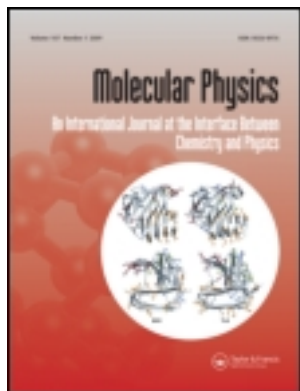


This article was downloaded by: [Australian National University]

On: 07 April 2013, At: 16:24

Publisher: Taylor & Francis

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Molecular Physics: An International Journal at the Interface Between Chemistry and Physics

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/tmph20>

Assessment of basis sets for F12 explicitly-correlated molecular electronic-structure methods

Florian A. Bischoff^a, Sandra Wolfsegger^a, David P. Tew^a & Wim Klopper^a

^a Lehrstuhl für Theoretische Chemie, Institut für Physikalische Chemie, Universität Karlsruhe (TH), D-76128 Karlsruhe, Germany

Version of record first published: 07 Oct 2010.

To cite this article: Florian A. Bischoff, Sandra Wolfsegger, David P. Tew & Wim Klopper (2009): Assessment of basis sets for F12 explicitly-correlated molecular electronic-structure methods, *Molecular Physics: An International Journal at the Interface Between Chemistry and Physics*, 107:8-12, 963-975

To link to this article: <http://dx.doi.org/10.1080/00268970802708942>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.tandfonline.com/page/terms-and-conditions>

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

INVITED ARTICLE

Assessment of basis sets for F12 explicitly-correlated molecular electronic-structure methods

Florian A. Bischoff, Sandra Wolfsegger, David P. Tew and Wim Klopper*

*Lehrstuhl für Theoretische Chemie, Institut für Physikalische Chemie,
Universität Karlsruhe (TH), D-76128 Karlsruhe, Germany*

(Received 7 November 2008; final version received 19 December 2008)

One-electron basis sets for F12 explicitly-correlated molecular electronic-structure methods are assessed by analysing the accuracy of Hartree–Fock energies and valence-only second-order correlation energies of a test set of 106 small molecules containing the atoms H, C, N, O and F. For these molecules, near Hartree–Fock-limit energies and benchmark second-order correlation energies (accurate to within 99.95% of the basis-set limit) are provided. Absolute energies are analysed as well as the Hartree–Fock and second-order correlation contributions to the atomisation energies of the molecules. Standard basis sets such as the Karlsruhe def2-TZVPP and def2-QZVPP sets and the augmented correlation-consistent polarised valence X -tuple zeta (aug-cc-pVXZ, $X = D, T, Q, 5$) sets are compared with the specialised cc-pVXZ-F12 ($X = D, T, Q$) sets that were recently optimised by Peterson and co-workers [J. Chem. Phys. **128**, 084102 (2008)] for use in F12 theory. The results obtained from F12 explicitly-correlated molecular electronic-structure calculations are compared with those that are obtained by standard electronic-structure calculations followed by basis-set extrapolation based on the X^{-3} convergence behaviour of the aug-cc-pVXZ basis sets. The most important conclusions are that the cc-pVXZ-F12 sets are the preferred basis sets for F12 theory and that the X^{-3} extrapolation from the aug-cc-pVQZ and aug-cc-pV5Z is slightly more accurate than F12 theory in the cc-pVTZ-F12 basis but less accurate than F12 theory in the cc-pVQZ-F12 basis.

Keywords: basis sets; explicitly-correlated wave functions; F12 methods; atomisation energies

1. Introduction

F12 explicitly-correlated molecular electronic-structure methods are emerging as practical tools for computational chemistry. These methods overcome the extremely slow convergence of orbital expansions with the basis size by using additional geminal basis functions, chosen to closely approximate the form of the correlation hole at short interelectronic distances. In the early days, high-end electronic-structure calculations focused on the complete-basis-set limit, using very large one-electron basis sets [1]. Examples include the calculation of the heat of formation of NCO [2], the calculation of barriers to linearity of small molecules such as H₂O [3,4] and SiC₂ [5], the calculation of anharmonic potential energy surfaces of molecules such as SiH₃⁻ [5], BH₅ [6], NH₃ [7] and H₃O⁺ [8], or the calculation of the water dimer [9,10] and water···H₂ [11,12] potentials in the complete-basis-set limit of subchemical accuracy. The requirement for large basis sets in combination with geminals linear in the interelectronic distances strongly restrained applications to larger chemical systems, although at the MP2

level, systems such as [10]annulene [13], ferrocene [14] and the benzene dimer [15,16] could be treated.

There have been many advances in F12 theory over the last six years (see Ref. [17] for a good review) and it is now possible to use the standard orbital basis sets of quantum chemistry [18], for example the def2-QZVPP basis, to obtain ~99.5% of the basis set limit correlation energy for the MP2 [19–21], CCSD [22–34] and CASPT2 [35] methods. Indeed, MP2-R12 calculations have recently been used to reduce basis set incompleteness errors in CCSD(T) energy differences to obtain highly accurate predictions of π -stacking and hydrogen bonding interactions of the nucleobase mimic 2-pyridone with fluorobenzenes [36,37]. A similar MP2-F12/def2-QZVPP approach has been used to compute the heat of formation of the HOSO₂ radical, an important intermediate in the atmospheric oxidation of SO₂ to SO₃, to an accuracy of 3 kJ mol⁻¹ [38]. Further examples include the calculation of the thermochemistry of the HOSO₂ + O₂ reaction and heat of formation of the HOSO₄ radical [39], calculations on rotational conformers of the alkanols n -propanol

*Corresponding author. Email: klopper@chem-bio.uni-karlsruhe.de

through *n*-pentanol [40], the computation of interactions between CO₂ and N-containing organic heterocycles [41], and the computation of hydrogenation energies of benzene and naphthalene [42], for which chemical accuracy was obtained. MP2-F12/def2-QZVPP calculations have been combined with CCSD(T) calculations to construct a potential energy surface for malonaldehyde, a prototype for hydrogen transfer reactions, which was used to compute the tunnelling splitting to within 1 cm⁻¹ of observation [43]. F12 methods have also been successfully combined with localisation techniques to approach linear scaling, which has permitted studies of enzyme reactions in a QM/MM framework [44,45].

Despite the continuing success of F12 methods, as yet fairly little is known about the relative performance of F12 calculations with the various orbital sets. Indeed, standard basis sets are far from optimal. Functions with a large angular momentum quantum number and large exponents are not necessary in F12 calculations, but diffuse functions are more important than in conventional methods [46]. Recently, Peterson and co-workers have constructed orbital basis sets for H, He, B–Ne, and Al–Ar, specifically optimised for F12 methods [47]. They report that their cc-pVXZ-F12 basis sets provide total and relative correlation energies equivalent to that obtained using aug-cc-pV(X+1)Z basis sets for MP2-F12 calculations. However, the number of functions for each angular momentum quantum number is not the same in the cc-pVXZ-F12 and aug-cc-pVXZ sets—overall the F12 sets are substantially larger—and it is not yet clear whether these new sets represent a saving in computational effort.

In this work, we provide frozen core (fc) MP2 correlation energies, converged to within 99.95% of the basis set limit, and use them to assess the new F12 basis sets together with the correlation-consistent aug-cc-pVXZ sets of Dunning and co-workers [48–50] and the Karlsruhe basis sets def2-TZVPP and def2-QZVPP [51]. This very high accuracy of 99.95% is necessary to rigorously assess F12 methods. We only consider the elements H, C, N, O and F. For our assessment set, we choose the set of 105 molecules compiled by Bakowies [52], to which we add H₂. This set represents a wide range of bonding situations for these five elements and is intended as a general testing set for F12 methods. We have used this same set to develop explicitly-correlated model chemistries, combining F12 calculations with post-CCSD(T) correlation contributions and relativistic corrections, aiming at chemical accuracy [53]. The structures for the 106 molecules have been optimised at the all-electron CCSD(T)/cc-pCVTZ level of theory and are reported

in the Supplementary Information which is available via the multimedia link on the article webpage.

This article is organised as follows: In Section 2, we present the benchmark fc-MP2 correlation energies, giving the details of how they were computed. In Section 3 we analyse the remaining basis set error and the errors arising from the resolution of the identity approximations intrinsic to the F12 methods and the density fitting employed in our program. In Section 4, we use our benchmark data to assess the performance of F12 and standard orbital basis sets when used in MP2-F12 calculations for both absolute and relative energies.

2. MP2 basis set limit

Since fc-MP2-F12 calculations with quadruple- ζ quality basis sets recover $\sim 99.5\%$ of the basis set limit correlation energy, it is necessary to compute the fc-MP2 limit to better than 99.95% to compare the performance of different quadruple- ζ sets. In Table 1 we present our reference fc-MP2 basis set limit correlation energies for the 106 molecules and the four atoms C, N, O and F in our test set. We also report the near basis set limit HF energies obtained using our decontracted aug-cc-pV6Z basis. The correlation energies are converged to within the required 99.95%. The 0.05% (0.05 kJ mol⁻¹ per valence electron) deviation from the basis set limit contains a variety of contributions, all of which are discussed and quantified in Section 3.

To achieve 99.95% accuracy, we use the MP2-F12 method with a decontracted aug-cc-pV6Z orbital basis. All primitives with exponents of less than 1000 a_0^{-2} were used as decontracted basis functions and contractions involving primitives with higher exponents were retained to avoid problems in density fitting. The MP2-F12 calculations were performed using a correlation factor of six Gaussian geminals, fitted to an exponential with exponent 1.4 a_0^{-1} [46,54], and the F12 amplitudes were optimised using the orbital invariant method [55]. Ansatz 2 and approximation B were used and the commutator integrals involving the kinetic energy were computed from the matrix representation of the core Hamiltonian. The auxiliary basis set for the resolution-of-the-identity (RI) approximation intrinsic to the MP2-F12 method was taken to be the union of the orbital basis and a complementary auxiliary basis (CABS) [56], which was chosen to be the aug-cc-pV5Z MP2-fitting basis (aug-cc-pV5Z cbas in TURBOMOLE jargon, *vide infra*) [57]. For the open-shell calculations, the ROHF wave function was used and UMP2-F12 calculations were performed using semi-canonical orbitals.

Table 1. Molecules in the test set and their reference near basis-set limit Hartree–Fock (HF) and MP2 frozen-core correlation energies (in E_h). B–A denotes the energy difference (in E_h) between the MP2-F12/2B and MP2-F12/2A frozen-core correlation energies.

No.	Molecule		HF energy	MP2 contribution	B–A
1	CFN	Cyanogen fluoride	-191.7869115	-0.6734136	0.0003255
2	CFN	Isocyanogen fluoride	-191.6701799	-0.6681564	0.0003199
3	CF ₂	Singlet difluoromethylene	-236.7797603	-0.7396863	0.0003906
4	CF ₂ O	Carbonyl fluoride	-311.7690426	-1.0179198	0.0005292
5	CF ₄	Tetrafluoromethane	-435.8611009	-1.3480501	0.0007405
6	CHF	Singlet fluoromethylene	-137.8254451	-0.4486484	0.0002237
7	CHFO	Formyl fluoride	-212.8538870	-0.7357389	0.0003595
8	CHF ₃	Trifluoromethane	-336.9423500	-1.0664251	0.0005611
9	CHN	Hydrogen cyanide	-92.9152377	-0.3871352	0.0001662
10	CHN	Hydrogen isocyanide	-92.8999603	-0.3740514	0.0001622
11	CHNO	Cyanic acid	-167.8048239	-0.6544582	0.0002937
12	CHNO	Isocyanic acid	-167.8447501	-0.6553452	0.0002937
13	CHNO	Formonitrile oxide	-167.7071831	-0.6848635	0.0002934
14	CHNO	Isofulminic acid	-167.7092725	-0.6482486	0.0002892
15	CH ₂	Singlet methylene	-38.8959361	-0.1559390	0.0000626
16	CH ₂ F ₂	Difluoromethane	-238.0207580	-0.7832349	0.0003889
17	CH ₂ N ₂	Cyanamide	-147.9793717	-0.6195516	0.0002591
18	CH ₂ N ₂	3H-Diazirine	-147.8975793	-0.6301166	0.0002591
19	CH ₂ N ₂	Diazomethane	-147.9120691	-0.6318115	0.0002512
20	CH ₂ O	Formaldehyde	-113.9230386	-0.4485051	0.0001954
21	CH ₂ O	Hydroxymethylene	-113.8468195	-0.4358324	0.0001911
22	CH ₂ O ₂	Dioxirane	-188.6951256	-0.7335711	0.0003291
23	CH ₂ O ₂	Formic acid	-188.8600034	-0.7188020	0.0003283
24	CH ₂ O ₃	Performic acid	-263.6387395	-0.9930592	0.0004664
25	CH ₃ F	Fluoromethane	-139.1097750	-0.5001195	0.0002245
26	CH ₃ N	Methanimine	-94.0762519	-0.4159873	0.0001582
27	CH ₃ NO	Formamide	-169.0186910	-0.6846840	0.0002934
28	CH ₃ NO ₂	Methyl nitrite	-243.7812749	-0.9679194	0.0004309
29	CH ₃ NO ₂	Nitromethane	-243.7809017	-0.9795674	0.0004314
30	CH ₄	Methane	-40.2170458	-0.2190333	0.0000639
31	CH ₄ N ₂ O	Urea	-224.1021999	-0.9176184	0.0004168
32	CH ₄ O	Methanol	-115.1017391	-0.4851798	0.0001915
33	CH ₅ N	Methylamine	-95.2629113	-0.4525078	0.0001566
34	CO	Carbon monoxide	-112.7902518	-0.4044170	0.0001939
35	CO ₂	Carbon dioxide	-187.7244395	-0.6867786	0.0003296
36	C ₂ F ₂	Difluoroacetylene (fixed)	-274.5945125	-0.9134635	0.0004456
37	C ₂ F ₄	Tetrafluoroethylene	-473.6525592	-1.5101832	0.0008251
38	C ₂ HF	Fluoroacetylene	-175.7291423	-0.6298874	0.0003019
39	C ₂ HF ₃	Trifluoroethylene	-374.7602074	-1.2255182	0.0006401
40	C ₂ H ₂	Acetylene	-76.8552215	-0.3454401	0.0001424
41	C ₂ H ₂ F ₂	1,1-Difluoroethylene	-275.8774293	-0.9420319	0.0004633
42	C ₂ H ₂ O	Ketene	-151.7979880	-0.6088949	0.0002631
43	C ₂ H ₂ O	Oxirene	-151.6575846	-0.6249065	0.0002747
44	C ₂ H ₂ O ₂	Glyoxal	-226.6972470	-0.8715650	0.0004052
45	C ₂ H ₃ F	Fluoroethylene	-176.9713289	-0.6574949	0.0002945
46	C ₂ H ₃ FO	Acetyl fluoride	-251.9204472	-0.9258834	0.0004359
47	C ₂ H ₃ N	Acetonitrile	-131.9837319	-0.5761492	0.0002317
48	C ₂ H ₃ N	Methyl isocyanide	-131.9526934	-0.5641634	0.0002287
49	C ₂ H ₄	Ethylene	-78.0705663	-0.3726636	0.0001305
50	C ₂ H ₄ O	Acetaldehyde	-152.9890345	-0.6397856	0.0002658
51	C ₂ H ₄ O	Oxirane	-152.9404392	-0.6484923	0.0002676
52	C ₂ H ₄ O ₂	Acetic acid	-227.9243915	-0.9094852	0.0004037
53	C ₂ H ₄ O ₂	Methyl formate	-227.8981132	-0.9076775	0.0004029
54	C ₂ H ₅ F	Fluoroethane	-178.1675871	-0.6917162	0.0002950

(continued)

Table 1. Continued.

No.	Molecule	HF energy	MP2 contribution	B-A	
55	C ₂ H ₅ N	Aziridine	-133.1016627	-0.6164873	0.0002303
56	C ₂ H ₆	Ethane	-79.2666426	-0.4096202	0.0001304
57	C ₂ H ₆ O	Dimethyl ether	-154.1409915	-0.6737526	0.0002640
58	C ₂ H ₆ O	Ethanol	-154.1578815	-0.6775961	0.0002612
59	C ₂ N ₂	Cyanogen	-184.6613685	-0.7549829	0.0003278
60	C ₃ H ₃ N	Acrylonitrile	-169.8386596	-0.7337807	0.0002993
61	C ₃ H ₄	Allene	-115.9147707	-0.5310153	0.0001983
62	C ₃ H ₄	Cyclopropene	-115.8742896	-0.5414329	0.0002050
63	C ₃ H ₄	Propyne	-115.9172745	-0.5363030	0.0002073
64	C ₃ H ₆	Cyclopropane	-117.1107221	-0.5733140	0.0002021
65	C ₃ H ₆	Propene	-117.1261419	-0.5657594	0.0001989
66	C ₃ H ₈	Propane	-118.3175998	-0.6026590	0.0001994
67	C ₃ O ₂	Carbon suboxide (fixed)	-263.3970576	-1.0186298	0.0004355
68	C ₄ H ₄	Butatriene	-153.7576701	-0.6937222	0.0002723
69	C ₄ H ₄	Cyclobutadiene	-153.7077028	-0.7045815	0.0002847
70	C ₄ H ₄	Tetrahydrane	-153.6640182	-0.7165988	0.0002872
71	C ₄ H ₄	Vinylacetylene	-153.7756023	-0.6944420	0.0002742
72	C ₄ N ₂	Dicyanoacetylene (fixed)	-260.3699559	-1.0809805	0.0004247
73	FH	Hydrogen fluoride	-100.0707863	-0.3197419	0.0001559
74	FHO	Hypofluorous acid	-174.8227081	-0.5880740	0.0002847
75	FHO ₂	Fluoroperoxide	-249.6239480	-0.8653412	0.0004190
76	FH ₂ N	Monofluoroamine	-155.0409719	-0.5476942	0.0002515
77	FH ₃ N ₂	Fluorohydrazine	-210.0684626	-0.7849995	0.0003509
78	FNO	Nitrosyl fluoride	-228.7304686	-0.8025419	0.0003847
79	F ₂	Difluorine	-198.7731963	-0.6118360	0.0003137
80	F ₂ N ₂	Difluorodiazene (cis)	-307.7266920	-1.0469854	0.0005221
81	F ₂ N ₂	Difluorodiazene (trans)	-307.7288330	-1.0425389	0.0005257
82	F ₂ O	Difluorine monoxide	-273.5865780	-0.8838321	0.0004496
83	F ₂ O ₂	Perfluoroperoxide	-348.3742509	-1.1833499	0.0005845
84	F ₃ N	Trifluoroamine	-352.7142390	-1.1280514	0.0005922
85	HNO	Nitrosylhydride	-129.8492567	-0.4962640	0.0002205
86	HNO ₂	Nitrous acid (cis)	-204.7423915	-0.7760086	0.0003540
87	HNO ₂	Nitrous acid (trans)	-204.7415982	-0.7773909	0.0003529
88	HNO ₂	Nitrous acid, H-NO ₂	-204.7219722	-0.7922894	0.0003530
89	HNO ₃	Nitric acid	-279.5864471	-1.0623180	0.0004918
90	HN ₃	Hydrogen azide	-163.9172349	-0.6876363	0.0002793
91	H ₂ N ₂	Diazene (cis)	-110.0396202	-0.4608058	0.0001852
92	H ₂ N ₂	Diazene (trans)	-110.0491822	-0.4603129	0.0001849
93	H ₂ N ₂	Diazene (iso)	-110.0206290	-0.4448095	0.0001875
94	H ₂ N ₂ O	Nitrosamide	-184.9201527	-0.7368870	0.0003205
95	H ₂ O	Water	-76.0673059	-0.3006048	0.0001233
96	H ₂ O ₂	Hydrogen peroxide	-150.8517704	-0.5701178	0.0002531
97	H ₃ N	Ammonia	-56.2247904	-0.2646166	0.0000894
98	H ₃ NO	Ammonia oxide	-131.0171087	-0.5309810	0.0002275
99	H ₃ NO	Hydroxylamine	-131.0554068	-0.5317602	0.0002187
100	H ₄ N ₂	Hydrazine	-111.2361412	-0.4961172	0.0001834
101	N ₂	Dinitrogen	-108.9925134	-0.4215422	0.0001879
102	N ₂ O	Nitrous oxide	-183.7656610	-0.7258487	0.0003154
103	N ₂ O ₃	Dinitrogen trioxide	-333.3899231	-1.2928074	0.0005934
104	N ₂ O ₄	Dinitrogen tetraoxide	-408.2303052	-1.5951352	0.0007361
105	O ₃	Ozone	-224.3637281	-0.8810446	0.0003830
106	H ₂	Dihydrogen	-1.1336066	-0.0342510	0.0000066
107	C	Carbon	-37.6886122	-0.0818513	0.0000417
108	N	Nitrogen	-54.4009236	-0.1137928	0.0000535
109	O	Oxygen	-74.8093817	-0.1792499	0.0001047
110	F	Fluorine	-99.4093241	-0.2477403	0.0001465

All calculations have been performed with the `ricc2` module [58] of the TURBOMOLE program package [59], which uses density fitting for all four-centre two-electron integrals [60]. The details of the RI-MP2 approach can be found in Refs. [61,62] and the details of the RI-MP2-F12 implementation in Refs. [40,63]. The MP2-F12 equations contain the Fock matrix in the unified orbital-plus-CABS space and six kinds of four-centre two-electron integrals involving the Coulomb operator, the correlation factor, or both together. The density fitting for the Coulomb and exchange contributions to the Fock matrix (RI-JK approximation) was performed using the aug-cc-pV5Z RI-JK-fitting auxiliary basis (aug-cc-pV5Z jkbas in TURBOMOLE jargon) of Weigend [64,65] and the density fitting for the remaining four-centre integrals was performed using a specially constructed auxiliary basis. Taking the aug-cc-pV6Z MP2-fitting auxiliary basis as a starting point, for each angular quantum number, additional functions were introduced with exponents bisecting those of the aug-cc-pV6Z basis. Furthermore, for each angular quantum number, three steep functions and three diffuse functions were added. The new functions extend the range of exponents represented in an even tempered way, using the highest, or lowest, two exponents to determine the series. In MP2-F12 theory, the density fitting basis must be suitable for both the orbital space and the CABS space. This extensive basis was thus necessary and ensured that the B matrix was positive definite for all but three molecules, difluoroacetylene (No. 36), carbon suboxide (No. 67) and dicyanoacetylene (No. 72). For these molecules the MP2-F12 energy was computed using F12 amplitudes predetermined by the coalescence conditions [66].

3. Error analysis

To obtain fc-MP2-F12 correlation energies very close to the basis set limit, one must obtain convergence with respect to the orbital basis and the various auxiliary basis sets employed. In our calculations, this includes the auxiliary basis for MP2 fitting (cbas), the auxiliary basis for constructing the Fock operator (jkbas) and the complementary auxiliary basis (cabs) for the many-electron RI. Furthermore, the errors introduced by neglecting contributions to the three- and four-electron integrals must also be controlled. In our MP2-F12 implementation we neglect some terms corresponding to Fock matrix elements between occupied and CABS orbitals (i.e. we assume that the generalised Brillouin condition (GBC) holds).

For conventional calculations, the behaviour of the error with respect to the orbital and auxiliary basis sets is well known. The error with respect to orbital basis converges as X^{-3} , where X is the cardinal number of the basis. The error in the energy with respect to the density fitting for the Hartree–Fock and MP2 integrals reduces quadratically with the size of the auxiliary basis sets. The situation for MP2-F12 is rather different. The convergence with respect to the orbital basis is predicted to reduce as X^{-7} , but the effect on the energy due to the density fitting and CABS-RI approximations is less well defined than in the conventional case because of the complexity of the equations. Nonetheless, we expect that comparing energies computed with two auxiliary basis sets representing cardinal numbers X and $X-1$ will give a conservative estimation of the remaining error associated with using the auxiliary basis X . This is the approach we adopt to assess the errors due to the auxiliary basis sets. To avoid problems with numerical stability, the fixed-amplitude method is used for these comparisons. Our investigations show that this restriction effects the reported errors only in the third significant figure.

To facilitate a fair and transparent comparison of the errors arising due to the various finite basis sets and approximations, we report errors per valence electron $\Delta = (E_{\text{app.}} - E_{\text{ref.}})/N_{\text{val}}$ for the molecules or reactions under consideration, in kJ mol^{-1} . This measure is also used for the error in the HF energy, which is reasonable because the error due to the core electrons is negligible in our orbital basis. In the discussion, we also refer to percentage errors, defined as $100(E_{\text{app.}} - E_{\text{ref.}})/E_{\text{ref.}}$. We report a statistical analysis of the errors in terms of the mean error $\bar{\Delta}$, the standard deviation σ , the mean absolute error $\Delta_{\text{abs.}}$, the root mean square error Δ_{rms} , and maximum error Δ_{max} :

$$\bar{\Delta} = \frac{1}{N} \sum_{i=1}^N \Delta_i, \quad (1)$$

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N (\Delta_i - \bar{\Delta})^2, \quad (2)$$

$$\Delta_{\text{abs.}} = \frac{1}{N} \sum_{i=1}^N |\Delta_i|, \quad (3)$$

$$\Delta_{\text{rms}}^2 = \frac{1}{N} \sum_{i=1}^N \Delta_i^2. \quad (4)$$

The dominant contribution to the deviation from the true basis set limit is the incompleteness of the orbital basis. Past experience, as well as our current calculations, show that the MP2-F12 energies

computed with approximation A (cf. Ref. [67] for the definition of approximations A and B) converge from above, whereas the energies computed using approximation B exhibit the usual convergence from below. To estimate the remaining basis set error, we assume that the MP2-F12/2A correlation energy, computed using the decontracted aug-cc-pV6Z basis, represents an upper bound. A conservative error estimate is thus obtained from a statistical analysis of the deviation of the MP2-F12/2B energies from the MP2-F12/2A energies. The resulting statistical measures of the error per valence electron due to the incompleteness of the orbital basis are reported in Table 2. The mean error and standard deviation of the orbital-basis error distribution are $0.042 \text{ kJ mol}^{-1}$ (0.041%) and $0.010 \text{ kJ mol}^{-1}$ (0.010%) respectively, and we therefore conclude that the fc-MP2 correlation energies are converged to within 0.05 kJ mol^{-1} (0.05%) of the basis set limit, per valence electron. As an aside, we mention that further decontraction of the aug-cc-pV6Z basis only lowered the correlation energy by $0.003 \text{ kJ mol}^{-1}$ per heavy atom, well below the basis set error due to higher angular momentum functions. The effect on the HF energy was even smaller.

The basis set limit correlation energies of three of the molecules in our set were computed using the fixed-amplitude method. For the remaining 103 molecules in our set, we were able to compare the MP2-F12 energies computed using fixed or variationally optimised amplitudes. The fixed-amplitude method systematically returns $\sim 0.004 \text{ kJ mol}^{-1}$ less correlation energy per valence electron than when the amplitudes are optimised. This error is of little consequence compared to the remaining 0.05% orbital basis-set-incompleteness error and we consider that we are justified in using the energies from the fixed-amplitude method as the reference basis set limit correlation energies for the three molecules difluoroacetylene (No. 36), carbon suboxide (No. 67) and dicyanoacetylene (No. 72).

In our MP2-F12 implementation, the exchange matrix elements are used to evaluate the integrals over $[K_1 + K_2, f_{12}]$ and the Fock matrix elements are used to evaluate the integrals over $[Q_{12}, F_1 + F_2]$ and also the coupling terms between the conventional and F12 amplitudes. F and K are the Fock and exchange operators respectively, f_{12} is the correlation factor and Q_{12} is the strong-orthogonality operator. All of these are comparatively small contributions and the error in the correlation energy arising from the RI-JK density fitting is thus expected to be tiny. To quantify this error, we have performed calculations using the aug-cc-pV5Z and aug-cc-pVQZ RI-JK auxiliary basis sets, keeping all other basis sets as those detailed in Section 2. An estimate of the density fitting error is obtained for each molecule as the difference between the MP2-F12/2B-fixed energies computed using these two RI-JK auxiliary basis sets. From the raw RI-JK of Table 2 we see that the aug-cc-pVQZ auxiliary basis gives correlation energies systematically smaller in magnitude than the aug-cc-pV5Z basis, by $0.0001 \text{ kJ mol}^{-1}$ per valence electron. The standard deviation of the error distribution is $0.00001 \text{ kJ mol}^{-1}$ and it is clear that the error in the RI-JK basis has essentially zero contribution to the deviation from the true basis set limit.

To estimate the error in the correlation energy arising from the density fitting of the MP2-type four-centre two-electron integrals, we have performed calculations using two density-fitting basis sets. We have used the specially constructed auxiliary set described in Section 2 and the aug-cc-pV6Z MP2-fitting auxiliary basis set. In Table 2, we present the statistical errors for the differences in the conventional fc-MP2 energies (denoted RI-MP2), in the F12 contribution to the fc-MP2-F12-fixed energies (denoted RI-F12) and the total fc-MP2-F12-fixed correlation energy (denoted MP2-Tot.). It is remarkable that the density fitting errors in the conventional

Table 2. Statistical measures of the errors per valence electron in the fc-MP2-F12 correlation energies (kJ mol^{-1}) arising from incompleteness in the orbital, CABS and density-fitting basis sets.

Error	$\bar{\Delta}$	σ	$\Delta_{\text{abs.}}$	Δ_{rms}	Δ_{max}	Molecule
Orb.	0.0421	0.0096	0.0421	0.0432	0.0586	79 F_2
CABS	-0.0024	0.0005	0.0024	0.0024	-0.0036	101 N_2
RI-MP2	0.0011	0.0002	0.0011	0.0011	0.0015	15 CH_2
RI-F12	-0.0012	0.0001	0.0012	0.0012	-0.0015	15 CH_2
RI-Tot.	-0.0001	0.0001	0.0001	0.0001	-0.0003	84 F_3N
RI-JK	0.0001	0.0000	0.0001	0.0001	0.0002	4 CF_2O
Fixed-Inv.	0.0042	0.0019	0.0042	0.0046	0.0086	5 CF_4

and F12 energy contributions consistently almost entirely cancel each other. The mean density-fitting error is $-0.0001 \text{ kJ mol}^{-1}$ per valence electron and the standard deviation is $0.0001 \text{ kJ mol}^{-1}$. Just as for the RI-JK error, the density fitting error is completely negligible in our calculations and even acts on average to cancel the RI-JK error.

The basis for the RI intrinsic to F12 theory is the union of the orbital and complementary auxiliary basis sets. This basis is used in many places to reduce three- and four-electron integrals, involving several combinations of operators, to two-electron integrals. It is also used for the matrix representation of the core Hamiltonian operator. We have performed calculations using the aug-cc-pV5Z and aug-cc-pVQZ MP2-fitting basis sets as CABS to estimate the error due to the incomplete basis for this RI error. The mean error is $-0.0024 \text{ kJ mol}^{-1}$ per valence electron, which represents a systematic shift to correlation energies that are too large in magnitude. The standard deviation is $0.0005 \text{ kJ mol}^{-1}$. Although this error is much larger than that of density fitting, it is still an order of magnitude smaller than the error due to the incomplete orbital basis.

Two further sources of error remain to be discussed. These are both related to the lack of convergence of the Hartree–Fock reference with respect to the orbital basis. The first is the GBC assumption, used in the evaluation of the B matrix in MP2-F12 theory. The second is the direct impact of using unconverged Hartree–Fock orbitals and orbital energies in the MP2 expressions. The importance of the Fock matrix elements between occupied and CABS orbitals may be assessed by computing the contribution to the second-order energy when these terms are treated as a perturbation and result in single excitations into orbitals in the CABS space [68]. The mean energy of this CABS singles correction is $-0.0026 \text{ kJ mol}^{-1}$ per valence electron, with a standard deviation of $0.0006 \text{ kJ mol}^{-1}$. Although this CABS-singles perturbation energy is not a direct measure of either of the two aforementioned errors, it is certainly an indication that these errors may safely be neglected in the decontracted aug-cc-pV6Z basis, particularly because the effect of unconverged orbitals on the correlation energy has previously been observed to be two orders of magnitude smaller than the effect on the Hartree–Fock energy.

An estimate of the overall deviation of our reference fc-MP2 correlation energies from the basis-set limit is obtained by Gaussian error propagation of the Orb., CABS, RI-Tot. and RI-JK errors. The resulting mean error is $0.040 \text{ kJ mol}^{-1}$ per valence electron and the standard deviation is 0.01 kJ mol^{-1} .

This error is completely dominated by the incompleteness of the orbital basis and our conservative estimates give a 95% confidence interval with a maximum error of 0.05 kJ mol^{-1} per valence electron (0.06%). For the Hartree–Fock basis set limit reference value we use the values computed with the decontracted aug-cc-pV6Z basis. No RI approximations were employed for the Hartree–Fock energy. Since the change in energy upon further decontraction or upon including the CABS singles correction was below $0.005 \text{ kJ mol}^{-1}$ per valence electron on average, the error in the total fc-MP2 electronic energies may also be taken to be 0.05 kJ mol^{-1} per valence electron.

4. Basis set assessment

4.1. Absolute Hartree–Fock and correlation energies

Having determined the fc-MP2 basis set limit to 99.95% accuracy, we may now proceed to assess the various orbital basis sets for use with the MP2-F12 method. The predominant basis sets in current usage are the aug-cc-pVXZ sets of Dunning and co-workers, the Karlsruhe def2-TZVPP and def2-QZVPP sets and the newly developed cc-pVXZ-F12 sets of Peterson and co-workers. We have computed the fc-MP2-F12/2B energies using each of these basis sets, always in combination with all of the large auxiliary basis sets described in Section 2. We used an exponent of $1.4 a_0^{-1}$ for the Slater-type correlation factor for all calculations except for the F12 basis sets, where we used the recommended values of 0.9, 1.0 and 1.1 a_0^{-1} for X = D, T and Q, respectively. Both the Hartree–Fock and correlation contributions are of interest when assessing the basis sets.

The statistical measures for the deviation of the computed Hartree–Fock energies from the reference values are presented in Table 3. In Figure 1, we plot the mean basis-set errors (in kJ mol^{-1}) against the number of functions in the basis set for a first-row main-group

Table 3. Statistical measures for the error per valence electron in the Hartree–Fock energy (kJ mol^{-1}).

Basis	$\bar{\Delta}$	σ	Δ_{abs}	Δ_{rms}	Δ_{max}	Molecule
cc-pVDZ-F12	4.05	0.53	4.05	4.08	7.42	106 H ₂
cc-pVTZ-F12	0.57	0.15	0.57	0.59	0.91	73 FH
cc-pVQZ-F12	0.06	0.01	0.06	0.06	0.09	73 FH
aug-cc-pVDZ	8.74	2.13	8.74	9.00	13.98	79 F ₂
aug-cc-pVTZ	2.14	0.58	2.14	2.22	3.49	79 F ₂
aug-cc-pVQZ	0.45	0.14	0.45	0.47	0.82	79 F ₂
aug-cc-pV5Z	0.05	0.01	0.05	0.05	0.07	79 F ₂
def2-TZVPP	1.35	0.25	1.35	1.37	1.78	84 F ₃ N
def2-QZVPP	0.17	0.04	0.17	0.17	0.27	98 H ₃ NO

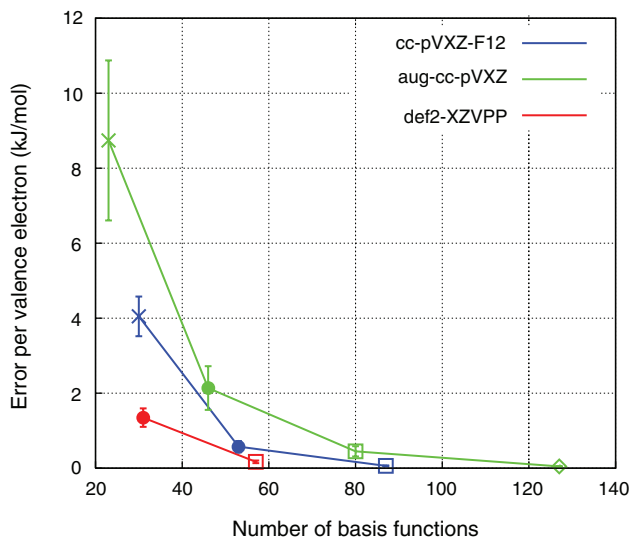


Figure 1. Mean errors per valence electron (kJ mol^{-1}) in the HF energy as a function of the size of the basis ($\times = \text{DZ}$, $\bullet = \text{TZ}$, $\square = \text{QZ}$, $\diamond = 5\text{Z}$).

element (C, N, O or F). Here and in all subsequent plots of this type, the standard deviation is indicated as an error bar. We see that the def2-TZVPP and def2-QZVPP basis sets are very well optimised with respect to the Hartree–Fock limit, whereas the aug-cc-pVDZ and cc-pVDZ-F12 basis sets have comparably large errors and error bars. The performance of the cc-pVTZ-F12 and cc-pVQZ-F12 basis sets is very convincing, however. Their errors are always found below the aug-cc-pVXZ curve, not only as a function of the cardinal number but also as a function of the number of basis functions. In MP2-F12 calculations the error in the HF energy can be significantly reduced by including a perturbative correction from single excitations into the CABS basis. In Table 4, we present the error statistics for the corrected HF energies. This correction greatly improves the Hartree–Fock energies, reducing the error by a factor of roughly 4, 3 and 2 for basis sets with cardinal numbers $X=2, 3$ and 4, respectively. We note that although we have used a very large CABS basis in our calculations, the convergence of the singles correction with the CABS basis is rapid and similar values are expected when smaller CABS basis sets are used.

The corresponding error distributions for the deviation of the correlation energies from the reference basis-set limit values are reported in Table 5 and depicted in Figure 2. The analogous values computed using F12 amplitudes determined using the cusp conditions rather than variational optimisation are presented in the Supporting Information. There is a clear improvement in the correlation energies when

Table 4. Statistical measures for the error per valence electron in the Hartree–Fock energy when the CABS singles correction is included (kJ mol^{-1}).

Basis	$\bar{\Delta}$	σ	$\Delta_{\text{abs.}}$	Δ_{rms}	Δ_{max}	Molecule
cc-pVDZ-F12	0.93	0.09	0.93	0.93	1.47	106 H_2
cc-pVTZ-F12	0.22	0.04	0.22	0.22	0.32	79 F_2
cc-pVQZ-F12	0.03	0.01	0.03	0.04	0.06	79 F_2
aug-cc-pVDZ	1.21	0.17	1.21	1.23	1.71	79 F_2
aug-cc-pVTZ	0.68	0.09	0.68	0.68	0.89	79 F_2
aug-cc-pVQZ	0.16	0.04	0.16	0.17	0.27	79 F_2
aug-cc-pV5Z	0.03	0.01	0.03	0.03	0.05	79 F_2
def2-TZVPP	0.43	0.06	0.43	0.43	0.64	98 H_3NO
def2-QZVPP	0.07	0.02	0.07	0.07	0.12	98 H_3NO

Table 5. Statistical measures for the error per valence electron (kJ mol^{-1}) in the fc-MP2-F12 correlation energy (invariant formulation) as well as in the extrapolated fc-MP2 correlation energy.

Basis	$\bar{\Delta}$	σ	$\Delta_{\text{abs.}}$	Δ_{rms}	Δ_{max}	Molecule
cc-pVDZ-F12	0.97	0.24	0.97	1.00	1.75	73 FH
cc-pVTZ-F12	0.26	0.08	0.26	0.27	0.41	79 F_2
cc-pVQZ-F12	0.07	0.02	0.07	0.07	0.10	79 F_2
aug-cc-pVDZ	1.71	0.22	1.71	1.73	2.19	79 F_2
aug-cc-pVTZ	0.63	0.18	0.63	0.65	1.02	79 F_2
aug-cc-pVQZ	0.23	0.07	0.23	0.24	0.40	79 F_2
aug-cc-pV5Z	0.09	0.03	0.09	0.09	0.17	79 F_2
def2-TZVPP	1.15	0.34	1.15	1.19	2.00	73 FH
def2-QZVPP	0.44	0.15	0.44	0.47	0.79	79 F_2
aug-(TQ)	0.64	0.27	0.64	0.69	1.30	79 F_2
aug-(Q5)	0.15	0.05	0.15	0.16	0.25	5 CF_4

progressing from the Karlsruhe sets, to the Dunning sets, to the F12 sets of Peterson. This trend is also shown clearly in Figure 3, where the mean basis set errors (kJ mol^{-1}) are plotted as a function of the number of basis functions of a first-row main-group element, as before. We see that with respect to the size of the basis, the Karlsruhe and Dunning sets lie on essentially the same curve, but the specially optimised F12 sets are a definite improvement. Although the cc-pVDZ-F12 and def2-TZVPP sets are approximately the same size, the F12 set has no f functions, replacing them with diffuse s and p functions, which are more important for F12 calculations.

Figure 2 also contains the error distributions for the deviation of the correlation energies that are obtained by means of the X^{-3} extrapolation scheme of Helgaker and co-workers [69]. The mean errors in the MP2 correlation energy are 0.64 and 0.15 kJ mol^{-1} per valence electron after extrapolating from the aug-cc-pVTZ/aug-cc-pVQZ—denoted as aug-(TQ)—or from the aug-cc-pVQZ/aug-cc-pV5Z results,

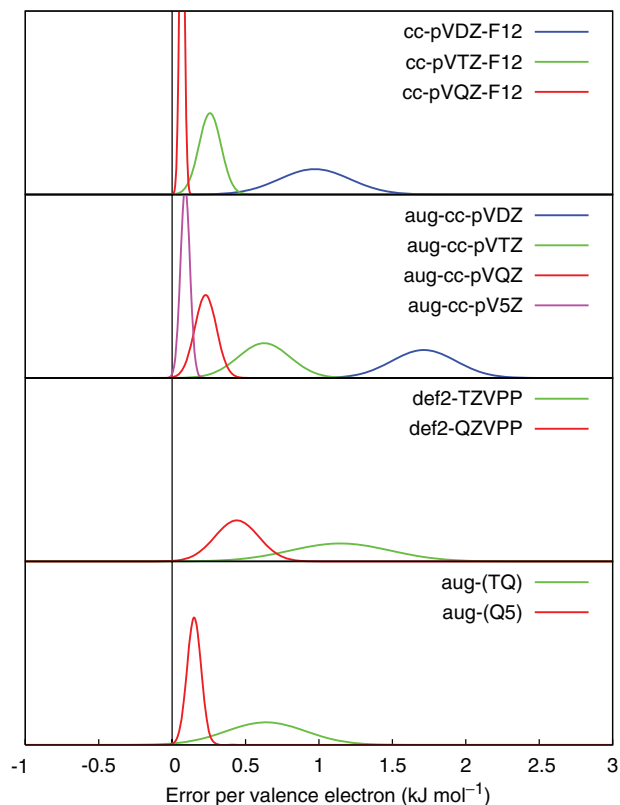


Figure 2. Performance of basis sets and extrapolation methods for absolute correlation energies.

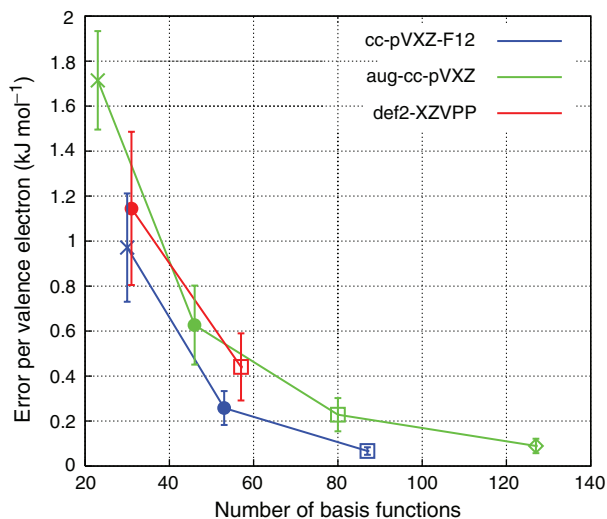


Figure 3. Mean errors per valence electron (kJ mol^{-1}) in the absolute correlation energy as a function of the size of the basis ($\times = \text{DZ}$, $\bullet = \text{TZ}$, $\square = \text{QZ}$, $\diamond = \text{5Z}$).

respectively (Table 5). The aug-(Q5) extrapolation is capable of recovering ca. 99.86% of the correlation energy on average with a standard deviation of only 0.04% (0.05 kJ mol^{-1} per valence electron). This is

Table 6. statistical measures for the error per valence electron in the total fc-MP2 electronic energy (kJ mol^{-1}).

Basis	$\bar{\Delta}$	σ	$\Delta_{\text{abs.}}$	Δ_{rms}	Δ_{max}	Molecule
cc-pVDZ-F12	1.90	0.30	1.90	1.92	3.00	73 FH
cc-pVTZ-F12	0.48	0.12	0.48	0.49	0.73	79 F ₂
cc-pVQZ-F12	0.10	0.03	0.10	0.10	0.16	79 F ₂
aug-cc-pVDZ	2.93	0.36	2.93	2.95	3.90	79 F ₂
aug-cc-pVTZ	1.30	0.26	1.30	1.33	1.91	79 F ₂
aug-cc-pVQZ	0.39	0.11	0.39	0.41	0.66	79 F ₂
aug-cc-pV5Z	0.12	0.04	0.12	0.13	0.22	79 F ₂
def2-TZVPP	1.57	0.38	1.57	1.62	2.58	73 FH
def2-QZVPP	0.51	0.17	0.51	0.53	0.92	79 F ₂

slightly better than the accuracy of F12/cc-pVTZ-F12 values (0.26 kJ mol^{-1} , 99.75%) but not as good as F12/cc-pVQZ-F12 (0.07 kJ mol^{-1} , 99.94%). We note, furthermore, that the aug-(TQ) and F12/cc-pVDZ-F12 levels are of comparable accuracy.

The error distributions for the deviation of the total MP2 energy (Hartree–Fock plus correlation plus CABS singles) are reported in Table 6. We find that the most accurate results are obtained at the cc-pVQZ-F12 and aug-cc-pV5Z levels, with mean errors per valence electron of 0.10 and 0.12 kJ mol^{-1} , respectively. The mean errors are slightly larger in the aug-cc-pVQZ, cc-pVTZ-F12 and def2-QZVPP basis sets (0.39, 0.48 and 0.51 kJ mol^{-1}). Although these three basis sets perform similarly well, the aug-cc-pVQZ basis is significantly larger than the def2-QZVPP and cc-pVTZ-F12 basis sets, which are almost equivalent in size and accuracy. We note, however, that the error in the def2-QZVPP basis is dominated by the correlation contribution, whereas the error in the cc-pVTZ-F12 basis is evenly distributed between Hartree–Fock and correlation.

4.2. Contributions to atomisation energies

It is important to assess the performance of the basis sets not only for total energies but also for relative energies. Using the fc-MP2-F12 method with the Peterson, Dunning and Karlsruhe basis sets, we have computed atomisation energies for the 106 molecules. Table 7 contains the statistical measures for the error per valence electron (kJ mol^{-1}) in the correlation contribution to the atomisation energies and we have also included the values obtained from extrapolating conventional fc-MP2 energies. The error distributions are displayed graphically in Figure 4, and in Figure 5 the mean errors are plotted against the size of the basis. The corresponding table and figure for the

Table 7. Statistical measures for the error per valence electron (kJ mol^{-1}) in the correlation contributions to the atomisation energies (invariant formulation).

Basis	$\bar{\Delta}$	σ	Δ_{abs}	Δ_{rms}	Δ_{max}	Molecule
cc-pVDZ-F12	0.21	0.31	0.31	0.37	0.84	105 O ₃
cc-pVTZ-F12	0.08	0.04	0.08	0.09	0.17	105 O ₃
cc-pVQZ-F12	0.02	0.01	0.02	0.02	0.04	105 O ₃
aug-cc-pVDZ	0.31	0.35	0.37	0.47	1.04	105 O ₃
aug-cc-pVTZ	0.15	0.07	0.15	0.16	0.28	83 F ₂ O ₂
aug-cc-pVQZ	0.12	0.02	0.12	0.12	0.17	67 C ₃ O ₂
aug-cc-pV5Z	0.06	0.01	0.06	0.06	0.08	67 C ₃ O ₂
def2-TZVPP	0.12	0.17	0.17	0.21	0.49	5 CF ₄
def2-QZVPP	0.11	0.05	0.11	0.12	0.23	5 CF ₄
aug-(TQ)	-0.09	0.08	0.10	0.12	-0.38	101 N ₂
aug-(Q5)	-0.09	0.03	0.09	0.10	-0.17	90 HN ₃

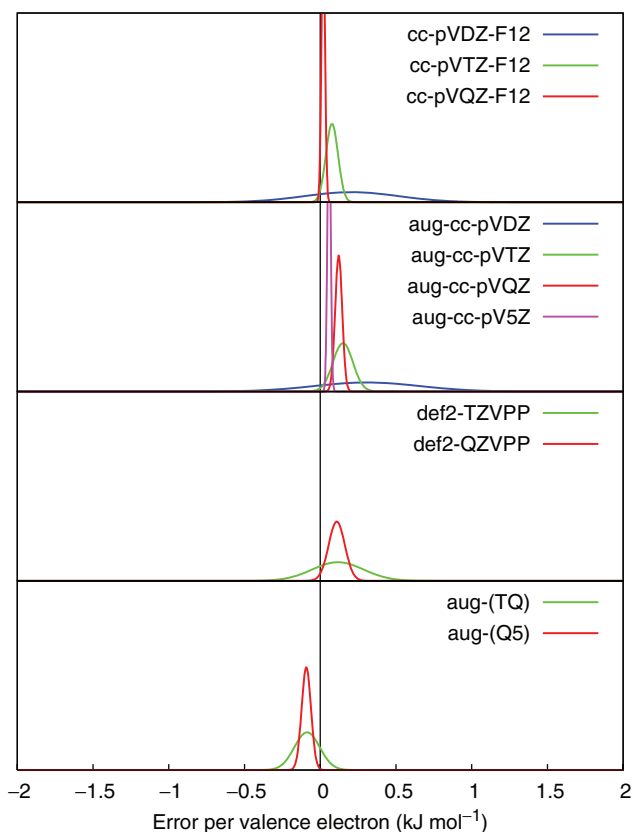


Figure 4. Performance of basis sets and extrapolation methods for atomisation energies.

Hartree–Fock (plus CABS singles) contribution are given in the Supporting Information.

Figure 4 shows that the results converge smoothly to the basis-set limits with the aug-cc-pVXZ and cc-pVXZ-F12 series, in terms of both mean errors and standard deviations. The error distributions for the smallest sets of the series (cc-pVDZ-F12,

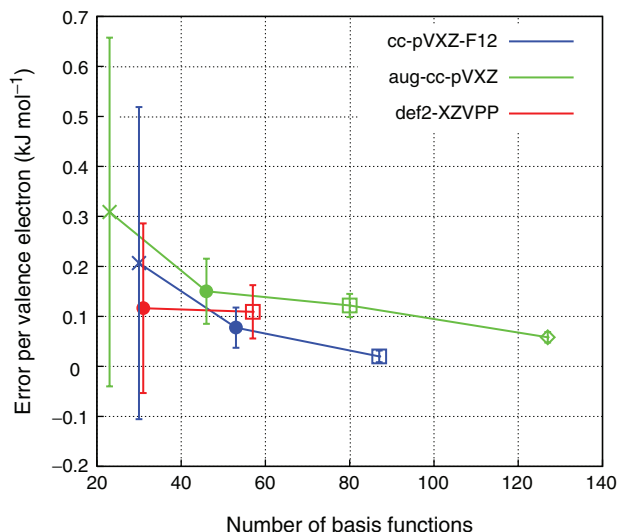


Figure 5. Mean errors per valence electron (kJ mol^{-1}) in the correlation contribution to the atomisation energy as a function of the size of the basis ($\times = \text{DZ}$, $\bullet = \text{TZ}$, $\square = \text{QZ}$, $\diamond = \text{5Z}$).

aug-cc-pVDZ and def2-TZVPP) are relatively broad, with standard deviations of 0.31, 0.34 and 0.17 kJ mol^{-1} per valence electron, respectively. Among these sets, the def2-TZVPP basis gives the smallest mean-absolute errors (0.17 kJ mol^{-1}). The results obtained from the cc-pVTZ-F12, aug-cc-pVQZ and def2-QZVPP basis sets are much better than from the smaller sets, with RMS deviations of 0.08, 0.12 and 0.12 kJ mol^{-1} , respectively. But note that the cc-pVTZ-F12 and def2-QZVPP basis sets are similar in size, whereas the aug-cc-pVQZ is significantly larger (Figure 5). The F12 basis sets of Peterson are a significant improvement over the Dunning sets, but the def2-TZVPP set is preferred over the cc-pVDZ-F12 set. We also note that the cc-pVTZ-F12 basis gives more accurate results than the (TQ) extrapolation, and similar accuracy to the (Q5) extrapolation.

In Table 8 we present the statistical errors for the sum of the fc-MP2 correlation energy plus Hartree–Fock contribution to the atomisation energy, including the correction obtained from the CABS single excitations (see also Figure 6). Such a sum corresponds to a realistic calculation of a relative energy such as an activation energy or heat of formation. We find that all F12 calculations studied in the present work yield atomisation energies accurate to within $\pm 0.5 \text{ kJ mol}^{-1}$ standard deviation, per valence electron (equivalent to an RMS error of 0.2% in the atomisation energies). In the cc-pVTZ-F12 and cc-pVQZ-F12 basis sets, the accuracy is better than $\pm 0.05 \text{ kJ mol}^{-1}$. When comparing the performance of the various basis sets,

Table 8. Statistical measures for the error per valence electron (kJ mol^{-1}) in the atomisation energies (Hartree–Fock plus frozen-core correlation contribution plus CABS singles).

Basis	$\bar{\Delta}$	σ	Δ_{abs}	Δ_{rms}	Δ_{max}	Molecule
cc-pVDZ-F12	-0.33	0.34	0.35	0.47	-1.84	106 H_2
cc-pVTZ-F12	-0.02	0.04	0.03	0.04	-0.18	106 H_2
cc-pVQZ-F12	0.01	0.01	0.01	0.01	-0.03	106 H_2
aug-cc-pVDZ	-0.21	0.33	0.29	0.39	-2.10	106 H_2
aug-cc-pVTZ	0.02	0.10	0.08	0.11	-0.64	106 H_2
aug-cc-pVQZ	0.09	0.04	0.10	0.10	-0.16	106 H_2
aug-cc-pV5Z	0.05	0.01	0.05	0.06	0.08	67 C_3O_2
def2-TZVPP	-0.04	0.23	0.17	0.23	-0.93	106 H_2
def2-QZVPP	0.08	0.06	0.09	0.10	0.21	5 CF_4

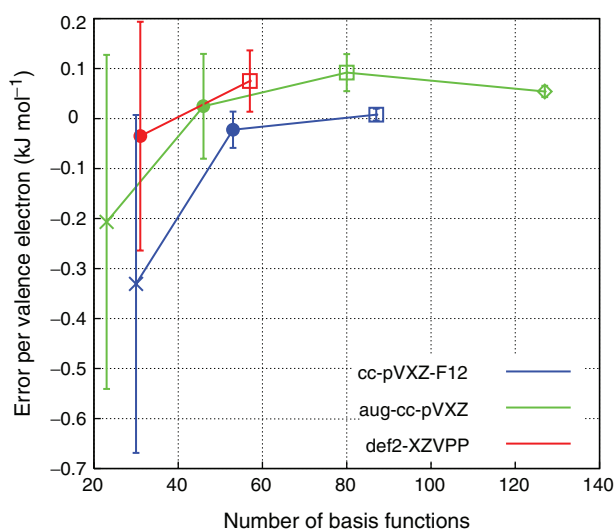


Figure 6. Mean errors per valence electron (kJ mol^{-1}) in the total atomisation energy (Hartree–Fock plus correlation contribution plus CABS singles) as a function of the size of the basis ($\times = \text{DZ}$, $\bullet = \text{TZ}$, $\square = \text{QZ}$, $\diamond = \text{5Z}$).

we note that the Hartree–Fock contribution converges from below and the correlation contribution from above. For the double-zeta sets, the error is dominated by that of the Hartree–Fock contribution, even though the CABS singles correction is included, but for the def2-QZVPP, aug-cc-pVQZ and aug-cc-pV5Z basis sets, the correlation error dominates. For the remaining sets, def2-TZVPP, cc-pVTZ-F12 and cc-pVQZ-F12, the mean Hartree–Fock and correlation errors are well balanced and also act to cancel each other. The cc-pVXZ-F12 results converge very smoothly to the reference values, and the standard deviations in these basis sets are significantly smaller than in the other sets. The only exception is the double-zeta set, where the def2-TZVPP set is preferred.

5. Conclusions

One-electron basis sets for F12 explicitly-correlated molecular electronic-structure methods have been assessed by analysing the accuracy of Hartree–Fock energies and valence-only second-order correlation energies (fc-MP2) of a test set of 106 small molecules containing the atoms H, C, N, O and F. Absolute energies as well as relative energies (atomisation energies) have been investigated.

The cc-pVXZ-F12 ($X = \text{D, T, Q}$) basis sets developed by Peterson and co-workers [47] especially for F12 calculations perform remarkably well, except perhaps for the smallest member in the series, cc-pVDZ-F12, which yields RMS errors of 0.47 kJ mol^{-1} per valence electron ($99.93 \pm 0.11\%$) of the atomisation energy. Using the cc-pVTZ-F12 and cc-pVQZ-F12 basis sets, the atomisation energies are obtained accurate to within 0.04 kJ mol^{-1} ($99.99 \pm 0.03\%$) and 0.01 kJ mol^{-1} ($99.99 \pm 0.02\%$), respectively, with respect to the basis-set-limit reference value (Table 8).

The X^{-3} extrapolation from the aug-cc-pVQZ and aug-cc-pV5Z is slightly more accurate than F12 theory in the cc-pVTZ-F12 basis but less accurate than F12 theory in the cc-pVQZ-F12 basis (Figure 4).

The cc-pVQZ-F12 and aug-cc-pV5Z basis sets are the preferred basis sets for highly accurate calculations using F12 methodology. If such high accuracy is not needed, then the basis sets aug-cc-pVQZ, cc-pVTZ-F12 and def2-QZVPP may be used almost interchangeably.

Acknowledgements

This research was supported by the Deutsche Forschungsgemeinschaft (DFG) in the framework of the Schwerpunktprogramm ‘First Principles Methods’ (SPP 1145, grant No. KL 721/2-3). F.A.B. gratefully acknowledges support by the Fonds der Chemischen Industrie through a Chemiefonds stipend. D.P.T. gratefully acknowledges support by the DFG through grant No. TE 644/1-1.

References

- [1] E.F. Valeev, W.D. Allen, R. Hernandez, C.D. Sherrill, and H.F. Schaefer III, *J. Chem. Phys.* **118**, 8594 (2003).
- [2] M.S. Schuurman, S.R. Muir, W.D. Allen, and H.F. Schaefer III, *J. Chem. Phys.* **120**, 11586 (2004).
- [3] G. Tarczay, A.G. Császár, W. Klopper, V. Szalay, W.D. Allen, and H.F. Schaefer III, *J. Chem. Phys.* **110**, 11971 (1999).
- [4] E.F. Valeev, W.D. Allen, H.F. Schaefer III, and A.G. Császár, *J. Chem. Phys.* **114**, 2875 (2001).

- [5] J.P. Kenny, W.D. Allen, and H.F. Schaefer III, *J. Chem. Phys.* **118**, 7353 (2003).
- [6] M.S. Schuurman, W.D. Allen, P. von Ragué Schleyer, and H.F. Schaefer III, *J. Chem. Phys.* **122**, 104302 (2005).
- [7] T. Rajamäki, M. Kállay, J. Noga, P. Valiron, and L. Halonen, *Mol. Phys.* **102**, 21 (2004).
- [8] T. Rajamäki, J. Noga, P. Valiron, and L. Halonen, *Mol. Phys.* **102**, 2259 (2004).
- [9] W. Klopper, J.G.C.M. van Duijneveldt-van de Rijdt, and F.B. van Duijneveldt, *Phys. Chem. Chem. Phys.* **2**, 2227 (2000).
- [10] G.S. Tschumper, M.L. Leininger, B.C. Hoffman, E.F. Valeev, H.F. Schaefer III, and M. Quack, *J. Chem. Phys.* **116**, 690 (2002).
- [11] A. Faure, P. Valiron, M. Wernli, L. Wiesenfeld, C. Rist, J. Noga, and J. Tennyson, *J. Chem. Phys.* **122**, 221102 (2005).
- [12] P. Valiron, M. Wernli, A. Faure, L. Wiesenfeld, C. Rist, S. Kedžuch, and J. Noga, *J. Chem. Phys.* **129**, 134306 (2008).
- [13] H.M. Sulzbach, H.F. Schaefer III, W. Klopper, and H.P. Lüthi, *J. Am. Chem. Soc.* **118**, 3519 (1996).
- [14] W. Klopper and H.P. Lüthi, *Chem. Phys. Lett.* **262**, 546 (1996).
- [15] M.O. Sinnokrot, E.F. Valeev, and C.D. Sherrill, *J. Am. Chem. Soc.* **124**, 10887 (2002).
- [16] M.O. Sinnokrot and C.D. Sherrill, *J. Phys. Chem. A* **110**, 10656 (2006).
- [17] W. Klopper, F.R. Manby, S. Ten-No, and E.F. Valeev, *Int. Rev. Phys. Chem.* **25**, 427 (2006).
- [18] W. Klopper and C.C.M. Samson, *J. Chem. Phys.* **116**, 6397 (2002).
- [19] H.-J. Werner, T.B. Adler, and F.R. Manby, *J. Chem. Phys.* **126**, 164102 (2007).
- [20] S. Ten-no, *J. Chem. Phys.* **126**, 014108 (2007).
- [21] D. Yamaki, H. Koch, and S. Ten-No, *J. Chem. Phys.* **127**, 144104 (2007).
- [22] D.P. Tew, W. Klopper, C. Neiss, and C. Hättig, *Phys. Chem. Chem. Phys.* **9**, 1921 (2007).
- [23] D.P. Tew, W. Klopper, C. Neiss, and C. Hättig, *Phys. Chem. Chem. Phys.* **10**, 6325, erratum (2008).
- [24] T.B. Adler, G. Knizia, and H.-J. Werner, *J. Chem. Phys.* **127**, 221106 (2007).
- [25] G. Knizia, T.B. Adler, and H.-J. Werner, *J. Chem. Phys.* (in press).
- [26] D.P. Tew, W. Klopper, C. Neiss, and C. Hättig, *Chem. Phys. Lett.* **452**, 326 (2008).
- [27] E.F. Valeev, *Phys. Chem. Chem. Phys.* **10**, 106 (2008).
- [28] T. Shiozaki, M. Kamiya, S. Hirata, and E.F. Valeev, *Phys. Chem. Chem. Phys.* **10**, 3358 (2008).
- [29] M. Torheyden and E.F. Valeev, *Phys. Chem. Chem. Phys.* **10**, 3410 (2008).
- [30] E.F. Valeev and T.D. Crawford, *J. Chem. Phys.* **128**, 244113 (2008).
- [31] T. Shiozaki, M. Kamiya, S. Hirata, and E.F. Valeev, *J. Chem. Phys.* **129**, 071101 (2008).
- [32] J. Noga, S. Kedžuch, J. Šimunek, and S. Ten-no, *J. Chem. Phys.* **128**, 174103 (2008).
- [33] D. Bokhan, S. Ten-no, and J. Noga, *Phys. Chem. Chem. Phys.* **10**, 3320 (2008).
- [34] A. Köhn, G.W. Richings, and D.P. Tew, *J. Chem. Phys.* **129**, 201103 (2008).
- [35] S. Ten-no, *Chem. Phys. Lett.* **447**, 175 (2007).
- [36] R. Leist, J.A. Frey, P. Ottiger, H.-M. Frey, S. Leutwyler, R.A. Bachorz, and W. Klopper, *Angew. Chem. Int. Ed.* **46**, 7449 (2007).
- [37] R.A. Bachorz, F.A. Bischoff, S. Höfener, W. Klopper, P. Ottiger, R. Leist, J.A. Frey, and S. Leutwyler, *Phys. Chem. Chem. Phys.* **10**, 2758 (2008).
- [38] W. Klopper, D.P. Tew, N. González-García, and M. Olzmann, *J. Chem. Phys.* **129**, 114308 (2008).
- [39] N. González-García, W. Klopper, and M. Olzmann, *Chem. Phys. Lett.* (in press).
- [40] S. Höfener, F.A. Bischoff, A. Glöb, and W. Klopper, *Phys. Chem. Chem. Phys.* **10**, 3390 (2008).
- [41] K.D. Vogiatzis, A. Mavrandonakis, W. Klopper and G.E. Froudakis, *Chem. Phys. Chem* (in press).
- [42] A. Bihlmeier, D.P. Tew, and W. Klopper, *J. Chem. Phys.* **129**, 114303 (2008).
- [43] Y. Wang, B.J. Braams, J.M. Bowman, S. Carter, and D.P. Tew, *J. Chem. Phys.* **128**, 224314 (2008).
- [44] F. Claeysens, J.N. Harvey, F.R. Manby, R.A. Mata, A.J. Mulholland, K.E. Ranaghan, M. Schütz, S. Thiel, W. Thiel, and H.-J. Werner, *Angew. Chem. Int. Ed.* **45**, 6856 (2006).
- [45] R.A. Mata, H.-J. Werner, S. Thiel, and W. Thiel, *J. Chem. Phys.* **128**, 025104 (2008).
- [46] D.P. Tew and W. Klopper, *J. Chem. Phys.* **123**, 074101 (2005).
- [47] K.A. Peterson, T.B. Adler, and H.-J. Werner, *J. Chem. Phys.* **128**, 084102 (2008).
- [48] T.H. Dunning Jr, *J. Chem. Phys.* **90**, 1007 (1989).
- [49] R.A. Kendall, T.H. Dunning Jr, and R.J. Harrison, *J. Chem. Phys.* **96**, 6796 (1992).
- [50] A.K. Wilson, T. van Mourik, and T.H. Dunning Jr, *J. Mol. Struct.: THEOCHEM* **388**, 339 (1996).
- [51] F. Weigend and R. Ahlrichs, *Phys. Chem. Chem. Phys.* **7**, 3297 (2005).
- [52] D. Bakowies, *J. Chem. Phys.* **127**, 084105 (2007).
- [53] W. Klopper, B. Ruscic, D.P. Tew, F.A. Bischoff and S. Wolfsegger, *Chem. Phys.* (in press).
- [54] S. Ten-no, *Chem. Phys. Lett.* **398**, 56 (2004).
- [55] W. Klopper, *Chem. Phys. Lett.* **186**, 583 (1991).
- [56] E.F. Valeev, *Chem. Phys. Lett.* **395**, 190 (2004).
- [57] C. Hättig, *Phys. Chem. Chem. Phys.* **7**, 59 (2005).
- [58] C. Hättig and F. Weigend, *J. Chem. Phys.* **113**, 5154 (2000).
- [59] TURBOMOLE v5.10, Turbomole GmbH, Karlsruhe (2008), see <http://www.turbomole.com>

- [60] F.R. Manby, *J. Chem. Phys.* **119**, 4607 (2003).
- [61] F. Weigend and M. Häser, *Theor. Chem. Acc.* **97**, 331 (1997).
- [62] F. Weigend, M. Häser, H. Patzelt, and R. Ahlrichs, *Chem. Phys. Lett.* **294**, 143 (1998).
- [63] F.A. Bischoff, S. Höfener, A. Glöß, and W. Klopper, *Theor. Chem. Acc.* **121**, 11 (2008).
- [64] F. Weigend, *J. Comput. Chem.* **29**, 167 (2008).
- [65] F. Weigend, unpublished.
- [66] S. Ten-no, *J. Chem. Phys.* **121**, 117 (2004).
- [67] W. Kutzelnigg and W. Klopper, *J. Chem. Phys.* **94**, 1985 (1991).
- [68] J. Noga, S. Kedžuch, and J. Šimunek, *J. Chem. Phys.* **127**, 034106 (2007).
- [69] T. Helgaker, W. Klopper, H. Koch, and J. Noga, *J. Chem. Phys.* **106**, 9639 (1997).