

01

The complete basis set limits of the correlation energy of diatomic molecules using distributive s -Gaussian functions

© V.N. Glushkov, S.I. Fesenko

Oles Honchar Dnepropetrovsk National University,
49000 Dnipro, Ukraine

e-mail: v.n_glushkov@yahoo.com, sergey.fesenko@gmail.com

Received February 12, 2022

Revised May 31, 2022

Accepted May 31, 2022

In this paper possibilities of distributive s -Gaussian functions to estimate the second-order correlation energy (MP2) are investigated to the complete basis set limit. In contrast to standard atom-centered, the proposed bases consist of functions distributed along the molecular axis and off-axis functions. We propose and investigate the possibilities of one of the models for choosing the alignment of such functions. The model is characterized by a set of subsets of off-axis s -functions uniformly distributed along circles whose planes are perpendicular to the molecular axis. The function parameters are determined by minimizing the Hartree-Fock energy (for functions located along the molecular axis) and minimizing the Hylleraas functional (for off-axis s -functions). The resulting basis sequences, combined with known extrapolation models, are used to predict the MP2 energy in the complete basis set limit. The efficiency of the models is demonstrated by calculating the energy MP2 of simple H_2 and LiH molecules.

Keywords: distributive off-axis functions, correlation energy, extrapolation models.

DOI: 10.21883/EOS.2022.09.54818.3292-22

Introduction

Modern methods of quantum chemistry have become so sophisticated, that the main source of errors in calculations *ab initio* is considered the incompleteness of single-particle basis sets. Therefore, the methods aimed at constructing new basis sets, as well as calculations of molecular properties in the limit of complete basis sets, have been actual for a long time. At the same time, the information on the convergence of non-empirical methods for calculating the characteristics of a system is extremely important, since it allows to perform extrapolation and estimate the errors in calculations of large systems, where the basis sets, which deliver the results close to the basic limit cannot be realized. It is known, that the convergence of the correlation energy with increasing basis set dimension is slower, than the convergence of the energy in the Hartree-Fock (HF) approximation. One of the approaches to speed up convergence uses the basis sets, which explicitly depend on the interelectronic distance r_{12} , the so-called R12 and F12 [1–3] schemes. Various extrapolation techniques, which allow to estimate the correlation energy in the complete basis set limit can be considered as an alternative approach [4–17]. In this case, for extrapolation it is sufficient to calculate the correlation energy obtained in two or three basis sets of various dimensions (depending on the model used). It should be noted, that one of the important points in the construction of such basis sets is the generated functions hierarchy concept, which leads to a systematic improvement of the results from one level in the hierarchical ladder to another. An example of such bases, in particular, are the

popular correlation-consistent basis sets aug-cc-pVXZ [18–20]. In this notation, X corresponds to the angular moment value and is called the cardinal number. For example, $X = 2$ defines a double-zeta basis set, aug-cc-pVDZ, $X = 3$ corresponds to a triple-zeta basis set, aug-cc-pVTZ and t.e. The hierarchical structure of the basis set can be used to build extrapolation models to predict the molecular characteristics values, in particular the correlation energy, in the CBS-complete basis set.

The purpose of this work is to evaluate the capabilities of distributive molecular basis set models for predicting the values of molecular characteristics in the limit of the complete basis set. In particular, we consider an attempt to construct such optimal basis sets consisting of s -functions centered outside the molecular axis (off-axis functions) to determine the correlation energy in the second order of Møller-Plesset perturbation theory (MP2). The resulting basis sequences, combined with popular extrapolation models, are used to predict the MP2 energy within the limit of the complete basis set. The questions of choosing cardinal numbers for distributive basis sets are discussed. The efficiency of the models is demonstrated by the example of calculating the energy MP2 of simple molecules H_2 and LiH .

Extrapolation model

As you know, in the general case, the extrapolation procedure is the transfer of conclusions made regarding any part of objects or phenomena to the entire set of these objects. In respect to our complete basis set problem,

the extrapolation procedure is as follows: the investigated property E (for example, energy, bond length and etc.) is calculated at a given theory level (for example, MP2), using at least three basis sets (for example, aug-cc-pVDZ, aug-cc-pVTZ and aug-cc-pVQZ). Then the obtained values are adjusted to the equation, which is determined by the model function. In this study we considered three most popular models:

$$E(k) = E_{\text{CBS}} + F e^{-\beta k}, \quad (1)$$

$$E(k) = E_{\text{CBS}} + F k^{-3}, \quad (2)$$

$$E(k) = E_{\text{CBS}} + F k^{-\beta}. \quad (3)$$

Here E_{CBS} is the energy value in the complete basis set limit, F and β — model parameters, k — the so-called cardinal number associated with the basis set. For example, as noted in the introduction, for aug-cc-pVDZ $k = 2$, for aug-cc-pVTZ $k = 3$ and etc.

The exponential form (1) was first considered in the study [7] and proved to be good for determining the energy in the HF approximation. A two-parameter model (2) was proposed in the study [11] to determine the correlation energy. The power extrapolation curve (3) was used in the studies by Bakowies [15].

The equations (1) and (3) contain three parameters E_{CBS} , F and β , while equation (2) requires only two parameters E_{CBS} and F . Below, using the example of the model function (1) and the basis sets aug-cc-pVDZ, aug-cc-pVTZ and aug-cc-pVQZ ($k = 2, 3, 4$), we will show, how these parameters can be determined. For this, the previously calculated values $E(k) = E_k \rightarrow E_2, E_3, E_4$ in the corresponding basis sets are used. On the other hand, the energies can be determined by extrapolation equations (1), i.e.

$$\begin{aligned} E_2 &= E_{\text{CBS}} + F e^{-2\beta}, \\ E_3 &= E_{\text{CBS}} + F e^{-3\beta}, \\ E_4 &= E_{\text{CBS}} + F e^{-4\beta}. \end{aligned} \quad (4)$$

Thus, we have a system of nonlinear equations (4) with respect to the unknowns E_{CBS} , F and β , which for this case has an analytical solution [21]:

$$\begin{aligned} E_{\text{CBS}} &= \frac{E_2 E_4 - E_3^2}{E_2 - 2E_3 + E_4}, \\ F &= \frac{E_2 - E_3^4}{(E_3 - E_4)^2 (E_2 - 2E_3 + E_4)}, \\ \beta &= \ln \frac{E_2 - E_3}{E_3 - E_4}. \end{aligned}$$

In a similar way, the parameters for other extrapolation models can be found. In this case, the corresponding systems of nonlinear equations, as a rule, do not have an analytical solution and should be solved numerically. It is useful to note, that there are online programs in the literature, which for the given values of energies (or other properties) and arbitrary cardinal numbers allow to determine the desired characteristics in the complete basis set limit [21] (complete basis set limit extrapolation calculator).

Table 1. Sequences of basis sets from off-axis of s -functions for the ground state of the H_2 molecule at interatomic distance $R = 1.4$ bohr, total Hartree-Fock energy $E^{\text{HF}}(30s) = -1.13362933$ Hartree, second-order correlation energy $E^{(2)}$ given in milliHartree and calculated in the basis sets $2 * (15s + k * 12_{\text{off}}) + 12_{\text{mid}}$, k — cardinal numbers.

k	Basis sets structure	$E^{(2)}$	$E^{(2)}/E_{\text{exact}}^{(2)}, \%$
	distributive basis ^(a)		
1	$2 * (15s + 12_{\text{off}}) + 12_{\text{mid}}(66s)$	-32.121	93.84
2	$2 * (15s + 2 * 12_{\text{off}}) + 12_{\text{mid}}(90s)$	-33.429	97.66
3	$2 * (15s + 3 * 12_{\text{off}}) + 12_{\text{mid}}(114s)$	-33.649	98.30
4	$2 * (15s + 4 * 12_{\text{off}}) + 12_{\text{mid}}(138s)$	-33.776	98.67
5	$2 * (15s + 5 * 12_{\text{off}}) + 12_{\text{mid}}(162s)$	-33.850	98.89
	atom-centered basis ^(b)		
2	aug-cc-pVDZ	-27.29	79.73
3	aug-cc-pVTZ	-31.99	93.46
4	aug-cc-pVQZ	-33.25	97.14
	MP2-R12/A (9s8p4d1f) „exact“	-34.23	100.00

Note. ^(a) — this study, ^(b) — [25].

Discussion of results

The first part of this section deals with the construction of basis sets from spherical Gaussian functions

$$\begin{aligned} \chi_p(r) &= \exp[-\xi_p |r - R_p|^2] \\ &= \exp -\xi_p [(x - X_p)^2 + (y - Y_p)^2 + (z - Z_p)^2] \end{aligned} \quad (5)$$

to calculate the correlation energy, $E^{(2)}$, in the second order of perturbation theory. The second part discusses the features of extrapolation of distributive basis sets, in particular, the choice of cardinal numbers.

The distributive basis sets have a long history and have shown their effectiveness in energy calculations with the HF [22,23] wave function and the Coulson-Fischer [24] function. The parameters of exponent ξ_p and centering R_p were determined based on different approaches. In one of them they were completely optimized from the requirement of HF energy minimum. In another approach the parameters were determined using the so-called anharmonic distribution model [23]. In this case the centers of functions were located on the molecular axis. However, it is clear, that a correct description of the correlation energy requires an extension of the basis by functions located outside the molecular axis, thus simulating the effects of polarization. The choice of simple molecules H_2 and LiH is based on the existing for them precision calculations of the energies MP2 of the ground state [25]. It is also worth noting, that the use of distributive functions is not limited to s -functions. It is also possible to use Gaussian functions with a large angular momentum quantum number. However, in this case the variational procedure will be more complicated.

In this paper we propose and explore the possibilities of one of the models for the distribution of such functions. It is determined by subsets off-axis of s -functions uniformly

Table 2. Convergence of the correlation energy with increasing dimension of off-axis of s -basis for the ground state of the LiH molecule at $R = 3.015$ bohr, HF energy $E^{\text{HF}}(36s) = -7.987329$ Hartree, values $E^{\text{HF}} + E^{(2)}$ are given in Hartree, values $E^{(2)}$ — are in milliHartree, bases like $36s + k * \text{areused}24_{\text{off}}$, k — cardinal numbers.

k	Basis structure	$E^{\text{HF}} + E^{(2)}$	$E^{(2)}$
	distributive basis ^(a)		
1	$36s + 1 * 24_{\text{off}}(60s)$	-8.052867	-65.538
2	$36s + 2 * 24_{\text{off}}(84s)$	-8.055916	-68.587
3	$36s + 3 * 24_{\text{off}}(108s)$	-8.057132	-69.803
4	$36s + 4 * 24_{\text{off}}(132s)$	-8.057779	-70.450
	atom-centered basis ^(b)		
	$(14s9p4d3f, 8s4p)$	—	-63.68
	$(14s9p4d, 8s4p3d)$	—	-65.24
	$(14s9p4d3f, 8s4p3d)$	—	-65.40
	MP2-R12/A — „exact“		
	$(11s8p6d5f, 9s8p6d5f)$	—	-72.76

Note. ^(a) this study, ^(b) — [25].

distributed along the circles, which planes are perpendicular to the molecular axis. Figure 1 allows to visualize such sets. The role of the parameters to be determined is played by the circles radii, ρ , and the centrations of the circle planes positions along the axis, Z_{off} . The number of functions N on the circle can be different. In our case most efficient was $N = 3$. Centrations of off-functions are conveniently defined in the cylindrical coordinate system (ρ, ϕ, Z) . Then $X = \rho \cos \phi$, $Y = \rho \sin \phi$, $Z = Z_{\text{off}}$ and $\phi = \frac{2\pi}{N}$. The structure of the basis sets indicated in Tables 1, 2 means the following: for example, for the LiH molecule the basis $36s + 3 * 24_{\text{off}}$ means, that it includes 36 s -functions located on axes, and three subsets of off-functions, each containing 24 off-functions. For the hydrogen molecule the basis $2 * (15s + 4 * 12_{\text{off}} + 12_{\text{mid}}$ contains 15 s -functions located on the axis and four subsets of off-functions, each of which includes 12 off-axis s -functions. All these functions fall on one atom H. The term 12_{mid} means the number of functions at the midpoint on the axis. The total number of s -functions is marked in brackets. The calculation algorithm is as follows.

1) Calculate the HF energy, E^{HF} , using the functions distributed along the axis. The basis parameters were determined by minimizing EHF. The calculations were performed for bases of various dimensions M . Finally a basis was chosen, which provided the value of the HF energy close to the HF limit. For example, for H_2 ($R = 1.4$ bohr) the basis of $2 \cdot 15 = 30$ of s -functions gave $E^{\text{HF}}(30s) = -1.13362933$ Hartree, which is comparable to „exact“ HF-value $E_{\text{limit}}^{\text{HF}} = -1.13362957$ Hartree. For LiH the basis consisted of 36 s -functions, which provided $E^{\text{HF}}(36s) = -7.987329$ Hartree (compare with the HF limit $E_{\text{limit}}^{\text{HF}} = -7.987351$ Hartree). In this case, the addition of off-functions practically does not change E^{HF} .

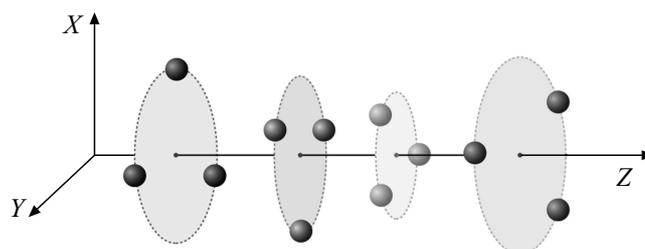


Figure 1. Distribution of off-axis features in the subset 12_{off} .

2) Add off-functions to the resulting basis and minimize $E^{(2)}$ by the parameters (ξ, Z, ρ) . In other words we apply the Hilleras variational method to determine these parameters.

The results of calculating the correlation energies H_2 and LiH in the second order of the perturbation theory are presented in the Tables 1 and 2, respectively. The calculations were carried out for the internuclear distance $R = 1.4$ bohr for H_2 and $R = 3.015$ bohr for LiH. For comparison, at the bottom of the tables we present the results obtained using popular traditional bases. To evaluate the efficiency of the model, we compare $E^{(2)}$ with the results calculated by the MP2-R12 method, which can be considered as „exact“. As you can see, our result for H_2 covers $\sim 99\%$ and for LiH more than 97% „of the exact“ value.

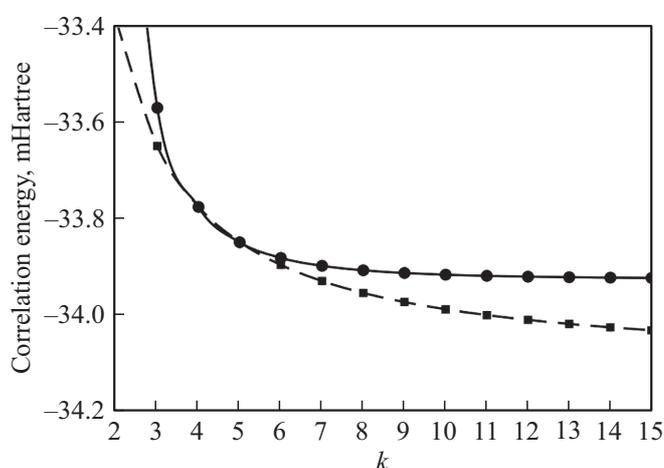
A feature of choosing sequences of distributive off-functions for constructing an extrapolation scheme is the definition of cardinal numbers. As noted, in the case of standard bases, the cardinal numbers depend on the value of the quantum number of the angular momentum l . For example, for aug-cc-pVDZ basis $k = l = 2$ and etc. In our case, the hierarchy of bases was determined by the number of subsets of off-functions, i.e. the cardinal number is equal to the number of subgroups consisting only of functions centered off the molecular axis. In the Tables 1 and 2 the cardinal number is indicated in the first column. So, for example, $k = 2$ means, that the basis contains two subsets of off-functions. Each of them consists of 12 off-functions for H_2 and 24 off-functions for LiH. It is worth noting, that the extrapolation results depend on the choice of the number of features in the subsets. We performed calculations with various types of subsets. Here are the most successful results in terms of accuracy. Note, that in case of using traditional bases, there is also arbitrariness in the choice of cardinal numbers. Obviously, different sets of cardinal numbers lead to different energy values in the full basis limit.

Below we consider MP2 energies, $E^{(2)}$, in the limit of a complete basis set. obtained as a result of extrapolation. For H_2 the basis sets with cardinal numbers 3, 4, 5 were used, and for LiH $k = 2, 3, 4$. Table 3 shows the correlation energies and optimal parameters for three extrapolation models named A, B, and C, which are defined by equations (1)–(3), respectively. As can be seen, for distributive

Table 3. Extrapolation results $E^{(2)} = E_{\text{CBS}}$ (in milliHartree) in complete basis set limit and optimized parameter values of model extrapolation functions

Model ^(a)	Basis	H ₂			LiH		
		E_{CBS}	F	β	E_{CBS}	F	β
A	distributive	-33.953	1.53832	0.540122	-71.186	9.1794	0.630973
	atom-centered	-33.711	89.3495	1.316445	-65.418	165.245	2.27726
B	distributive	-33.928	9.70492	–	-70.922	30.2166	–
	atom-centered	-34.169	58.8454	–	-65.516	7.47243	–
C	distributive	-34.112	1.57459	1.11341	-72.869	7.57998	0.823755
	atom-centered	-34.302	46.7168	2.73590	-65.444	70.3731	5.31791

Note.^(a) A $\rightarrow E(k) = E_{\text{CBS}} + Fe^{-\beta k}$, B $\rightarrow E(k) = E_{\text{CBS}} + Fk^{-3}$, C $\rightarrow E(k) = E_{\text{CBS}} + Fk^{-\beta}$.

**Figure 2.** Correlation energy MP2 of the H₂ molecule as a function of the cardinal number: model B (circles), model C (squares).**Table 4.** Comparison of the correlation energies MP2 for the ground state of H₂ and LiH molecules with the results of precision calculations.

Source	Method	$E^{(2)}$	
		H ₂	LiH
This study	CBS limit, off-axis functions	-34.112	-72.869
[25]	MP2-GG1 (a)	-34.247	-72.877
[26]	MP2-R12/A	-34.23	-72.76
[26]	MP2-R12/B	-34.17	-72.16

Note.^(a) Gaussian geminals were used as basis functions.

bases, the power model C is the most successful, i.e., $e. E(k) = E_{\text{CBS}} + Fk^{-\beta}$, while for atom-centered bases this is a two-parameter model (model B). The presented parameter values allow us to reproduce our results. As an illustrative example of convergence, we present in Fig. 2 the values of the correlation energy of the H₂ molecule as a function

of cardinal numbers for models B and C. As can be seen, the models B and C show close behavior in the range of cardinal numbers 4 and 5. This becomes clear, when we consider, that the specified cardinal numbers are common to „fitting“ of both extrapolation models.

Table 4 shows our best values of the MP2 correlation energy (model C) in the complete basis limit compared to „exact values“ obtained by methods using the MP2-R12 method. The analysis shows, that our results are in good agreement with the results of precision calculations.

In conclusion, we note, that a further improvement of the results can be achieved by combining atom-centered and distributive bases centered outside the molecular axis. For example, for more complex molecules, higher quantum number orbitals (including *d* functions) are built as atom-centered ones, while distributive functions are useful for describing the outer part of the electron cloud.

Funding

The work was done within the personal interest of the authors.

Conflict of interest

The authors have no conflict of interest.

References

- [1] E.A. Hylleraas. *Z. Phys.*, **54**, 29 (1929).
- [2] W. Cencek, J. Rychlewski. *J.Chem.Phys.*, **98**, 125 (1994).
- [3] K.A. Peterson, M.K. Kercharwani, J.M.L. Martin. *Mol. Phys.*, **113**, 1551 (2015).
- [4] W. Klopper, F.R. Manby, S. Ten-No, E.F. Valeev. *Int. Rev. Phys. Chem.*, **25**, 427 (2006).
- [5] R. Bukowski, B. Jeziorski, S. Rubak, K. Szalewicz. *J. Chem. Phys.*, **102**, 888 (1995).
- [6] L. Kong, F.A. Bischoff, E.F. Valeev. *Chem. Rev.*, **112**, 75 (2012).
- [7] M.R. Nyden, G.A. Petersson. *J. Chem. Phys.*, **75**, 1843 (1981).

- [8] D. Feller. *J. Chem. Phys.*, **98**, 7059 (1993).
- [9] W. Klopper. *J. Chem. Phys.*, **102**, 6168 (1995).
- [10] D. Feller, K.A. Peterson, J.G. Hill. *J. Chem. Phys.*, **135**, 044102 (2011).
- [11] A. Halkier, T. Helgaker, P. Jørgensen, W. Klopper, H. Koch, J. Olsen, A.K. Wilson. *Chem. Phys. Lett.*, **286**, 243 (1998).
- [12] T. Helgaker, W. Klopper, H. Koch, J. Noga. *J. Chem. Phys.*, **106**, 9639 (1997).
- [13] Jan M.L. Martin. *Chem. Phys. Lett.*, **259**, 669 (1996).
- [14] A. Halkier, T. Helgaker, P. Jørgensen, W. Klopper, J. Olsen. *Chem. Phys. Lett.*, **302**, 437 (1999).
- [15] N.N. Pansini, A.C. Neto, A.J.C. Varandas. *Theor. Chem. Acc.*, **135**, 261 (2016).
- [16] D. Bakowies. *J. Chem. Phys.*, **127**, 084105 (2007).
- [17] J. Seino, H. Nakai. *J. Comp. Chem.*, **37**, 2304 (2016).
- [18] T.H. Dunning Jr. *J. Chem. Phys.*, **90**, 1007 (1989).
- [19] D.E. Woon, T.H. Dunning Jr. *J. Chem. Phys.*, **100**, 2975 (1994).
- [20] A.K. Wilson, T. van Mourik, T.H. Dunning Jr. *J. Mol. Struct. (Theochem)*, **388**, 339 (1996).
- [21] V. Vasilyev. *Comp. Theor. Chem.*, **1115**, 1 (2017).
- [22] V.N. Glushkov, J. Kobus, S. Wilson. *J. Phys. B*, **41**, 205102 (2008).
- [23] V.N. Glushkov, S. Wilson. *Int. J. Quant. Chem.*, **107**, 2632 (2007).
- [24] V.N. Glushkov, S. Wilson. *Mol. Phys.*, **12**, 3092 (2014).
- [25] P. Dahle, T. Helgaker, D. Jonsson, P.R. Taylor. *Phys. Chem. Chem. Phys.*, **93**, 112 (2007).
- [26] W. Klopper, W. Kutzelnigg. *J. Chem. Phys.*, **94**, 2020 (1991).