
Spanning the Hilbert Space With an Even Tempered Gaussian Basis Set

IRA CHERKES,¹ SHACHAR KLAIMAN,² NIMROD MOISEYEV^{1,2}

¹Russell-Berrie Nanotechnology Institute, Technion–Israel Institute of Technology, Haifa 32000, Israel

²Shulich Department of Chemistry and Minerva Center for Nonlinear Physics of Complex Systems, Technion–Israel Institute of Technology, Haifa 32000, Israel

Received 20 November 2008; accepted 5 January 2009

Published online 30 April 2009 in Wiley InterScience (www.interscience.wiley.com).

DOI 10.1002/qua.22090

ABSTRACT: In this work we show how and why the well-known Gaussian even-tempered basis spans the Hilbert space evenly. This basis is thought to be chosen based mainly on empirical results. However as shown here it can also be deduced based on rigorous numerical ground using a Gram-Schmidt orthogonalization procedure.

© 2009 Wiley Periodicals, Inc. Int J Quantum Chem 109: 2996–3002, 2009

Key words: even-tempered; gaussian basis; Gram-Schmidt orthogonalization

Introduction

Since being proposed by Boys in 1950 [1], Gaussian basis sets have been widely used in electronic structure calculations. Gaussian basis sets made *ab-initio* calculations feasible by allowing analytical evaluations of certain integrals necessary for the calculations. Still a many electron Self Consistent field (SCF) computation scales as N^4 where N is the number of basis functions. This enforces a strict limit on the number of basis functions that can be used in the computation. The use of Gaussians as

basis functions can also lead to a (quasi-)linear dependence between the different basis functions. The linear dependence is a consequence of the use of a nonorthogonal basis set. It is imperative that one avoids such a situation in any calculation since it makes the calculation numerically unstable.

The above two reasons led to a nonlinear exponent optimization of the Gaussian functions. An optimization procedure in which all the exponents of the primitive functions are optimized is extremely time consuming and can be done in practice only to a limited number of primitive functions at a time. To avoid full optimization many methods were devised to fit some functional for the Gaussian exponents that will give results close to fully optimized basis while at the same time will have much less optimization parameters. The first suggested functional was named even-tempered [2, 3]

Correspondence to: I. Cherkas; e-mail: ira.cherkes@gmail.com

Contract grant sponsor: ISF.

Contract grant number: 96/07.

Contract grant sponsors: Technion, Israel ministry of science, culture and sports.

with only two optimization parameters. Other functionals with more optimization parameters where proposed, among them are: well-tempered sets [4], the Klobukowskis' expansion of exponents as a polynomial [5], and the Petersson and Zhong expansion using Legendre polynomials [6]. All of these methods try and match some functional form to the exponents, hoping that given sufficient parameters one can get closer to the SCF limit while avoiding the need to fully optimize all the exponents. Despite the additional parameters of other functionals, the even-tempered basis is still considered to be the one that spans the Hilbert space most evenly.

The need to avoid exponent optimization for each new problem has forwarded the use of contracted basis functions which are formed from the superposition of several optimized primitive gaussian functions. The optimization of the primitive function's exponents is usually done to fit a certain property of the studied problem, e.g., minimum energy, asymptotic form of the wavefunction, etc.. The optimization is therefore problem dependent and today one can find such contracted functions for most atoms in the periodic table. However, these contracted functions are mostly optimized such that they minimize the ground state energy and thus are usually good for describing localized states. However, if one wishes to describe delocalized states, in particular resonance states, one has to add to the contracted basis functions additional delocalized Gaussian functions. (see [7–12]). If one wishes to avoid exponents optimization and to avoid the need for information on the specific problem the even-tempered basis set will often be the best choice.

Reeves and Harrison [3] were the first to suggest the even-tempered Gaussian basis set based on intuitive reasoning. Later this basis was proposed again by Ruedenberg et al.[2] This time it was based on the empirical observation that when full optimization is performed the optimized exponents obtained follow an almost straight line on the logarithmic scale. The exponents in the even-tempered set also form such a straight line, reducing the number of parameters to optimize to two. For example, consider an even tempered set made of N s-type Gaussian functions:

$$g_k(r) = \left(\frac{2\alpha_k}{\pi}\right)^{\frac{3}{4}} e^{-\alpha_k r^2}, \quad (1)$$

where $k = 1, 2, \dots, N$ and the exponents are chosen as:

$$\alpha_k = \alpha_1 \varepsilon^{k-1}. \quad (2)$$

Here the two parameters that should be optimized are ε and α_1 . The functional in Eq. (2) is applied also for Gaussians of other angular momentum $l = 0, 1, 2$ (corresponding to s-, p-, d-orbitals). Recently, the even-tempered Gaussian set has been shown to be asymptotically complete [13] when taking α_1 and ε to be functions of N as proposed by Schmidt and Ruedenberg [14]. However, when ε remains constant it is possible to show that the even-tempered basis is incomplete [15].

The overlap between two Gaussian functions with the same angular momentum depends only on the ratio between the exponents:

$$\langle g_k | g_{k'} \rangle = \left(\frac{2\sqrt{\alpha_k/\alpha_{k'}}}{1 + \alpha_k/\alpha_{k'}} \right)^{\frac{2l+3}{2}} = \cos \vartheta_{k,k'}, \quad (3)$$

where $\vartheta_{k,k'}$ is the angle between the functions. In the even-tempered basis the ratio between two adjacent exponents is the constant ε . This property was noted by Reeves and Harrison and made them intuitively suggest that such constant overlap will lead to even coverage of the Hilbert space. The common explanation is that the overlap can be considered as the cosine of the angle between the functions. In the even-tempered basis this angle is constant, i.e. $\cos \vartheta_{k,k+1}$ is independent of k . If one would like to decrease the overlap between adjacent functions this would necessarily increase the overlap with a different function. This point will be elaborated later. Thus, the even tempered sequence of Gaussian exponents gives the largest value for the minimum angle between adjacent functions. This results in an even coverage of the space. However in what sense the coverage is even or why the last statements is true beside its intuitive nature, has never been demonstrated.

In this work we show how and why the even-tempered basis spans the Hilbert space evenly, taking the explanation given above one step further. We begin with a general, problem independent method of choosing a Gaussian set in a given interval and then use the results to prove that the most even representation of the space is obtained when using the even-tempered basis. Our initial analysis is problem independent and therefore we concentrate on the overlap of the functions. Afterwards we

extend the method and examine also the question of exponent selection for Gaussian basis sets for a more general metric. This gives further insight into the extensively used even-tempered basis set.

Suppose one wishes to construct a basis set composed of N one-dimensional s-type Gaussian functions: $g_k(x) = (2\alpha_k/\pi)^{1/4}e^{-\alpha_k x^2}$, whose exponents are to be chosen in the interval $\alpha_{k \geq 3} \in [\alpha_1, \alpha_2]$. For simplicity, we shall restrict ourselves to s-type Gaussian functions while the extension to p- and d-type functions is straightforward. Since the overlap integrals treated here are separable in cartesian coordinates it is sufficient to carry out our analysis in the one-dimensional case. In principle the value of α_1 should be equal to zero in order to describe the long range exponential decay, and $\alpha_2 = \infty$ in order to describe the cusp at the nuclei [16]. Of course in real calculations in order to prevent the overlap matrix from becoming singular due to the limited accuracy of the numerical calculations, α_1 should be a small number (but not zero value) and α_2 should get very large value but not infinitely large. The questions how small α_1 and how large α_2 should be are out of the scope of the present study.

Let us start with the first nontrivial case in which one wishes to use only three Gaussian functions, i.e., $N = 3$. Two exponents are immediately chosen to be α_1 and α_2 , but what of the third exponent? What is the best choice for the third exponent so that the resulting Gaussian functions will best represent the space between $[\alpha_1, \alpha_2]$?

Here we propose to use the Gram-Schmidt orthogonalization (GSO) procedure to determine the third exponent. The GSO tells us how to construct three orthogonal functions from our three Gaussian functions. Once the first two Gaussian functions have been orthogonalized, the third orthogonal function needs to be found. Let us denote the orthogonal set by χ_k , where $k = 1, \dots, N$. The orthogonal functions can be found from the following prescription

$$\begin{aligned}\chi_1(x) &= g_1(x) \\ \chi_k(x) &= g_k(x) - \sum_{p=1}^{k-1} \frac{\langle \chi_p | g_k \rangle}{N_p} \chi_p(x).\end{aligned}\quad (4)$$

Note that with exception to χ_1 , the new orthogonal functions are not normalized and their norm is given by:

$$N_k = \langle \chi_k | \chi_k \rangle.\quad (5)$$

As can be seen from Eq. (4) the major benefit in the GSO process is that it allows for a step by step selection of the Gaussian basis functions and the construction of the corresponding orthogonal basis functions.

We choose the next exponent such that it maximizes the norm of the newly constructed orthogonal function. This criterion allows the addition of the function that is the least linearly dependent on all the orthogonal functions and therefore also on all the previously chosen Gaussian functions in the set. This criterion is also equivalent to adding of a function that contributes the most to describing a missing part of the space, as will be explained below. The resulting norm as a function of α_3 is depicted in Figure 1(a). As can be seen in Figure 1(a), the maximal norm occurs just in the middle of the logarithmic α axis, i.e., at $\alpha_3 = 10^{(\log_{10}\alpha_1 + \log_{10}\alpha_2)/2} = \sqrt{\alpha_2\alpha_1}$ hereby dividing the logarithmic α axis into two.

One can now continue this process and add more Gaussian functions into the nonorthogonal basis. With each Gaussian added, the Gaussian's exponent is optimized by choosing the exponent which gives the maximal norm of the newly constructed orthogonal function.

The norm can be calculated from the previous orthogonalized functions using the following formula:

$$N_k = \langle \chi_k | \chi_k \rangle = 1 - \sum_{p=1}^{k-1} \frac{|\langle \chi_p | g_k \rangle|^2}{N_p}.\quad (6)$$

Or as proved in [17] it can also be calculated using:

$$N_k = \frac{\det S_k}{\det S_{k-1}},\quad (7)$$

where S_k is the overlap matrix of the k Gaussian functions and S_{k-1} is the overlap matrix of all Gaussians except the one for which we calculate the norm. This allows one to skip the actual orthogonalization of the functions.

Figure 1 shows the results for the next few Gaussian functions chosen using the maximal norm criterion. Whenever there are several maxima with the same height one is chosen arbitrarily while in all other cases the global maximum is selected. Two important features can be detected in Figure 1. First, it is interesting that the α for which the maximal norm is obtained seems to always divide the axis

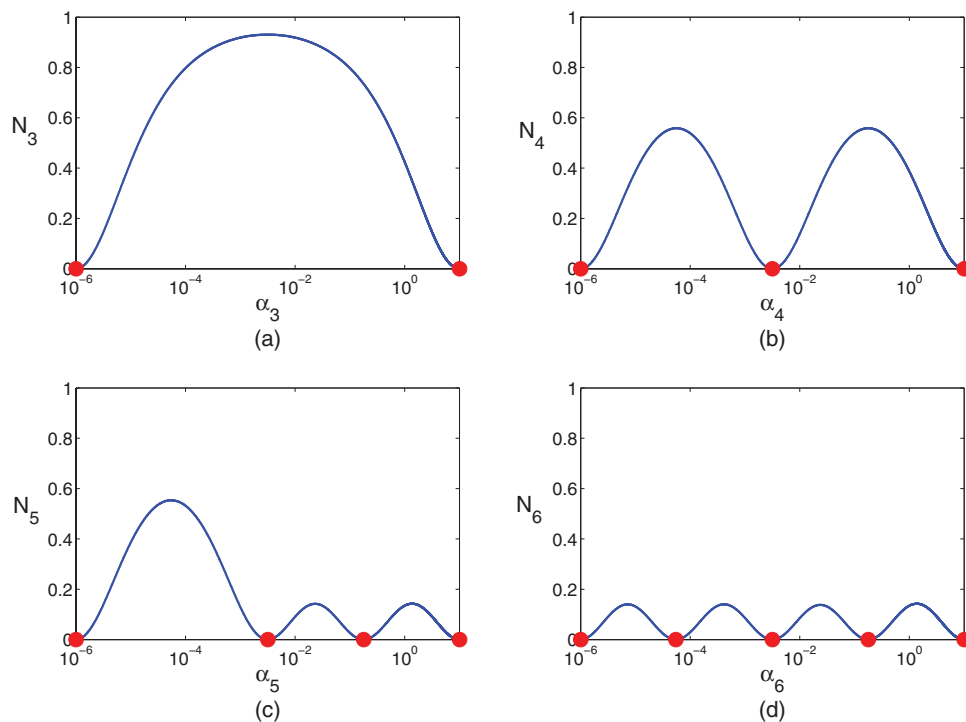


FIGURE 1. The norm- N_k , of the k th orthogonal (via Gram-Schmidt) basis function as a function of the Gaussian exponent α_k lying in the interval $[\alpha_1 = 10^{-6}, \alpha_2 = 10]$. The red dots represent exponents that are already in the basis set. Note that the maximal norm appears in the centers of segments between existing exponents on the logarithmic scale. Also the maximal value of the norm decreases significantly when the logarithmic segment becomes smaller. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

between two previous exponents into two on the logarithmic scale. The second important feature is that the maximal norm always appears in the largest interval. Note that the value of the norm at the maximum drops as more Gaussian functions are chosen and the intervals between exponents become smaller.

Figure 2 depicts the exponents chosen by the GSO method described above for 14 Gaussian functions in the interval: $[\alpha_1 = 10^{-6}, \alpha_2 = 10]$. The results clearly show that the maximal norm criteria partitions the largest interval between two previously chosen exponents into equal segments on a logarithmic scale.

The choice of maximal norm is equivalent to adding the Gaussian that is least represented by the existing orthogonal set, $\{\chi_{i1}\}^{k-1}$, and therefore is the least represented by the original Gaussian set, $\{g_{i1}\}^{k-1}$. We work in the overlap metric. The function

$$\tilde{g}_k = \sum_{p=1}^{k-1} \frac{\langle \chi_p | g_k \rangle}{N_p} \chi_p \quad (8)$$

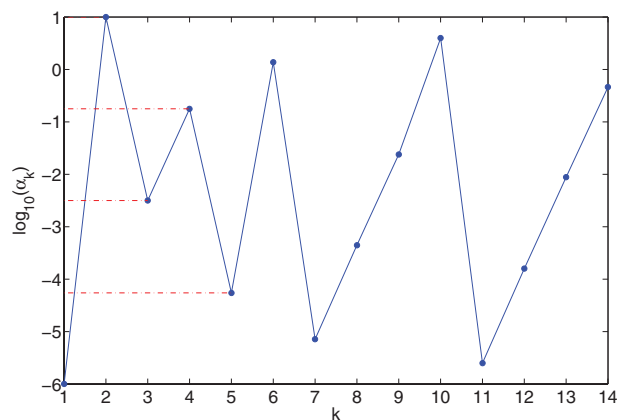


FIGURE 2. The optimal exponents chosen by the maximal norm criteria in the order of their selection. The results show that the maximal norm criteria partitions the largest interval between two previously chosen exponents into equal segments on a logarithmic scale. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

is an approximation to the examined Gaussian, g_k . To evaluate the error of the approximation in the overlap matrix we look at [18, 19]

$$\varepsilon_k = \langle g_k | g_k \rangle - \langle \tilde{g}_k | \tilde{g}_k \rangle = \langle g_k - \tilde{g}_k | g_k - \tilde{g}_k \rangle = N_k \quad (9)$$

as can be shown by simple substitution of Eq. (4). Therefore our method searches for the Gaussian that has the largest error in its representation using the already existing functions in the set. The criterion suggested above for selecting the k 'th Gaussian is therefore equivalent to looking for a Gaussian that has a dominant contribution to the description of the missing part of the space.

Keep in mind that the norm is inversely-proportional to how well this function can be spanned by the existing basis, and that the maximal norm value is obtained for the exponent in the middle of the largest segment on a logarithmic scale between two existing exponents in the bases. We now use these results to prove that the even-tempered basis spans the Hilbert space evenly.

We now turn and examine an already existing basis. We do not try to add functions to it but simply to determine its properties using a norm criterion. Eq. (7) is used in order to calculate the norm.

Suppose one wishes to choose a basis of N functions such that it will best represent any given function. The exponents are allowed to be chosen only in the range $[\alpha_1, \alpha_2]$, however there are many ways to choose exponents in this region. Which basis of all those spans the space better? We ask if we would have added another function to that basis how much would it contribute. To answer this we use the norm criterion. The higher the norm is the less this function is represented. For a basis to effectively span all Gaussians in this range we look for the highest norm, as a function of the missing exponent, to be as low as possible. Let us examine the results presented in Figure 3. Two different choices of four exponent Gaussian basis sets in the region $[\alpha_1 = 10^{-6}, \alpha_2 = 10]$ have been selected. The sets are represented by red dots. The even-tempered with even intervals between exponents is described in (a) and exponents selected using GSO procedure are described in (b). In Figure 3 we show the variation of the norm with an additional exponent $\alpha \in [\alpha_1, \alpha_2]$ that is not in the basis. Now we address the question of which basis set spans the Hilbert space better. It would be desirable to see in which set there is a Gaussian that is poorly repre-

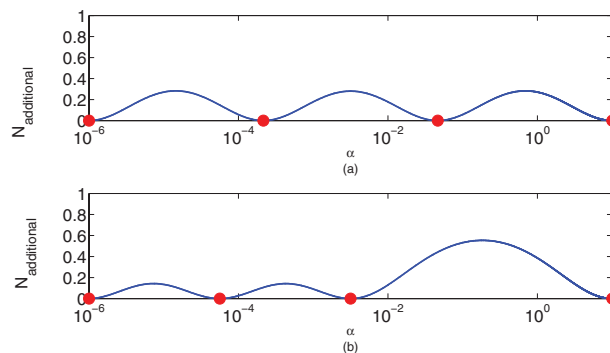


FIGURE 3. The norm of all Gaussians with exponents in the range $[\alpha_1 = 10^{-6}, \alpha_2 = 10]$ for two different four function basis sets. The exponents of the basis sets are represented by red dots. (a) Even-tempered basis set and (b) exponents selected using GSO procedure. The even-tempered basis spans the space more evenly since the maximal N_k of all possible missing Gaussians is lower. Meaning there is no one particular Gaussian in the range that is poorly represented with respect to other. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

sented compared to the others. Such question can be answered by looking at the norm of all functions in the range $[\alpha_1, \alpha_2]$. Figure 3 shows that the maximal norm is smaller in the even-tempered case than in the other set, rendering the representation of the Hilbert space by the even-tempered basis more even. This conclusion holds for any choice of exponents in Figure 3(b).

In general for better representation of the space it is advisable that the norm of a function that can be added to the N base functions will be as small as possible. Since the norm is proportional to the size of the interval between existing exponents on the logarithmic scale, it is best to avoid large segments. The way to do so is to divide the logarithmic space to even segments, meaning to use an even-tempered basis set. Thus the even-tempered spans the space evenly in the sense that there is no one particular Gaussian in the region that is extremely poorly represented with respect to the other.

We now consider a generalization to our approach. We started our discussion as a problem independent analysis and as such it was done in the overlap metric only. However, what would happen if we try to apply it to a different metric? We can redefine the inner product from $\langle g_k | g_k \rangle$ to $\langle g_k | \hat{O} | g_k \rangle$, where \hat{O} is a positive definite operator. For example we examine here the Coulombic metric which is of large interest in molecular computations [18, 20]. We consider the three

dimensional radial potential $1/r$ for s-type radial Gaussian functions, however, the result are the same for p-type and d-type as well. The inner product is redefined as

$$(g_k|g_{k'}) = \left\langle g_k \left| \frac{1}{r} \right| g_{k'} \right\rangle = \int_0^\infty g_{k'}^- g_k r^2 dr = \int_0^\infty g_k r g_{k'} dr. \quad (10)$$

We repeat the same GSO process as before with all Dirac brackets in Eqs. (4)–(6) replaced by round brackets. In each step we wish to choose the Gaussian that is least represented in the Coulombic metric by the already existing functions in the basis. Hence the criterion for selecting the exponents remains the maximal norm, N_k . Note that the orthogonal functions are still a linear combination of Gaussian functions, however now the coefficients are different than in the overlap metric case, giving different weight to the Gaussian functions. Note also that $\chi_{p'}$ are orthogonal in the Coulombic metric, i.e. $\langle \chi_{p'} | 1/r | \chi_{p'} \rangle = N_p \delta_{p,p'}$.

The resulting norm as a function of the selected exponent for the first few steps is depicted in Figure 4. As can be seen in Figures 1 and 4, the norm, N_k , as a function of the exponent, α , is not the same when calculated using the Coulomb and overlap metrics. Still two common features can be identified. The maximal norm obtained partitions the largest interval between two previously chosen exponents into equal segments on a logarithmic scale. Furthermore, just as with the norm calculated in the overlap metric, the value of the maximal norm decreases with the length of the interval partitioned. These results indicate that the basis that span the space most evenly in the Coulombic metric is once more the well known even-tempered basis. We further checked the effect of different metrics of the form $\langle g_k(r) | r^m | g_{k'}(r) \rangle = \int_0^\infty r^m g_k g_{k'} dr$. Up to $m = 10$ the resulting norm retained the properties mentioned above. This suggests that an even-tempered selection of the exponents spans the space evenly for any metric of the form $\langle g_k(r) | r^m | g_{k'}(r) \rangle$.

To conclude, basis selection using Gram–Schmidt sheds light on the behavior of the long known even-tempered basis. It allows to prove that indeed an even-tempered set spans the space evenly and shows exactly the meaning of even expansion. Moreover we saw here that even-tempered basis spans the space in the most even way

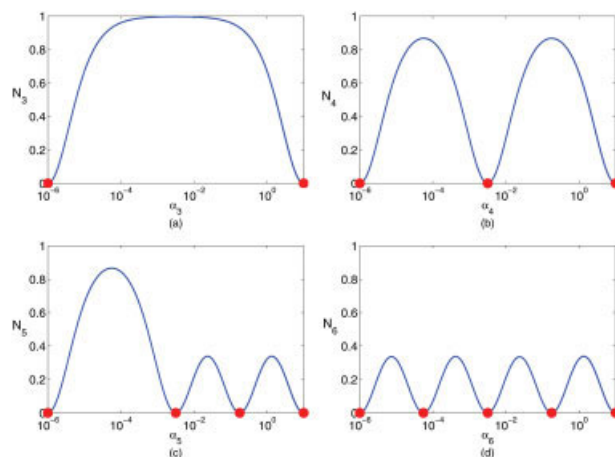


FIGURE 4. The norm— N_k in the Coulombic metric, of the k th orthogonal (via Gram–Schmidt) basis function as a function of the Gaussian exponent α_k lying in the interval $[\alpha_1 = 10^{-6}, \alpha_2 = 10]$. The red dots represent exponents that are already in the basis set. Note that the same properties that were valid in the overlap metric appears also here. Maximal norm appears in the centers of segments between existing exponents on the logarithmic scale. Also the maximal value of the norm decreases significantly when the logarithmic segment becomes smaller. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

also in other metrics in particular in the Coulomb metric. The fact that even-tempered Gaussian basis spans reasonably well a variety of functions explains the empirical observation of Ruedenberg, Raffennetti and Bardo that fully optimized exponents for many different problems form almost straight line on the logarithmic scale.

ACKNOWLEDGMENTS

The authors thank Dr. M. Sindelka and Dr. Y. Sajeev for very helpful discussions.

References

- Boys, S. F. Proc R Soc Ser A 1950, 200, 542.
- Ruedenberg, K.; Raffennetti, R. C.; Bardo, R. D. In Energy, Structure and Reactivity, Proceedings of the 1972 Boulder Conference on Theoretical Chemistry; Wiley: New York, 1973.
- Reeves, C. M.; Harrison, M. C. J Chem Phys 1963, 39, 11.

4. Huzinaga, S.; Klobukowski, M. *Chem Phys Lett* 1985, 120, 509.
5. Klobukowski, M. *Chem Phys Lett* 1993, 214, 166.
6. Petersson, G. A.; Zhong, S.; Montgomery, J. A., Jr.; Frisch, M. J. *J Chem Phys* 2003, 118, 1101.
7. Sajeev, Y.; Sindelka, M.; Moiseyev, N. *J Chem Phys* 2008, 128, 061101.
8. Sajeev, Y.; Moiseyev, N. *Phys Rev B* 2008, 78, 075316.
9. Cabrera-Trujillo, R.; Öhrn, Y.; Deumens, E.; Sabin, J. R. *Phys Rev A* 2004, 70, 042705.
10. Cabrera-Trujillo, R.; Deumens, E.; Öhrn, Y.; Sabin, J. R. *Nuc Instrum Methods Phys Res Sect B* 2000, 168, 484.
11. Kryzhevoi, N. V.; Cederbaum, L. S. *J Chem Phys* 2005, 123, 154308.
12. Fronzoni, G.; Stener, M.; Decleva, P.; Wang, F.; Ziegler, T.; van Lenthe, E.; Baerends, E. J. *Chem Phys Lett* 2005, 416, 56.
13. Kryachko, E. S.; Wilson, S. *Int J Quantum Chem* 2003, 93, 112.
14. Schmidt, M. W.; Ruedenberg, K. *J Chem Phys* 1979, 71, 3951.
15. Davidson, E. R.; Feller, D. *Chem Rev* 1986, 86, 861.
16. Kutzelnigg, W. *Int J Quantum Chem* 1994, 51, 447.
17. Morse, P. M.; Feshbach, H. *Methods of Theoretical Physics*; Vol. 2, McGraw-Hill: New York, 1953; p 928.
18. Reine, S.; Tellgren, E.; Krapp, A.; Kjaergaard, T.; Helgaker, T.; Jansik, B.; Hoest, S. S.; Salek, P. *J Chem Phys* 2008, 129, 104101.
19. Baerends, E. J.; Ellis, D. E.; Ros, P. *Chem Phys* 1973, 2, 41.
20. Varga, S. *Int J Quantum Chem* 2008, 108, 1518.