

Whereas the error is not reduced by the extrapolation to the limit of a complete basis set based on the $X = 5$ and 6 numbers (56 extrapolation), we can indeed significantly improve the raw $X = 3$ data by performing an extrapolation based on the $X = 2$ and 3 numbers (DT extrapolation). The mean error, mean absolute error, rms error, and maximum absolute error of the DT extrapolation are 2, 4, 6 and 11 kJ mol^{-1} , respectively, to be compared with 9, 10, 11 and 17 kJ mol^{-1} for the $X = 3$ results (table 5). In view of the low costs of the $X = 2$ and 3 calculations, the successful DT extrapolation might constitute a useful computational tool (Truhlar 1998, Chuang and Truhlar 1999). Similarly, the TQ extrapolation (rms error = 2 kJ mol^{-1}) is an improvement over the direct cc-pVQZ calculation (rms error = 3 kJ mol^{-1}).

9. Conclusions

The results of our calculations are summarized in figures 6–8. The figures show the convergence of the computed enthalpies of reaction with respect to both the computational method (figure 7) and the one-electron basis set (figure 8). Results accurate to within 1 kJ mol⁻¹ are obtained at the CCSD(T)/cc-pV5Z level (using CCSD(T)(full)/cc-pCVQZ geometries and including a core correlation correction computed at the CCSD(T)/cc-pCV5Z level).

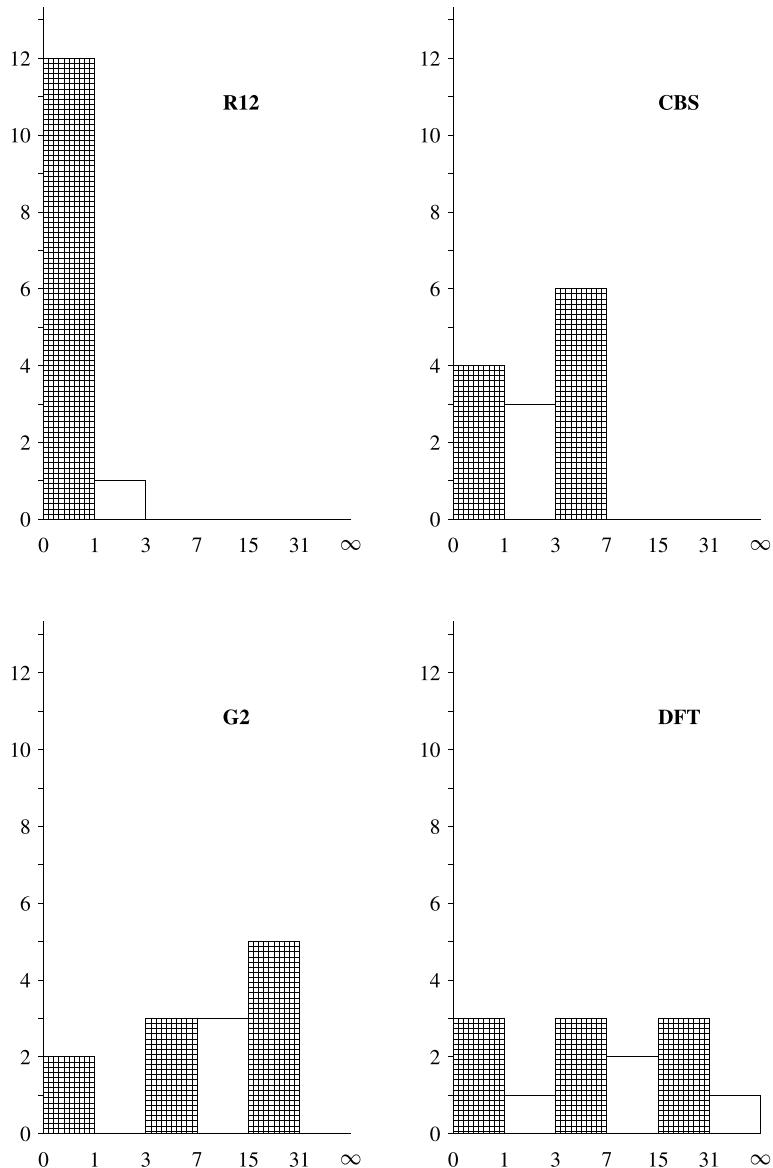


Figure 6. Histograms representing the error distribution of the computed enthalpies of reaction. The intervals are given in kJ mol⁻¹. For example, for the G2 model chemistry, five errors are found between 15 and 31 kJ mol⁻¹.

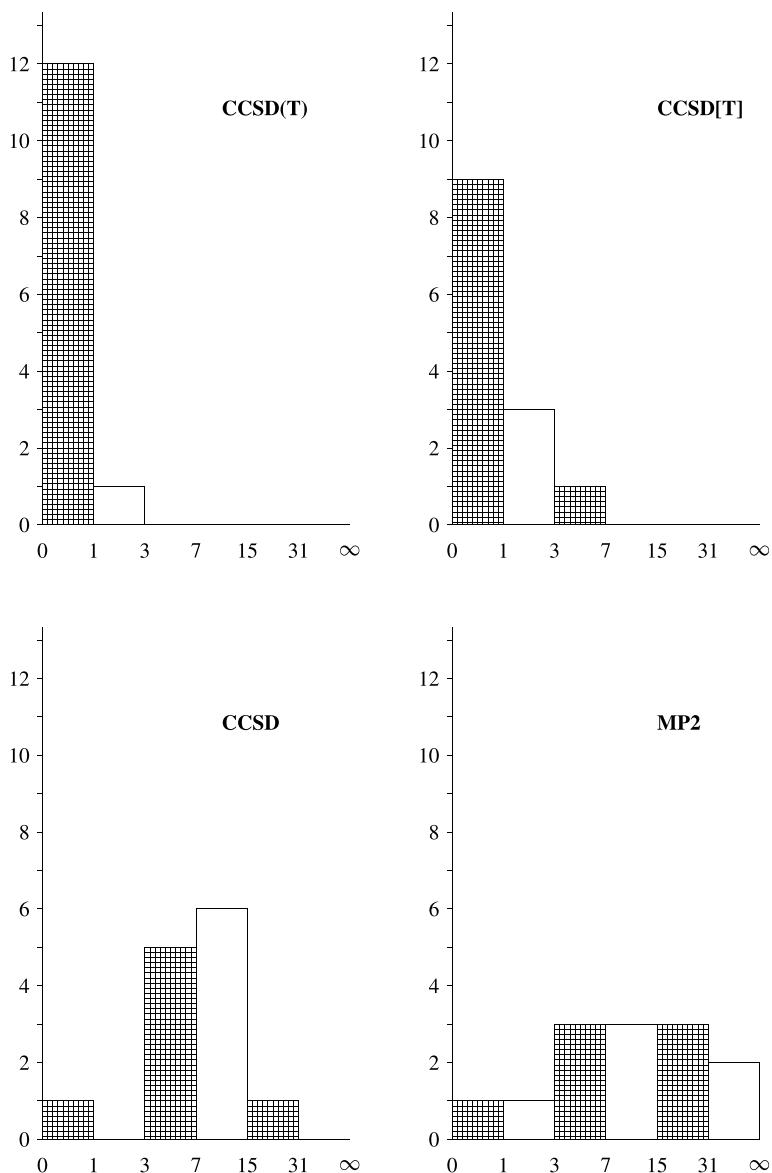


Figure 7. Histograms representing the error distribution of the computed enthalpies of reaction. The intervals are given in kJ mol^{-1} . For example, at the CCSD level, five errors are found between 3 and 7 kJ mol^{-1} . All results have been obtained from R12 calculations.

The CCSD(T)/cc-pV5Z calculations represent a proper level of description where the errors in the one- and n -electron spaces are balanced and small relative to the uncertainties in the experimental reaction enthalpies. For other properties, different levels may be appropriate. The CCSD(T)-R12 model gives a better description of the Coulomb cusp and the short-range correlation and therefore approaches the basis-set limit with smaller basis sets. For the enthalpies of isogeric reactions, there is a large cancellation of the correlation contributions for the reactants and the products since the number of paired electrons is

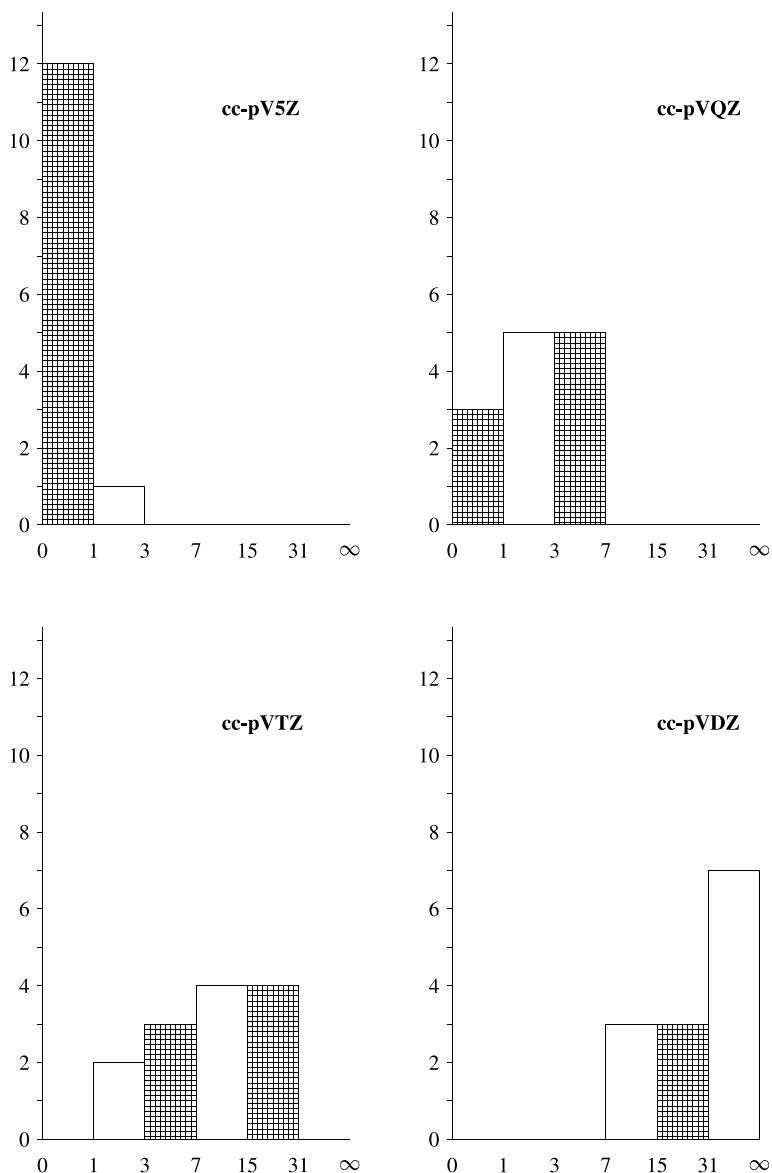


Figure 8. Histograms representing the error distribution of the computed enthalpies of reaction. The intervals are given in kJ mol⁻¹. For example, with the cc-pVQZ basis set, five errors are found between 3 and 7 kJ mol⁻¹. All results correspond to CCSD(T)(FC)/cc-pVXZ calculations supplemented with a core correlation correction obtained with the cc-pCV5Z basis set.

conserved. Because of this cancellation, accurate reaction enthalpies can be calculated using standard basis sets. In contrast, since atomizations do not conserve the number of paired electrons, accurate energies of atomization cannot be calculated using standard basis sets—for agreement with experiment, one must either carry out extrapolations or, better, employ an explicitly correlated model, in which the short-range electronic interactions are accurately described by means of the interelectronic coordinate. For the CBS, G2 and DFT methods, it

is difficult to obtain a significantly improved accuracy in the reaction enthalpies since these methods rely on empirically adjusted parameters and extrapolations. We note that the simple DT extrapolation gives results of the same accuracy as these methods. The rms error of 6 kJ mol⁻¹ of the DT extrapolation is twice the CBS error but less than half the G2 and DFT errors.

Acknowledgments

We thank Professor G A Petersson for helpful comments on the manuscript. This work has received support from the Danish Natural Sciences Research Council (grant no 9600856) and from the Research Council of Norway (Programme for Supercomputing) through grants of computing time (grant nos NN2694K and NN2794K). The research of WK has been made possible by a fellowship of the Royal Netherlands Academy of Arts and Sciences.

References

- Almlöf J and Taylor P R 1987 *J. Chem. Phys.* **86** 4070
 —— 1992 *Adv. Quantum Chem.* **22** 301
- Bakken V, Helgaker T, Klopper W and Ruud K 1999 *Mol. Phys.* **96** 653
- Bartlett R J 1995 *Modern Electronic Structure Theory* ed D R Yarkony (Singapore: World Scientific) p 1047
- Becke A D 1993 *J. Chem. Phys.* **98** 5648
- Bukowski R, Jeziorski B and Szalewicz K 1996 *J. Chem. Phys.* **104** 3306
 —— 1999 *J. Chem. Phys.* **110** 4165
- Bunge C F 1970 *Theor. Chim. Acta* **16** 126
- Byron F W Jr and Joachain C J 1967 *Phys. Rev.* **157** 1
- Carroll D P, Silverstone H J and Metzger R M 1979 *J. Chem. Phys.* **71** 4142
- Cencek W and Rychlewski J 1993 *J. Chem. Phys.* **98** 1252
 —— 1995 *J. Chem. Phys.* **102** 2533
- Chase M W Jr 1998 *NIST-JANAF Thermochemical Tables* 4th edn (*J. Phys. Chem. Ref. Data Monograph* vol 9) (Woodbury, NY: American Institute of Physics)
- Chuang Y and Truhlar D G 1999 *J. Phys. Chem. A* **103** 651
- Császár A G, Allen W D and Schaefer H F III 1998 *J. Chem. Phys.* **108** 9751
- Curtiss L A, Carpenter J E, Raghavachari K and Pople J A 1992 *J. Chem. Phys.* **96** 9030
- Curtiss L A, Jones C, Trucks G W, Raghavachari K and Pople J A 1990 *J. Chem. Phys.* **93** 2537
- Curtiss L A and Raghavachari K 1995 *Quantum Mechanical Electronic Structure Calculations with Chemical Accuracy* ed S R Langhoff (Dordrecht: Kluwer) p 139
- Curtiss L A, Raghavachari K and Pople J A 1993 *J. Chem. Phys.* **98** 1293
- Curtiss L A, Raghavachari K, Redfern P C and Pople J A 1997 *J. Chem. Phys.* **106** 1063
- Curtiss L A, Raghavachari K, Redfern P C, Rassolov V and Pople J A 1998c *J. Chem. Phys.* **109** 7764
- Curtiss L A, Raghavachari K, Redfern P C and Stefanov B B 1998a *J. Chem. Phys.* **108** 692
- Curtiss L A, Raghavachari K, Trucks G W and Pople J A 1991 *J. Chem. Phys.* **94** 7221
- Curtiss L A, Redfern P C, Raghavachari K and Pople J A 1998b *J. Chem. Phys.* **109** 42
- Dunning T H Jr 1989 *J. Chem. Phys.* **90** 1007
- Feller D 1992 *J. Chem. Phys.* **96** 6104
 —— 1993 *J. Chem. Phys.* **98** 7059
- Feller D and Peterson K A 1998 *J. Chem. Phys.* **108** 154
- Flores J R 1992 *Phys. Rev. A* **46** 6063
- Flores J R and Redondo P 1993 *J. Phys. B: At. Mol. Opt. Phys.* **26** 2251
- Frisch M J *et al* 1995 *Gaussian 94, Revision E.2* (Pittsburgh PA: Gaussian Inc)
- Gdanitz R J 1993 *Chem. Phys. Lett.* **210** 253
 —— 1998a *Chem. Phys. Lett.* **283** 253
 —— 1998b *Chem. Phys. Lett.* **288** 590 (erratum)
 —— 1998c *Chem. Phys. Lett.* **295** 540 (erratum)
 —— 1998d *J. Chem. Phys.* **109** 9795
 —— 1999 *J. Chem. Phys.* **110** 706

- Gdanitz R J and Röhse R 1995 *Int. J. Quantum Chem.* **55** 147
—1996 *Int. J. Quantum Chem.* **59** 505 (erratum)
- Halkier A, Helgaker T, Jørgensen P, Klopper W, Koch H, Olsen J and Wilson A K 1998 *Chem. Phys. Lett.* **286** 243
- Halkier A, Koch H, Christiansen O, Jørgensen P and Helgaker T 1997 *J. Chem. Phys.* **107** 849
- Helgaker T, Jørgensen P and Olsen J 1999 unpublished
- Helgaker T *et al* 1997a DALTON, an ab initio electronic structure program, Release 1.0
- Helgaker T, Klopper W, Koch H and Noga J 1997b *J. Chem. Phys.* **106** 9639
- Helgaker T and Taylor P R 1995 *Modern Electronic Structure Theory* ed D R Yarkony (Singapore: World Scientific) p 725
- Hill R N 1985 *J. Chem. Phys.* **83** 1173
- Jankowski K and Malinowski P 1980 *Phys. Rev. A* **21** 45
- Kendall R A, Dunning T H Jr and Harrison R J 1992 *J. Chem. Phys.* **96** 6796
- Klopper W 1991 *Chem. Phys. Lett.* **186** 583
—1995 *J. Chem. Phys.* **102** 6168
—1998 *The Encyclopedia of Computational Chemistry* ed P v R Schleyer *et al* (Chichester: Wiley) p 2351
- Klopper W and Almlöf J 1993 *J. Chem. Phys.* **99** 5167
- Klopper W and Helgaker T 1998 *Theor. Chem. Acc.* **99** 265
- Klopper W and Kutzelnigg W 1987 *Chem. Phys. Lett.* **134** 17
—1991 *J. Chem. Phys.* **94** 2020
- Klopper W and Lüthi H P 1996 *Chem. Phys. Lett.* **262** 546
—1999 *Mol. Phys.* **96** 559
- Klopper W, Lüthi H P, Brupbacher Th and Bauder A 1994 *J. Chem. Phys.* **101** 9747
- Klopper W and Noga J 1995 *J. Chem. Phys.* **103** 6127
- Klopper W, Röhse R and Kutzelnigg W 1991 *Chem. Phys. Lett.* **178** 455
- Klopper W, Van Duijneveldt-van de Rijdt J G C M and Van Duijneveldt F B unpublished
- Komasa J and Rychlewski J 1997 *Mol. Phys.* **91** 909
- Korona T, Williams H L, Bukowski R, Jeziorski B and Szalewicz K 1997 *J. Chem. Phys.* **106** 5109
- Kutzelnigg W 1985 *Theor. Chim. Acta* **68** 445
- Kutzelnigg W and Klopper W 1991 *J. Chem. Phys.* **94** 1985
- Kutzelnigg W and Morgan J D III 1992a *J. Chem. Phys.* **96** 4484
—1992b *J. Chem. Phys.* **97** 8821 (erratum)
- Lee C, Yang W and Parr R G 1988 *Phys. Rev. B* **37** 785
- Lindgren I and Salomonson S 1980 *Phys. Scr.* **21** 335
- Mårtensson-Pendrill A-M, Alexander S A, Adamowicz L, Oliphant N, Olsen J, Öster P, Quiney H M, Salomonson S and Sundholm D 1991 *Phys. Rev. A* **43** 3355
- Martin J M L 1996 *Chem. Phys. Lett.* **259** 669
—1997a *Theor. Chem. Acc.* **97** 227
—1997b *J. Mol. Struct. (Theochem.)* **398** 135
—1997c *Chem. Phys. Lett.* **273** 98
- Martin J M L and Taylor P R 1997 *J. Chem. Phys.* **106** 8620
- Montgomery J A Jr, Ochterski J W and Petersson G A 1994 *J. Chem. Phys.* **101** 5900
- Montgomery J A Jr, Petersson G A and Matsunaga N 1989 *Chem. Phys. Lett.* **155** 413
- Noga J and Klopper W 1995 DIRCCR12, a coupled-cluster program (unpublished)
- Noga J, Klopper W and Kutzelnigg W 1997 *Recent Advances in Coupled-Cluster Methods* ed R J Bartlett (Singapore: World Scientific) p 1
- Noga J and Kutzelnigg W 1994 *J. Chem. Phys.* **101** 7738
- Noga J, Kutzelnigg W and Klopper W 1992 *Chem. Phys. Lett.* **199** 497
- Noga J, Tunega D, Klopper W and Kutzelnigg W 1995 *J. Chem. Phys.* **103** 309
- Nyden M R and Petersson G A 1981 *J. Chem. Phys.* **75** 1843
- Ochterski J W, Petersson G A and Montgomery J A Jr 1996 *J. Chem. Phys.* **104** 2598
- Ochterski J W, Petersson G A and Wiberg K B 1995 *J. Am. Chem. Soc.* **117** 11 299
- Persson B J and Taylor P R 1996 *J. Chem. Phys.* **105** 5915
- Petersson G A and Al-Laham M A 1991 *J. Chem. Phys.* **94** 6081
- Petersson G A, Bennett A, Tensfeldt T G, Al-Laham M A, Shirley W A and Mantzaris J 1988 *J. Chem. Phys.* **89** 2193
- Petersson G A and Braunstein M 1985 *J. Chem. Phys.* **83** 5129
- Petersson G A and Licht S L 1981 *J. Chem. Phys.* **75** 4556
- Petersson G A, Malick D K, Wilson W G, Ochterski J W, Montgomery J A Jr and Frisch M J 1998 *J. Chem. Phys.* **109** 10 570

- Petersson G A and Nyden M R 1981 *J. Chem. Phys.* **75** 3423
Petersson G A, Tensfeldt T G and Montgomery J A Jr 1991 *J. Chem. Phys.* **94** 6091
Petersson G A, Yee A K and Bennett A 1985 *J. Chem. Phys.* **83** 5105
Pople J A, Head-Gordon M, Fox D J, Raghavachari K and Curtiss L A 1989 *J. Chem. Phys.* **90** 5622
Pople J A, Head-Gordon M and Raghavachari K 1987 *J. Chem. Phys.* **87** 5968
Purvis G D and Bartlett R J 1982 *J. Chem. Phys.* **76** 1910
Rablen P R, Lockman J W and Jorgensen W L 1998 *J. Phys. Chem. A* **102** 3782
Raghavachari K and Curtiss L A 1995 *Modern Electronic Structure Theory* ed D R Yarkony (Singapore: World Scientific) p 991
Raghavachari K, Trucks G W, Pople J A and Head-Gordon M 1989 *Chem. Phys. Lett.* **157** 479
Sasaki F and Yoshimine M 1974 *Phys. Rev. A* **9** 17
Schäfer A, Horn H and Ahlrichs R 1992 *J. Chem. Phys.* **97** 2571
Schäfer A, Huber C and Ahlrichs R 1994 *J. Chem. Phys.* **100** 5829
Schwartz C 1962 *Phys. Rev.* **126** 1015
—1963 *Methods in Computational Physics* ed B Alder *et al* (New York: Academic) p 241
Stanton J F, Gauss J, Watts J D, Lauderdale W J and Bartlett R J ACESII, *an ab initio program system*
Tarczay G, Császár A G, Klopper W, Szalay V, Allen W D and Schaefer H F III 1999 *J. Chem. Phys.* at press
Termath V, Klopper W and Kutzelningg W 1991 *J. Chem. Phys.* **94** 2002
Truhlar D G 1998 *Chem. Phys. Lett.* **294** 45
Williams H L, Korona T, Bukowski R, Jeziorski B and Szalewicz K 1996 *Chem. Phys. Lett.* **262** 431
Wilson A K and Dunning T H Jr 1997 *J. Chem. Phys.* **106** 8718
Wilson A K, van Mourik T and Dunning T H Jr 1996 *J. Mol. Struct. (Theochem.)* **388** 339
Woon D E and Dunning T H Jr 1993 *J. Chem. Phys.* **98** 1358
—1994 *J. Chem. Phys.* **100** 2975
—1995 *J. Chem. Phys.* **103** 4572
Woon D E, Peterson K A and Dunning T H Jr 1998 *J. Chem. Phys.* **109** 2233