Theor Chem Acc (2005) 113: 152–160
DOI 10.1007/s00214-005-0619-2

REGULAR ARTICLE

Shant Shahbazian · Mansour Zahedi

Towards a complete basis set limit of Hartree–Fock method: correlation-consistent versus polarized-consistent basis sets

Received: 22 June 2004 / Accepted: 15 October 2004 / Published online: 2 April 2005

Abstract In this paper the convergence pattern of correlation-consistent (cc-pVxZ) and polarized-consistent (PC-*n*) hierarchies relative to the complete basis set limit have been considered in a small set of diatomic molecules. Using the sequence of these basis sets it was demonstrated that potential energy surfaces derived from basis-set-dependent solution of the Hartree–Fock equations achieves the exact numerical derived potential energy surfaces (PESs) in an ordered manner. So it was possible to compute the spectroscopic parameters in the complete basis set limit with considerable accuracy using the most extended members of both hierarchies. On the other hand, for the first time the detailed convergence patterns of total energies in three separate inter-nuclear distances have been considered in these molecules and it was demonstrated that the total energies arrive at microhartree accuracy at a considerable rate. Possible performance of extrapolation schemes is discussed and it was demonstrated that reliable extrapolation procedures indeed exist. A successful test of the proposed extrapolation method, using the three most extended members of polarized-consistent basis sets, has been accomplished on selected polyatomic molecules.

Keywords Hartree–Fock method · Correlation-consistent · Polarized-consistent · Complete basis set limit · Extrapolation schemes

Introduction

In recent years, modern electronic structure theory has achieved great advances in precise, quantitative description of molecular systems. Thanks to this progress, it is now possible, using sophisticated methods like CCSD(T), to solve the Schrödinger equation for real-life molecules with reasonable accuracy [1]. In contrast to these developments, even now reaching exact solutions is a formidable task due to the incompleteness of basis sets that are used to solve the algebraic equations and which are employed in the vast majority of current computer codes due to their algorithmic simplicity. Knowing the fact that direct numerical solution of approximate equations like coupled cluster equations could not achieve the same popularity in near future, it is necessary to develop strategies to estimate the complete basis set limit of a particular property like total energy using a hierarchy of finite basis sets. After construction of correlation-consistent basis sets by Dunning [2], the first evidence was introduced for a practical scheme to reach this goal. Since then, a decade of practice to estimate the complete basis set limit of different correlated methods using this hierarchy by different research groups has shown correlation-consistent basis sets to be the best candidate for this propose, at least for post-Hartree–Fock methods [3–7]. In contrast to the reliability of this sequence of basis sets to estimate the correlation energy of molecular systems, less attention has been devoted to the Hartree–Fock energies. This does not seem surprising due to the known intrinsic difference between patterns for reaching the complete basis set limit by HF and post-HF methods [8]. It is evident that, to have good precision in estimating the complete basis set limit of a particular property like energy, it is necessary to reach both the HF as well as the correlated basis set limit with the same accuracy.

In pursuit of such a goal, in 1999 a systematic study of the HF limit of the total and relative energies for some diatomic molecules, comparing numerical solutions with those derived from correlation consistent basis sets, appeared [9]. Using different extrapolation procedures to estimate the complete basis set limit, the authors came to the conclusion that "there is no general reliable scheme to extrapolate complete basis set limit". So according to their final judgment, direct computation by utilizing the largest basis set available was proposed as the only safe method to estimate the complete basis sets limit with reasonable accuracy (at least sub-millihartree).

S. Shahbazian · M. Zahedi (✉)
Department of Chemistry, Faculty of Sciences,
Shahid Beheshti University,
Evin, Tehran, Iran, 19839,
P.O.Box 19395-4716
E-mail: m-zahedi@cc.sbu.ac.ir

In this paper we have reexamined the complete basis set limit of the HF method using correlation-consistent basis sets and also newly developed polarized-consistent basis sets [10–12], comparing with refined numerical solutions of the Hartree–Fock equations.

Computational details and selected species

Main computational procedure

To arrive at a reliable estimate of quality of basis sets used to attain the complete basis set limit, precise basis-set-free calculated values of the property under consideration are needed. So accurate numerical solution of the HF equations are invoked for such comparison.

Although first attempts at numerical solution of the HF equations goes back to the late 1970s and early 1980s [13, 14], real progress was achieved later in the 1980s [15–18], although the quality of solution was not clear at that time. As we will present in this paper, part of the disappointment arising in this previous investigation [9], comes from a certain ambiguity. To have reliable numerical values for properties considered in this investigation, all desired properties have been recalculated using newly developed code by Kobus [19–22]. The precision of this code has been tested elsewhere [19,23], thus giving us confidence in its reliability. The differences from previous literature are discussed in the subsequent section.

The correlation-consistent basis sets [2], and newly developed polarized-consistent basis sets [10–12] have been used for algebraic solution. A newly developed code [24], has been used for basis-set-dependent calculations and the results have been checked with Gaussian and Gamess-US ab initio codes as standard for assurance. To compute the first and second derivatives of energy with respect to nuclear coordinates, Peterson's code for Dunham analysis [25,26] has been used. The equilibrium distance of inter-nuclear separations, harmonic vibrational frequencies, vibrational anharmonicity factors and vibrational-rotational coupling constants have been offered as partial outcome of Dunham analysis throughout this paper. Comparison of calculated bond length and vibrational frequencies with standard ab initio codes noted above, gives us confidence in the reliability of Peterson's code.

Selecting the molecular set

A large set of diatomic molecules has been used for numerical HF calculations. We have selected six of them as prototypes, namely: H_2, C_2, N_2, CO, FH and F_2. These molecules may be regarded as a subset of the molecular set used in previous investigations [9]. Although we have done same calculations on the remaining members of that set as well as some diatomic molecules not considered by others, due to large amount of data we have hesitated in discussing them here. There are also two additional points that are worth mentioning. First, in contrast to neutral species, cationic and anionic species are biased with the kind of basis set used for algebraic calculation. The cations, due to their contracted wave functions, are usually better described using basis sets designed for neutral species, whereas anionic wave functions, due to their spatial extension, need to be augmented with diffuse functions to gain reliable results. So, the choice of mix of neutral, positively and negatively charged species could be misleading. Other members omitted in this paper are those for which no polarized-consistent basis sets exist in the coarse of preparing this paper and particularly, those molecules containing elements from the third and higher rows of the periodic table. So in this primary report, we have confined ourselves to the six molecules mentioned above. A more complete set of molecules is under consideration and the results will be offered in future, although many of the main results described in this paper also seem to be true for that extended set.

Results and discussion

Numerical HF calculations

As mentioned above, in last two decades, different numerical HF calculations have appeared in literature. They have been mainly used to calibrate and evaluate basis-set-dependent results. It seems there is a tendency to believe that numerical results are always superior to basis-set-dependent solutions. Although this is true as an ideal case, but it must be born in mind that, like basis set truncation error, there are "grid truncation" and "practical infinity truncation" errors in "finite difference" numerical solutions [23] of Hartree–Fock equations and without checking or estimating these errors, it is possible that results can be completely illusory. This is particularly important and problematic when older literature is used to derive numerical data. So, we have redone the entire numerical calculations using newly developed code described previously to test the quality of the older calculations.

To be sure about the effect of grid selection on the precision of the final results, two different grids have been used for calculations whereas the practical infinity has been estimated from previous experience. As has been discussed elsewhere [21,23], the Hartree–Fock wavefunction in diatomic molecules could be described in prolate spheroidal coordinates. So, the two grid parameters offer the numerical partitioning of the two independent variables (leaving the azimuthal angle as the third variable) in this coordinate system. Also, the practical infinity has been used as the asymptotic limit of boundary condition. First, a 115×115 grid has been used. Then a tighter 16×193 grid has been utilized for final calculations. Previous experience demonstrated the reliability of a 169×193 grid for many diatomic molecules containing first row elements [23]. In all cases, after the establishment of the equilibrium inter-nuclear distance (R_e), belonging to the minimum of the potential energy surface (with at least ±0.0001 Bohr precision), a complete scan of potential energy surface (PES) has been done from ($R_e-0.3$) to ($R_e+0.7$)

Table 1 Potential energy surface (PES) derived properties from Dunham analysis using numerical HF method

Molecule	Grid[a]	E_e(hartrees)	D_e(kcal.mol^{-1})[b]	R_e(Bohr)	ω_e(cm^{-1})	$\omega_e x_e$(cm^{-1})	α_e(cm^{-1})
H_2	169*193;35	−1.133667	83.88	1.3862	4,583.11	108.05	2.71053
C_2	169*193;40	−75.406685	18.48	2.3408	1,905.04	12.46	0.01693
CO	169*193;25	−112.792703	184.92	2.0821	2,426.78	11.27	0.01510
N_2	169*193;40	−108.996600	122.20	2.0133	2,730.45	10.90	0.01372
FH	169*193;40	−100.071294	101.59	1.6951	4,473.80	84.93	0.74869
F_2	169*193;40	−198.779813	−24.40	2.5063	1,267.43	6.58	0.00812

[a] (Grid size; practical infinity)

[b] Dissociation energies have been calculated using atomic HF total energies from reference [29]

with a step size of 0.05 Bohr. So, 21 points have been calculated in each case. These points have been subsequently utilized for a Dunham analysis. The performance of this step size and boundary selections, has been demonstrated elsewhere [27].

Table 1 offers the final results gained form potential energy surface (PES) analysis. Only the values generated with 169 × 193 grid has been shown in this table. A detailed analysis of the PES generated with the two grids mentioned above reveals interesting patterns. For the molecules H_2, C_2, N_2 and CO using these two grids the mean energy differences in the whole PES are always smaller than one microhartree, whereas the maximum difference does not exceed ∼ 1.5 microhartree. Consequently, for practical proposes, the final results must be regarded as accurate (standard chemical accuracy). On the other hand, for the FH molecule the maximum differences reaches ∼ 3.1 microhartree, although this also seems acceptable, but for F_2 the maximum difference goes to ∼ 48.9 microhartree, giving serious doubt about the reliability of the final results. It is interesting to note that F_2 has been mentioned as a special case in previous investigation [9].

Table 2 Comparison of total energies (in hartrees) derived from numerical HF solutions in this study and those cited in the literature

Molecule	R (Bohr)	Grid	Total energy	Reference
H_2	1.4	169*193;30	−1.13362957135	19
	1.4	319*595;65	−1.1336295715	38
	1.4	169*193;35	−1.1336295	Present work
C_2	2.358	?	−75.406565	9
	2.358	169*193;40	−75.406565	Present work
CO	2.132	81*161;?	−112.79095	16
	2.132	169*193;25	−122.7909073	39
	2.132	169*193;40	−122.790907	40
	2.132	169*193;25	−122.7909072	41
	2.132	169*193;25	−122.790906	19
	2.132	169*193;25	−122.790907	Present work
N_2	2.068	81*161;?	−108.99381	16
	2.068	229*229;25	−108.9938257	19
	2.068	169*193;40	−108.9938256	40
	2.068	793*793,40	−108.993825622	42
	2.068	169*193;40	−108.9938257	Present work
FH	1.7328	81*161;?	−100.07082	16
	1.7328	169*193;40	−100.070757	Present work
	1.7328	199*199;40	−100.0707578	Present work
F_2	2.668	?	−198.773323	9
	2.668	169*193;40	−198.773310	J. Kobus*
	2.668	247*439;160	−198.773443	J. Kobus*
	2.668	247*439;160	−198.773443	11
	2.668	169*193;40	−198.773310	Present work
	2.668	199*199;40	−198.773311	Present work

* Private communication

For a comparison, Table 2 offers the total energy values derived from previous calculations cited in the literature and those from the present study. As may be seen from this table, there is a good agreement between values in literature and those derived from our calculations except for FH and F_2. In the case of the FH molecule, the only cited value comes form the mid 1980s work of Pyykkö et al. [16]. Due to the lower precision of the computed total energies of that paper for N_2 and CO molecules relative to more recent calculations, and because of better agreement of our results with the more recent calculations, we have preferred values derived from our calculations for the FH molecule. It is worth mentioning that properties calculated for the FH molecule show a good agreement with numerical values of spectroscopic constants offered by Styszynski in his calculations on the FH molecule [28], although we did not have access to his total energies for direct comparison. On the other hand, in the case on the F_2 molecule, the problem originates from "grid truncation" error. As Table 2 offers, calculations using tighter grids change the situation dramatically. So, problems with the F_2 molecule in the previous study come from their underestimated total energy. Due to the demanding grids used by Kobus for high-precision computations on the F_2 molecule, it was impossible for us to do a complete scan with that grid on F_2's PES, so we have used an "additive approximation" to modify values due to grid truncation. These modified values have been used in checking the basis set extrapolated values.

As a final conclusion, it seems that in the case of the FH and F_2 molecules, the total energies used in the previous investigation [9] as exact numerical values are not the

Table 3 Calculated parameters from Dunham analysis of correlation-consistent-based PESs

Molecule	Basis set	E_e	D_e	R_e	ω_e	$\omega_e x_e$	α_e
H_2	cc-pVDZ	−1.128746	81.69	1.4134	4,583.9	105.59	2.55167
	cc-pVTZ	−1.132990	83.69	1.3879	4,587.4	108.56	2.73618
	cc-pVQZ	−1.133496	83.84	1.3866	4,582.0	107.52	2.69613
	cc-pV5Z	−1.133645	83.87	1.3862	4,583.3	108.02	2.70878
	cc-pV6Z	−1.133663	83.88	1.3862	4,583.3	108.05	2.71014
C_2	cc-pVDZ	−75.387049	13.97	2.3666	1,914.3	12.45	0.01635
	cc-pVTZ	−75.401450	17.65	2.3449	1,904.6	12.27	0.01684
	cc-pVQZ	−75.405784	18.40	2.3412	1,905.1	12.34	0.01690
	cc-pV5Z	−75.406545	18.45	2.3408	1,904.9	12.39	0.01692
	cc-pV6Z	−75.406659	18.47	2.3408	1,905.0	12.37	0.01692
CO	cc-pVDZ	−112.750151	176.69	2.0979	2,431.6	10.91	0.01467
	cc-pVTZ	−112.781813	183.28	2.0871	2,425.0	11.06	0.01488
	cc-pVQZ	−112.790626	184.75	2.0825	2,427.3	11.27	0.01507
	cc-pV5Z	−112.792412	184.85	2.0821	2,426.9	11.32	0.01510
	cc-pV6Z	−112.792649	184.90	2.0821	2,426.8	11.28	0.01510
N_2	cc-pVDZ	−108.955559	112.15	2.0358	2,758.3	11.20	0.01342
	cc-pVTZ	−108.986557	120.38	2.0165	2,731.7	10.72	0.01361
	cc-pVQZ	−108.994470	121.81	2.0137	2,729.7	10.82	0.01370
	cc-pV5Z	−108.996188	122.04	2.0133	2,730.3	10.86	0.01371
	cc-pV6Z	−108.996515	122.16	2.0133	2,730.4	10.87	0.01372
FH	cc-pVDZ	−100.019707	93.72	1.7036	4,440.8	86.78	0.76438
	cc-pVTZ	−100.058462	100.06	1.6969	4,481.8	84.67	0.74157
	cc-pVQZ	−100.068185	101.19	1.6949	4,477.2	84.06	0.74828
	cc-pV5Z	−100.070926	101.53	1.6950	4,475.1	85.38	0.74855
	cc-pV6Z	−100.071233	101.60	1.6951	4,474.1	85.08	0.74877
F_2	cc-pVDZ	−198.688936	−33.40	2.5467	1,180.1	5.60	0.00800
	cc-pVTZ	−198.758042	−25.32	2.5117	1,266.8	6.71	0.00803
	cc-pVQZ	−198.774463	−24.78	2.5086	1,263.8	6.50	0.00812
	cc-pV5Z	−198.779058	−24.60	2.5070	1,265.9	6.65	0.00812
	cc-pV6Z	−198.779645	−24.47	2.5066	1,266.6	6.55	0.00812

Total energies and bond distances are in atomic units, dissociation energies in kcal.mol^{-1} and other quantities in cm^{-1}

best choices. Now we are in a position to compare numerical values of total energies with those, which come form basis-set-dependent algebraic solutions.

Basis-set-dependent algebraic HF calculations

Both correlation-consistent and polarized-consistent basis sets have been used for construction of potential energy surfaces of molecules mentioned previously. The same initial, final and step size have been used to scan PESs as described in the previous section. Tables 3 and 4 contain the final results of Dunham analysis on PESs. A brief look at these tables offers some general trends. In most cases enlargement of the basis set causes the increase of dissociation energies as well as contraction of bond lengths. Aside from the F_2 molecule, in all other cases the bond lengths computed at cc-pV6Z and PC-4 completely coincide with those calculated with the numerical solution. The slight difference in the case of the F_2 molecule is in accord with the grid convergence problem discussed previously. On the other hand, there are less clear systematic trends in the case of spectroscopic constants. The computed properties utilizing the first member in both hierarchies, namely cc-pVDZ and PC-0, show a vast difference with the values calculated using other members of their family. So these two small basis sets seem unreliable relative to their more extended counterparts. The PC-0 basis set does not contain polarized functions so its erratic behavior does not seem strange. Therefore, this set has been excluded from further investigations in this paper.

A comparison with Table 1 confirms that for all molecules, except for the total energies that will be discussed in next section, all properties calculated from PES Dunham analysis show a nice convergence using both hierarchies. Particularly, the rate of convergence of spectroscopic properties is quite remarkable. For any practical purposes, the spectroscopic constants derived using PESs computed with cc-pVQZ and PC-2 basis sets could be viewed as the complete basis set limit. Moreover, there are cases for which the results originating from PESs constructed from the cc-pVTZ basis set could be viewed as reliable. The same trends also seem to hold for the dissociation energies. It is important to realize that such a fast convergence does not necessarily reveal the total energy convergence pattern. A balanced basis set with a truncation error, which is relatively independent of inter-nuclear distance, could produce an accurate PES nearly parallel with the exact PES computed in the complete basis set limit. On the other hand, it is evident from basis-set-dependent calculations that the basis set superposition error (BSSE) [30] completely diminished using extended basis sets and there is no serious need for any modification of the PES using methods like the counterpoise procedure [31].

So both correlation-consistent and polarized-consistent basis sets show a similar convergence pattern for spectro-

Table 4 Calculated parameters from Dunham analysis of polarized-consistent-based PESs

Molecule	Basis set	E_e	D_e	R_e	ω_e	$\omega_e x_e$	α_e
H_2	PC-0	−1.120671	80.30	1.3989	4,614.5	107.95	2.73423
	PC-1	−1.130291	83.53	1.3997	4,640.1	108.34	2.64134
	PC-2	−1.133238	83.76	1.3880	4,589.2	107.74	2.70340
	PC-3	−1.133650	83.87	1.3863	4,583.5	108.10	2.70942
	PC-4	−1.133665	83.88	1.3862	4,583.1	108.05	2.71053
C_2	PC-0	−75.158103	1.62	2.4323	1,702.7	10.29	0.01587
	PC-1	−75.358568	15.11	2.3486	1,932.1	12.79	0.01675
	PC-2	−75.400652	18.72	2.3404	1,907.9	12.39	0.01702
	PC-3	−75.405742	18.48	2.3406	1,906.2	12.36	0.01692
	PC-4	−75.406610	18.38	2.3408	1,905.1	12.34	0.01693
CO	PC-0	−112.333871	131.57	2.1770	2,128.3	8.96	0.01426
	PC-1	−112.712424	178.73	2.0935	2,447.9	11.37	0.01499
	PC-2	−112.783296	182.77	2.0830	2,426.4	11.18	0.01515
	PC-3	−112.791669	183.11	2.0821	2,427.9	11.32	0.01508
	PC-4	−112.792617	182.99	2.0821	2,426.8	11.29	0.01511
N_2	PC-0	−108.508147	35.86	2.1284	2,312.7	9.47	0.01341
	PC-1	−108.915315	113.35	2.0276	2,761.2	11.63	0.01380
	PC-2	−108.987260	121.54	2.0137	2,732.1	10.83	0.01382
	PC-3	−108.995599	122.23	2.0132	2,731.2	10.88	0.01371
	PC-4	−108.996524	122.19	2.0133	2,730.5	10.89	0.01372
FH	PC-0	−99.670078	88.59	1.7557	4,287.9	73.85	0.65588
	PC-1	−99.989216	96.04	1.7117	4,451.5	93.00	0.75736
	PC–2	−100.063578	100.16	1.6946	4,474.2	86.98	0.76926
	PC-3	−100.070673	100.37	1.6951	4,474.9	84.61	0.74854
	PC-4	−100.071239	100.33	1.6951	4,473.9	85.45	0.74851
F_2^a	PC-0	−198.032516	−20.44	2.6613	1,305.7	4.97	0.00537
	PC-1	−198.618064	−35.82	2.5634	1,176.3	5.32	0.00744
	PC-2	−198.763665	−27.93	2.5141	1,258.6	6.56	0.00801
	PC-3	−198.778343	−27.05	2.5065	1,267.1	6.54	0.00811
	PC-4	−198.779670	−27.00	2.5065	1,267.3	6.68	0.00812

Total energies and bond distances are in atomic units, dissociation energies in kcal.mol^{-1} and other quantities in cm^{-1}

[a] There is almost 2.5 kcal.mol^{-1} difference between D_e values derived from PC-4 and those from numerical HF and cc-pV6Z calculations. This difference comes from atomic HF calculations that do not use degenerate atomic orbitals (p_x, p_y, p_z) of fluorine atom. The same problem can be observed elsewhere (see [43])

scopic constants as well as dissociation energies. Regarding the larger number of basis functions used in each of correlation consistent basis sets, to reach the same accuracy relative to their PC counterparts, it seems that using polarized consistent basis sets are more economic to gain a reliable PES.

Convergence pattern of total energies in basis-set-dependent solutions

As discussed in previous section, both cc-pVxZ and PC-*n* hierarchies seem to reach the complete basis set limit of spectroscopic constants derived from Dunham analysis with reasonable accuracy. Although this is promising, it is evident that in diatomic molecules constructed by elements from higher rows of the periodic table as well as polyatomic molecules, the same accuracy could not be achieved by direct calculation and an extrapolation scheme is desirable to reach the complete basis set limit for total energies. On the other hand, no one has yet done a systematic study on the performance of extrapolation schemes relative to different inter-nuclear distances. It is evident that any attempt to survey the PES around the equilibrium geometry demands an extrapolation scheme insensitive to slight geometrical changes. An attempt will be done in this section to pursue the convergence pattern of basis-set-dependent solutions. The extrapolation and related material will be discussed in hte next section.

Although in a previous study [9], both total and dissociation energies have been used to seek a reliable extrapolation scheme, we hesitate before using dissociation energies in our analysis. As has been discussed in a previous study [9], there is no need for high angular basis functions for a complete description of HF wavefunctions for atoms, whereas for molecules such polarized basis functions are certainly mandatory for a proper description. So, there is no reason to believe that the same dependency of total HF energy and maximum angular momentum of proper basis set used persist for both atoms and corresponding molecules. On the other hand, for atoms the HF total energies have been calculated precisely and tabulated [29] so, after extrapolating the total energies of diatomic molecules, it is possible to reach the complete basis set limit of dissociation energies using these tabulated data. It is important to note that in polyatomic molecules the dissociation process usually produces smaller molecular fragments, and consequently in those cases the extrapolation could be done both for the main molecule and for its fragments to estimate the exact dissociation energy at Hartree–Fock level.

Table 5 Total energy differences of basis-set-dependent calculations with numerical solution at three different inter-nuclear distances

Molecule	Basis set	R_e−0.3	R_e	R_e+0.7	Basis set	R_e−0.3	R_e	R_e+0.7
H_2	cc-pVDZ	10,827	5,074	2,299	PC-1	6,982	3,414	3,085
	cc-pVTZ	1,109	678	538	PC-2	883	430	377
	cc-pVQZ	221	172	146	PC-3	26	17	24
	cc-pV5Z	26	22	30	PC-4	2	2	3
	cc-pV6Z	4	4	11				
	Numerical	−1.107687	−1.133667	−1.082686		−1.107687	−1.133667	−1.082686
C_2	cc-pVDZ	31,428	19,920	15,760	PC-1	53,064	48,143	48,249
	cc-pVTZ	6,950	5,242	4,499	PC-2	6,139	6,033	6,338
	cc-pVQZ	1,069	901	904	PC-3	919	943	1100
	cc-pV5Z	165	140	126	PC-4	78	75	84
	cc-pV6Z	35	26	24				
	Numerical	−75.354434	−75.406685	−75.308816		−75.354434	−75.406685	−75.308816
CO	cc-pVDZ	56,298	42,747	38,819	PC-1	92,185	80,382	78,480
	cc-pVTZ	14,677	10,910	9,757	PC-2	10,334	9,408	9,172
	cc-pVQZ	2,443	2,077	2,197	PC-3	1,109	1,034	1,155
	cc-pV5Z	361	291	295	PC-4	100	86	86
	cc-pV6Z	77	54	53				
	Numerical	−112.6922534	−122.792703	−112.615022		−112.692253	−122.792703	−112.615022
N_2	cc-pVDZ	69,000	41,566	33,909	PC-1	10,0576	81,494	76,217
	cc-pVTZ	13,598	10,053	9,038	PC-2	10,282	9,340	9,201
	cc-pVQZ	2,470	2,130	2,056	PC-3	1,065	1,001	1,122
	cc-pV5Z	460	412	383	PC-4	91	76	81
	cc-pV6Z	96	85	93				
	Numerical	−108.8671449	−108.996600	−108.7653578		−108.8671449	−108.996600	−108.7653578
FH	cc-pVDZ	54,560	51,613	48,763	PC-1	88,515	82,179	7,7791
	cc-pVTZ	13,668	12,833	12,842	PC-2	7,784	7,716	7,500
	cc-pVQZ	3,127	3,109	3,264	PC-3	644	621	664
	cc-pV5Z	357	368	408	PC-4	64	55	56
	cc-pV6Z	56	61	74				
	Numerical	−100.0218915	−100.071294	−99.988419		−100.0218915	−100.071294	−99.988419
F_2	cc-pVDZ	96,902	91,296	78,595	PC-1	17,2206	16,2602	14,7059
	cc-pVTZ	23,150	21,778	20,974	PC-2	17,785	16,164	14,358
	cc-pVQZ	5,749	5,351	4,780	PC-3	1,506	1,470	1,453
	cc-pV5Z	845	755	521	PC-4	165	143	95
	cc-pV6Z	207	168	53				
	Numerical	−198.743389	−198.779813	−198.707934		−198.743389	−198.779813	−198.707934

Energy differences offered in microhartrees and distances in atomic units

Table 5 contains the total energies of numerical solution and the difference with those derived from basis set dependent solution of Hartree-Fock equations in three inter-nuclear separations. In the case of the F_2 molecule as discussed in previous sections, the numerical standard grid in this study is not sufficient for microhartree precision, so an additive approximation has been used to estimate the numerical Hartree–Fock total energy for the three inter-nuclear distances considered in this paper. Equation (1) offers this estimation method. According to our previous experiences, it seems to us that the results gained by this procedure are superior to those from the standard grid.

$$E(R)_{\text{exact}} = E(R)_{(169\times 193)} + \Delta E(R)_{(169\times 193-115\times 115)} + \Delta E(R_0 = 2.668\text{a}_0)_{(247\times 439-193\times 169)} \qquad (1)$$

In this equation, the second and third terms modify the result for the dependence of inter-nuclear value on grid quality and grid truncation errors, respectively. Utilizing both hierarchies it is evident that total energies, in contrast to correlation energy [8], converge with a remarkable rate relative to basis sets cardinal numbers. Even a brief look at Table 5, reveals that the addition of new basis functions with higher angular numbers to each basis set and using the next member in both hierarchies (with one higher cardinal number), causes the error to be reduced by almost an order of magnitude. In this regard, with cc-pV5Z or PC-3, millihartree accuracy could be reached. On the other hand, in all cases cc-pV6Z and PC-4 yield an ideal accuracy and practically could be regarded as the complete basis set limit. This is in accord with previous investigation [9].

Although, as mentioned above, there are many common features in the convergence pattern of both hierarchies, there are also subtle differences. By comparison of the similar members of both hierarchies (1 = DZ, 2 = TZ, 3 = QZ and 4 = 5Z), it is evident that in all cases (except the PC-1 basis set) the absolute accuracy of the polarized-consistent basis sets is superior to the correlation-consistent basis set, independent of the largest member of the correlation-consistent hierarchy. A comparable accuracy could be gained only with the PC-4 and cc-pV6Z basis sets. Due to the smaller size of PC-4 relative to cc-pV6Z it seems that for direct calculations the former is superior. So these observations confirm the higher rate of convergence of polarized-consistent basis sets, relative to correlation-consistent basis sets.

As mentioned previously, Table 5 offers the total energy difference for three separate inter-nuclear distances, namely, at R_e, $R_e - 0.3$ and $R_e + 0.7$. In contrast to our expectation, we have not found a general trend for the convergence pattern when changing the inter-nuclear distance. Although in many cases the contraction of bond length deteriorates the quality of basis-set-dependent results, there are also exceptions to this trend. This is particularly true for basis sets with higher cardinal numbers. It seems plausible that this behavior of the convergence for different inter-nuclear separations is due to the dependence of the optimum polarized functions exponents on the internuclear distance. The basis set exponents are derived as a suitable average for a series of molecules. For the present cases, the convergence pattern depends on where the exponent is relative to the optimum value for the specific system. On the other hand, with the exception of F_2, the most extended basis sets in both hierarchies (PC-4 and cc-pV6Z) yield absolute errors relatively independent of inter-nuclear distance. Whether this independence is an exception or a general rule remains to be considered in future studies.

Extrapolation to the complete basis set limit

Diatomic molecules

As demonstrated in previous section, it is possible to reach millihartree and even near microhartree precision easily with direct computation for the molecular collection under study. Although the rate of convergence of basis-set-derived total energies relative to the complete basis set limit are remarkable, it is evident that in larger systems with more electrons the same accuracy could not be reached with direct computation. A reliable estimation of Hartree–Fock energies is even more important for large systems. This is particularly vital in the case of isodesmic reactions [32], which could be used for a reliable estimation of thermodynamic quantities of large polyatomic molecules without considering the correlation energy contribution directly.

In a previous study, Halkier et al. [9], used different equations for extrapolation and came to the conclusion that an exponential form is the best among others (Eq. (2)).

$$E_R(L) = E_R(\infty) + Ae^{-\mathrm{BL}} \tag{2}$$

In this equation $E_R(L)$ is the total energy computed at the inter-nuclear distance R and L denotes the highest angular functions of the basis sets used in extrapolation. In cc-pVxZ hierarchy L ranges from 2 (cc-pVDZ) to 6 (cc-pV6Z), whereas, in the PC-*n* hierarchy it ranges from $L = 2$ (PC-1) to $L = 5$ (PC-4). Also, $E_R(\infty)$ is the total energy in complete basis set limit. The quantities, A and B are empirical constants without physical significance. Since there are three unknown quantities in this equation ($E_R(\infty)$, A and B), at least three consecutive basis sets are needed for extrapolation. Restricting ourselves to this minimum requirement, we have used three member groups (L, $L + 1$, $L + 2$) in each hierarchy. In this regard, the correlation-consistent hierarchy could be extrapolated with $\{(2, 3, 4), (3, 4, 5)$ and $(4, 5, 6)\}$ collections whereas, for the polarized-consistent hierarchy the possible collections are $\{(2, 3, 4)$ and $(3, 4, 5)\}$.

Table 6 depicts the final results. Only the total energy differences between extrapolated and numerical values have been offered for comparison. Different trends could be deduced from this table. It is evident that extrapolations using collections containing the smallest members of each hierarchy could not be regarded reliable. This is not surprising since the absolute errors according to Table 5 decrease an order of magnitude with every unit increase in cardinal number and so the smallest basis sets in each hierarchy by no means could be viewed as reliable. In this regard the general validity of extrapolation schemes appearing in literature using the smallest basis sets of these hierarchies seems questionable [33]. So, we hesitate before considering them further.

Although the situation in the case of the (3, 4, 5) collection for correlation consistentcy is not as worse as the (2, 3, 4) collection containing the cc-pVDZ basis set, the errors are almost in the same order of magnitude as those for the cc-pV5Z basis set. The interesting trend is the general overestimation of the complete basis set limit of this collection. The exact total energy is somewhere between the cc-pV5Z and extrapolated value of the (3, 4, 5) collection. The same specifications also hold for the extrapolated total energies from PC-3 and the (2, 3, 4) collection, although in this case the extrapolated values are slightly better than those corresponding to the correlation-consistent hierarchy. The general overestimation of extrapolated values also holds for this set.

The situation is reversed while utilizing the (4, 5, 6) collection for extrapolation of correlation-consistent basis sets. This time the extrapolated values are generally better than cc-pV6Z. A general underestimation could be traced in extrapolated values in contrast to the (3, 4, 5) collection. The same is true for the (3, 4, 5) collection extrapolated total energies of the polarized-consistent hierarchy. They are usually better than PC-4 but overestimated the complete basis set limit, although this overestimation is not as much as for the total energies derived from the (2, 3, 4) collection. In this regard it seems that for a reasonable extrapolation only the three most extended basis sets in each hierarchy could be assumed reasonable. Even in these cases, the errors of the most extended basis set and the extrapolated values are in the same order of magnitude. This is not surprising since the PC-4 and cc-pV6Z attain remarkable accuracy in the molecules studied in this paper.

It is also interesting to consider the A and B values of Eq. (2) in the fitting procedure. In contrast to what we expected according to previous study [9], not only a vast difference has been found among A and B values corresponding to different molecules and fitting collections, but also regarding to inter-nuclear distance. This variance is more pronounced for A than B. So no general predictable pattern seems to exist for these quantities in fitting the collections studied. We hesitate to consider them further in this study. It is important to note that this observation does not invalidate the "two-point" extrapolation proposed in previous study [9]. This point will be discussed further in the next section.

Table 6 Total energy differences (microhartree) of extrapolated as well as some basis-sets-derived values with numerical HF solutions

Molecule	Extrap. type*	$R_e-0.3$	R_e	$R_e+0.7$	Extrap. type*	$R_e-0.3$	R_e	$R_e+0.7$
H_2	(2, 3, 4)	132	106	34	(2, 3, 4)	−114	−49	−29
	(3, 4, 5)	−29	−41	−19	(3, 4, 5)	1	1	2
	(4, 5, 6)	1	2	7				
	cc-pVQZ	221	172	146	PC-2	883	430	377
	cc-pV5Z	26	22	30	PC-3	26	17	24
	cc-pV6Z	4	4	11	PC-4	2	2	3
C_2	(2, 3, 4)	−791	−922	−782	(2, 3, 4)	266	243	352
	(3, 4, 5)	1	−22	−89	(3, 4, 5)	−84	−103	−160
	(4, 5, 6)	13	6	9				
	cc-pVQZ	1,069	901	904	PC-2	6,139	6,033	6,338
	cc-pV5Z	165	140	126	PC-3	919	943	1,100
	cc-pV6Z	35	26	24	PC-4	78	75	84
CO	(2, 3, 4)	−2650	−1315	−461	(2, 3, 4)	−63	−86	106
	(3, 4, 5)	−66	−162	−344	(3, 4, 5)	−24	−35	−78
	(4, 5, 6)	32	18	18				
	cc-pVQZ	2,443	2,077	2,197	PC-2	10,334	9,408	9,172
	cc-pV5Z	361	291	295	PC-3	1,109	1,034	1,155
	cc-pV6Z	77	54	53	PC-4	100	86	86
N_2	(2, 3, 4)	−327	−531	−669	(2, 3, 4)	17	−89	15
	(3, 4, 5)	17	−64	−144	(3, 4, 5)	−24	−39	−73
	(4, 5, 6)	16	8	32				
	cc-pVQZ	2,470	2,130	2,056	PC-2	10,282	9,340	9,201
	cc-pV5Z	460	412	383	PC-3	1,065	1,001	1,122
	cc-pV6Z	96	85	93	PC-4	91	76	81
FH	(2, 3, 4)	−534	−145	−218	(2, 3, 4)	−49	−126	−72
	(3, 4, 5)	−630	−708	−805	(3, 4, 5)	13	6	−3
	(4, 5, 6)	19	22	30				
	cc-pVQZ	3,085	3,064	3,219	PC-2	7,742	7,671	7,455
	cc-pV5Z	315	323	363	PC-3	602	576	619
	cc-pV6Z	14	16	29	PC-4	22	10	11
F_2	(2, 3, 4)	376	268	−1550	(2, 3, 4)	−412	−169	63
	(3, 4, 5)	−1079	−1030	−999	(3, 4, 5)	45	11	−65
	(4, 5, 6)	112	82	−5				
	cc-pVQZ	5,749	5,351	4,780	PC-2	17,785	16,164	14,358
	cc-pV5Z	845	755	521	PC-3	1,506	1,470	1,453
	cc-pV6Z	207	168	53	PC-4	165	143	95

* See text for notation used in this table

Polyatomic molecules

To check the reliability of the extrapolation scheme of the (3, 4, 5) collection using contracted polarized-consistent basis sets, as the most economic and reliable scheme, two polyatomic molecules, namely H_2O and CO_2, have been investigated. Unfortunately in contrast to diatomic molecules no general numerical Hartree–Fock method exists for polyatomic molecules so, in the latter case, very large basis sets calculations could be used as exact values. In this regard the “distributed universal even-tempered” basis sets (DUET) seems to be the best candidate (for a recent review see [34]). A decade of development and detailed comparison with numerical results on diatomic molecules [34], makes us confident about DUET’s reliability for polyatomic molecules. Particularly for the two molecules mentioned above, there are reliable estimates of total Hartree–Fock energy in their experimental geometry (see [35] for H_2O and [36] for CO_2).

Table 7 contains the total energies calculated using the PC-2, PC-3 and PC-4 basis sets as well as extrapolated values using the (3, 4, 5) collection. Like the diatomic molecules considered previously in this paper, a considerable improvement could be achieved by going to more extended basis sets of this hierarchy. On the other hand, the extrapolated total energies reveal the same slight overestimation for these two cases as could be observed for the diatomic molecules considered in Table 6. But what is more important is the quality of the extrapolated total energies. In both cases, the extrapolated total energies are considerably better than the total energies derived from the PC-4 basis set.

So, it seems that the extrapolation procedure used is at least as successful in the case of polyatomic molecules as for the diatomic molecules considered herein.

Final conclusion and future prospects

A relatively complete survey of basis set convergence pattern on a small subset of molecules has been carried out in this paper. Our computations on both numerical and basis-set-dependent Hartree–Fock methods demonstrate that there are very general trends in the basis set convergence pattern. In

Table 7 Total as well as difference energies of water and carbon dioxide using PC basis sets and extrapolated values relative to exact values

Molecule	Basis set	Total energy*	Difference*
H_2O***	PC-2	−76.061108	6,380
	PC-3	−76.066894	594
	PC-4	−76.067434	54
	(3, 4, 5)	−76.067490	−2
	exact**	−76.067488	–
CO_2***	PC-2	−187.710281	15,127
	PC-3	−187.723900	1,508
	PC-4	−187.725286	122
	(3, 4, 5)	−187.725443	−35
	exact**	−187.725408	–

* Total energies offered in hartree with energy differences in microhartree

** See references [35] for water and [36] for carbon dioxide.

*** The geometries: H_2O ($R_{O–H}$= 1.80885 Bohr, ∠ H-O-H= 104.52°) and linear CO_2 ($R_{C–O}$ = 1.160 Angstroms)

particular, it seems that the polarized-consistent hierarchy is a reliable and economic set to reach the complete basis set limit of the PES derived spectroscopic constants as well as total energies with a considerable convergence rate. This is promising since the correlation energy estimation needs more extended basis sets and computationally demanding methods.

On the other hand, thanks to current hardware technology and high performance computer codes, in the case of small molecules it is possible to use direct computation as the best method to reach the complete basis set limit with microhartree accuracy, but in the case of larger systems an extrapolation scheme seems mandatory. The simple exponential-type functions like Eq. (2) work relatively well, but it seems that in the case of the polarized-consistent hierarchy a general slight overestimation exists. So possible modification of Eq. (2) opens the door for a better estimation of the complete basis set limit. Although any complication of this extrapolation equation may increase the number of fitting parameters and so make it mandatory to use basis sets with lower cardinal numbers (and consequently a lower precision), using a "two-point" exponential equation (with a fixed B value) [9] with a "modification term" could possibly help the better estimation of the complete basis set limit. Such a modification is one of the main topics in our future research program.

Among the problems left to be considered, knowing the fact that density-functional-based methods behave like the Hartree–Fock method in their basis set convergence pattern [10–12,37], polarized-consistent basis sets and the extrapolation scheme described in this paper could also be tested for them. These issues will be discussed in a separate paper.

Acknowledgements Shant Shahbazian is grateful to Jack Kobus and Kirk Peterson for making their computer codes available to him, to Frank Jensen for making the polarized-consistent basis sets available to him, and to D. Moncrieff for drawing his attention to the high-quality results for the CO_2 molecule. Also special thanks to Jack Kobus for his valuable comments to Shant Shahbazian during the research program and to an unknown referee for his or her elegant suggestions.

References

1. Dunning TH Jr (2000) J Phys Chem A 104:9062
2. Dunning TH Jr (1989) J Chem Phys 90:1007
3. Feller D (1992) J Chem Phys 96:6104
4. Feller D (1993) J Chem Phys 98:7059
5. Martin JML (1996) Chem Phys Lett 259:669
6. Martin JML, Taylor PR (1997) J Chem Phys 106:8620
7. Varandas AJC (2000) J Chem Phys 113:8880
8. Helgaker T, Klopper W, Koch H, Noga J (1997) J Chem Phys 106:9639
9. Halkier A, Helgaker T, Jorgenson P, Klopper W, Olsen J (1999) Chem Phys Lett 302:437
10. Jensen F (2001) J Chem Phys 115:9113
11. Jensen F (2002) J Chem Phys 116:7372
12. Jensen F (2002) J Chem Phys 117:9234
13. Christiansen PA, McCullough EA Jr (1977) J Chem Phys 67:1877
14. Laaksonen L, Pyykkö P, Sundholm D (1983) Int J Quantum Chem 23:309
15. Laaksonen L, Pyykkö P, Sundholm D (1983) Int J Quantum Chem 23:319
16. Sundholm D, Pyykkö P, Laaksonen L (1985) Mol Phys 56:1411
17. Pyykko P, Sundholm D, Laaksonen L (1987) Mol Phys 60:597
18. Pyykko P, Dircksen GHF, Müller-Plathe F (1987) Chem Phys Lett 134:575
19. Kobus J (1993) Chem Phys Lett 202:7
20. Kobus J (1994) Comp Phys Comm 78:247
21. Kobus J (1996) Comp Phys Comm 98:346
22. Kobus J (1997) Adv Quantum Chem 28:1
23. Kobus J Handbook of molecular physics and quantum chemistry, Vol. 2, Part 7, Chap 3
24. Sajadi A, Shahbazian S FER-BOS-HF computer code (under construction). This code has been developed for comparison of mean field approximation for both fermionic and bosonic systems. It will be described completely in a separate paper.
25. Dunham JL (1932) Phys Rev 41:721
26. Peterson KA Fit1D computer code, University of Wisconsin-Madison, March 1989
27. Peterson KA, Dunning TH Jr (1997) J Mol Struct (Theochem), 400:93
28. Styszynski J (2000) Chem Phys Lett 317:351
29. Fischer CF (1977) The Hartree–Fock method for atoms: a numerical approach. Wiley, New York
30. Chalasinski G, Szczesniak MM (2000) Chem Rev 100:4227, and references therein
31. Boys SF, Bernardi F (1970) Mol Phys 19:553
32. Hehre WJ, Radom L, Schleyer PvR, Pople JA (1986) Ab initio molecular orbital theory. Wiley, New York
33. Truhlar DG (1998) Chem Phys Lett 294:45
34. Wilson S, Moncrieff D, Kobus J (2001) In: Maruani J (ed.) New trends in quantum systems in chemistry and physics. Kluwer, Dordrecht, Vol 1, p. 115
35. Moncrieff D, Wilson S (1996) J Phys B At Mol Opt Phys 29:6009
36. Moncrieff D, Wilson S (1995) J Phys B At Mol Opt Phys 28:4007
37. Christensen KAa, Jensen F (2000) Chem Phys Lett 317:400
38. Jensen F (1999) J Chem Phys 110:6601
39. Kobus J, Moncrieff D, Wilson S (1995) Mol Phys 86:1315
40. Kobus J, Moncrieff D, Wilson S (2000) Phys Rev A 62 No. 062503
41. Kobus J, Moncrieff D, Wilson S (1994) J Phys B 27:5139
42. Kobus J, Quiney HM, Wilson S (2001) J Phys B 34:2045
43. Peterson KA (1998) J Chem Phys 108:4761