



Perspectives on Basis Sets Beautiful: Seasonal Plantings of Diffuse Basis Functions

Ewa Papajak, Jingjing Zheng, Xuefei Xu, Hannah R. Leverentz, and Donald G. Truhlar*

Department of Chemistry, University of Minnesota, Minneapolis, Minnesota 55455-0431, United States

S Supporting Information

ABSTRACT: We present a perspective on the use of diffuse basis functions for electronic structure calculations by density functional theory and wave function theory. We especially emphasize minimally augmented basis sets and calendar basis sets. We base our conclusions on our previous experience with commonly computed quantities, such as bond energies, barrier heights, electron affinities, noncovalent (van der Waals and hydrogen bond) interaction energies, and ionization potentials, on Stephens et al.'s results for optical rotation and on our own new calculations (presented here) of polarizabilities and of potential energy curves of van der Waals complexes. We emphasize the benefits of partial augmentation of the higher-zeta basis sets in preference to full augmentation at a lower ζ level. Benefits and limitations of the use of fully, partially, and minimally augmented basis sets are reviewed for different electronic structure methods and molecular properties. We have found that minimal augmentation is almost always enough for density functional theory (DFT) when applied to ionization potentials, electron affinities, atomization energies, barrier heights, and hydrogen-bond energies. For electric dipole polarizabilities, we find that augmentation beyond minimal has an average effect of 8% at the polarized triple- ζ level and 5% at the polarized quadruple- ζ level. The effects are larger for potential energy curves of van der Waals complexes. The effects are also larger for wave function theory (WFT). Even for WFT though, full augmentation is not needed for most purposes, and a level of augmentation between minimal and full is optimal for most problems. The calendar basis sets named after the months provide a convergent sequence of partially augmented basis sets that can be used for such calculations. The jun-cc-pV(T+d)Z basis set is very useful for MP2-F12 calculations of barrier heights and hydrogen bond strengths.

INTRODUCTION

It has been realized for a long time^{1,2} that basis sets for electronic structure calculations of many molecular properties must contain diffuse basis functions for quantitative accuracy. It has also been shown many times that diffuse functions have a relatively small effect for a number of molecular properties (for example, for ionization potentials and stationary point geometries of many molecules). It is best to avoid using diffuse basis functions when they are unnecessary or to avoid using more than the required number when they are necessary because adding diffuse functions to a basis set not only increases the cost of the calculation but can cause problems with SCF convergence and can increase basis set superposition error (BSSE). However, systematic studies of how to include the diffuse space efficiently that are based on large sets of data for various molecular properties and barrier heights are few and far between.

We have recently conducted a series of such studies including barrier heights, electron affinities, ionization potentials, atomization energies, and hydrogen bond energies.^{3–6} We conclude that much current practice includes more diffuse functions than are needed. Often, better accuracy could be achieved if the additional cost were invested in higher- ζ basis set or more polarization functions. Our conclusions on how to achieve higher accuracy on a per-cost basis can be widely useful for practical electronic structure calculations, and so we summarize them in this article. In addition to considering our own results, we comment briefly on the optical rotation studies of Stephens, Sadlej, and co-workers.^{7–9}

In addition to reviewing and integrating previously published work, this paper presents new calculations of polarizability and potential energy curves, which provide interesting challenges for efficient basis set selection.

MINIMALLY AUGMENTED BASIS SETS

In the correlation consistent basis sets, Dunning and co-workers¹⁰ defined “aug-” (which denotes “augmented”) to mean adding one diffuse basis function to every atom for every angular momentum of basis functions already present on that atom. We call such basis sets fully augmented. For example, the cc-pVTZ basis set for methane has s, p, d, and f functions on C and s, p, and d functions on H, so aug-cc-pVTZ adds diffuse s, p, d, and f subshells to C and diffuse s, p, and d subshells to H. Dunning and co-workers individually optimized the exponential parameters for each aug- basis set.

In contrast, the earlier “plus” basis sets originally systematized by Pople and co-workers¹¹ added only diffuse s and p subshells to nonhydrogenic functions and no diffuse functions to hydrogen atoms. We call this minimal augmentation. In Pople-type basis sets, one uses the same exponential parameters for diffuse functions in any basis set for a given atom. For example, 6-31+G(d,p) and 6-311+G(2df,2p) use the same exponential parameters for diffuse functions on C.

We have defined three kinds of minimal augmentation, denoted by “+”, “maug-”, and “ma-”. We next explain these in turn.

Received: February 15, 2011

Published: August 05, 2011

The “+” kind of minimal augmentation can be applied to any basis set for any atom for which exponential parameters have been defined for adding diffuse functions to the 6-31G or 6-311G basis sets. We then add a plus suffix. For example, cc-pVTZ+ denotes a minimally augmented cc-pVTZ basis set with the exponential parameters for diffuse functions from ref 11. Some basis sets of the + type have other names. For example, the MG3S basis set¹² is minimally augmented, and in fact is the same as 6-311+G(3d2f,2df,2p) for H through Si but is improved¹³ for S, P, and Cl. The generalization of the + approach to basis sets other than 6-31+G and 6-311+G was introduced in ref 3.

The maug- kind of minimal augmentation can be applied to any aug- type correlation consistent basis set. One simply truncates the diffuse space to the minimal augmentation level. For example maug-cc-pVTZ retains the diffuse s and p functions on carbon with the exponential parameters optimized for the aug-case but deletes all other diffuse functions. This approach was introduced in refs 3 and 4. We note that other levels of augmentation between minimal (maug-) and full (aug-) have also been defined for correlation-consistent basis sets, and these are discussed in ref 6 and below.

The ma- kind of minimal augmentation⁵ may be used with any nondiffuse basis set. One simply creates diffuse functions for atoms heavier than He by dividing the smallest s and p exponential parameters already present by a factor of 3. The ma- type of minimal augmentation was originally proposed for the def2-basis sets,¹⁴ which are the second-generation default basis sets of the TURBOMOLE program, as developed in Karlsruhe. For example, the ma-TZVP basis set is obtained from the def2-TZVP basis set for carbon by adding diffuse functions obtained this way.⁵ For basis sets other than the Karlsruhe def2- series, one simply adds ma- as a prefix, e.g., ma-LANL2DZ.

Examples of applications of the efficient “+” and “maug-” basis sets to problems involving large molecules are available both in work by our group^{15–17} and in work by other research groups.^{18,19}

It is sometimes assumed that adding extra basis functions is a useful cautionary step, so it does not hurt to use aug- when a smaller basis set would be sufficient. However, using full augmentation raises the cost, and if it is done at the expense of not increasing the size of the valence space, it can be harmful. For example, it is often more beneficial to increase a basis set from maug-cc-pVDZ to maug-cc-pVTZ rather than increase it to aug-cc-pVDZ, even for properties that are sensitive to including diffuse functions, for example, barrier heights, hydrogen bond energies, and electron affinities. When applying density functional theory, the only exceptions to this general finding in our tests were B3LYP calculations of hydrogen bonding energies, which also showed no improvement in accuracy from augmentation of the quadruple- ζ basis set, and which showed higher accuracy of maug than aug at the triple- ζ level; these findings indicate that the error is dominated for B3LYP by the intrinsic error in the density functional, not by the lack of completeness of the basis set.

There are also other ways that adding extra diffuse basis functions could be harmful. For example, it can increase basis set superposition error. It is sometimes speculated that adding more functions to the basis set increases its completeness and should therefore reduce the basis set superposition error; however, it is especially the functions that overlap other centers that cause BSSE, so it is also possible to increase BSSE by adding diffuse functions. We have studied this and have shown³ that the increased diffuse space of the fully augmented basis sets does not

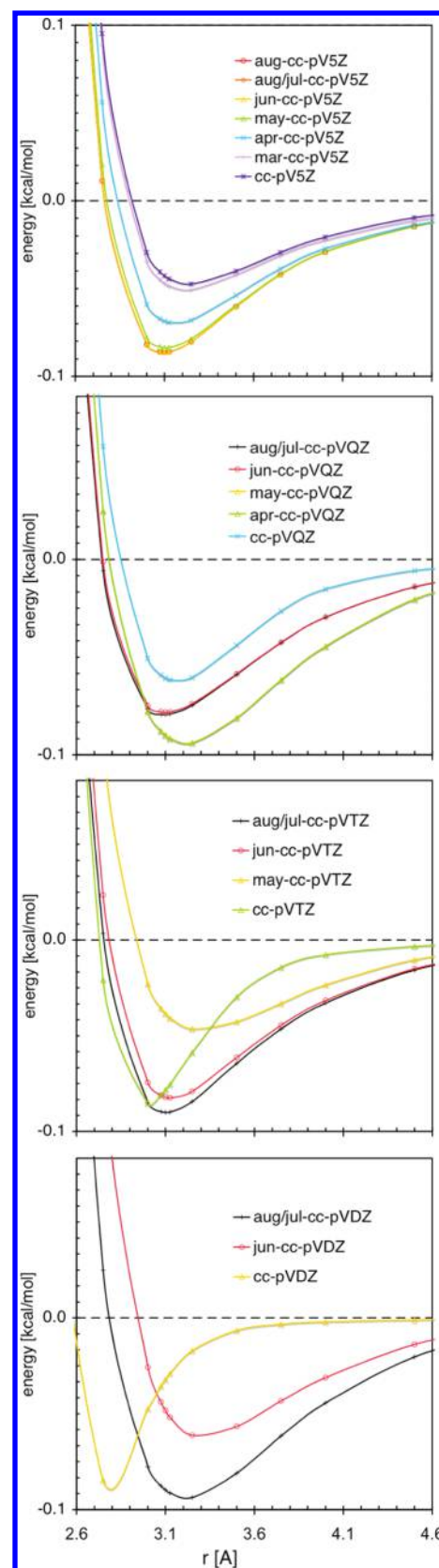


Figure 1. CCSD(T) interaction energy (in kcal/mol) for Ne_2 relative to the energy of Ne atoms at infinite separation for quintuple-, quadruple-, triple-, and double- ζ basis sets. The abscissa r is the distance between the Ne atoms. No counterpoise corrections are applied in Figures 1–6.

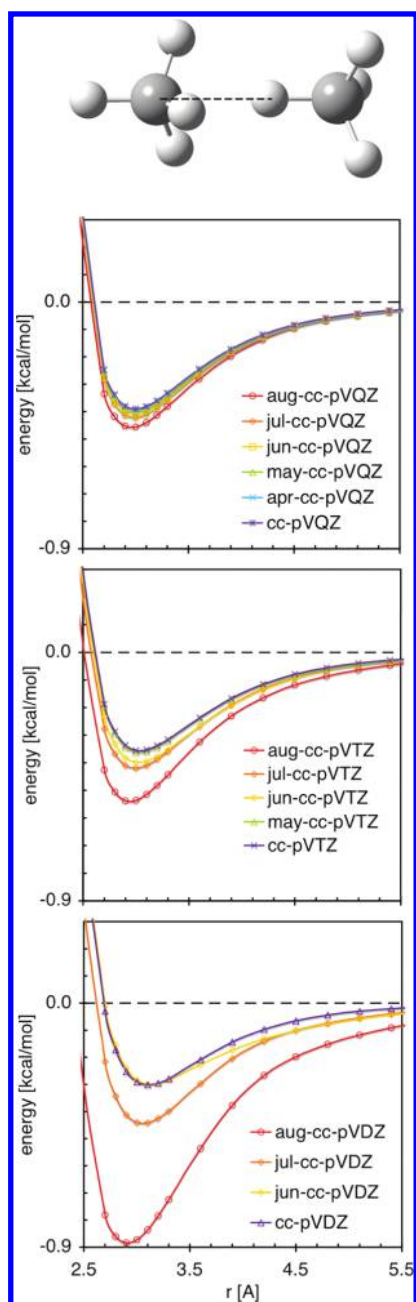


Figure 2. CCSD(T) interaction energy (in kcal/mol) for the CH_4 dimer relative to the energy of two CH_4 molecules at infinite separation for quadruple-, triple-, and double- ζ basis sets. The abscissa r is the distance between the C and H atoms, as indicated by a dashed line (---) in the figure. The structures have C_{3v} symmetry.

necessarily decrease the counterpoise correction for basis set superposition error, and in fact it can increase it. One could hypothesize that this is an indication of the limitations of counterpoise corrections, not necessarily increased BSSE. In any case, one can conclude from this that an expensive full augmentation followed by an also expensive counterpoise correction calculation is not necessarily accurate, besides being uneconomical.

DENSITY FUNCTIONAL THEORY

We have found^{3–5} that minimal augmentation is almost always enough for density functional theory (DFT) when applied to

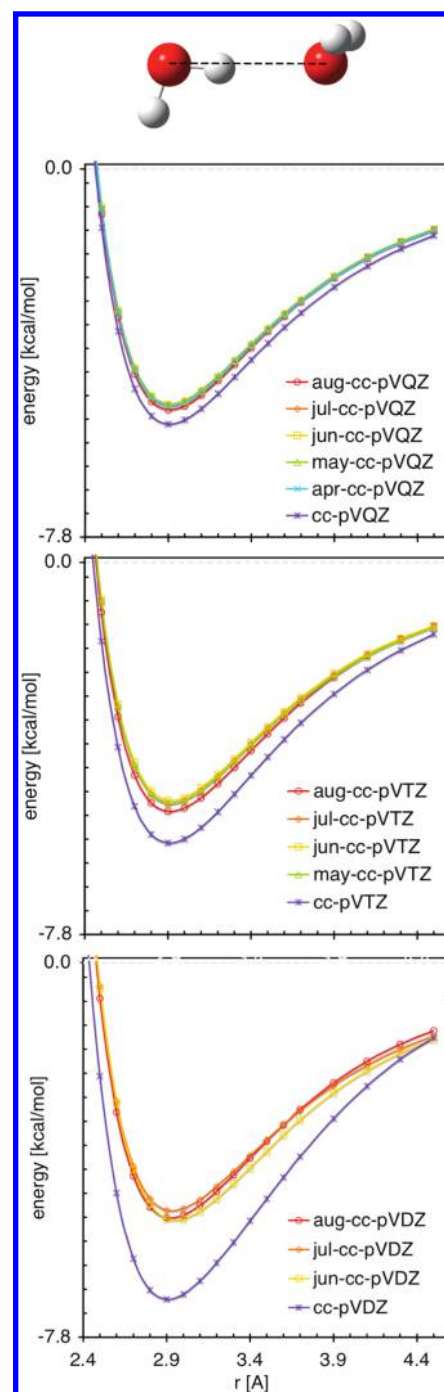


Figure 3. CCSD(T) interaction energy (in kcal/mol) for the $(\text{H}_2\text{O})_2$ relative to the energy of H_2O molecules at infinite separation for quadruple-, triple-, and double- ζ basis sets. The abscissa r is the distance between the O atoms as indicated by a dashed line (---) in the figure.

ionization potentials, electron affinities, atomization energies, barrier heights, and hydrogen-bond energies. Indeed, it has advantages over full augmentation not only in cost but also in sometimes reducing slow SCF convergence and basis set superposition error. We also note that minimal augmentation (relative to no augmentation) is often more advantageous in DFT than in wave function theory (WFT).²⁰ Minimal augmentation is especially important for barrier heights, hydrogen bonding, and electron affinities but less important for bond energies and ionization potentials.³

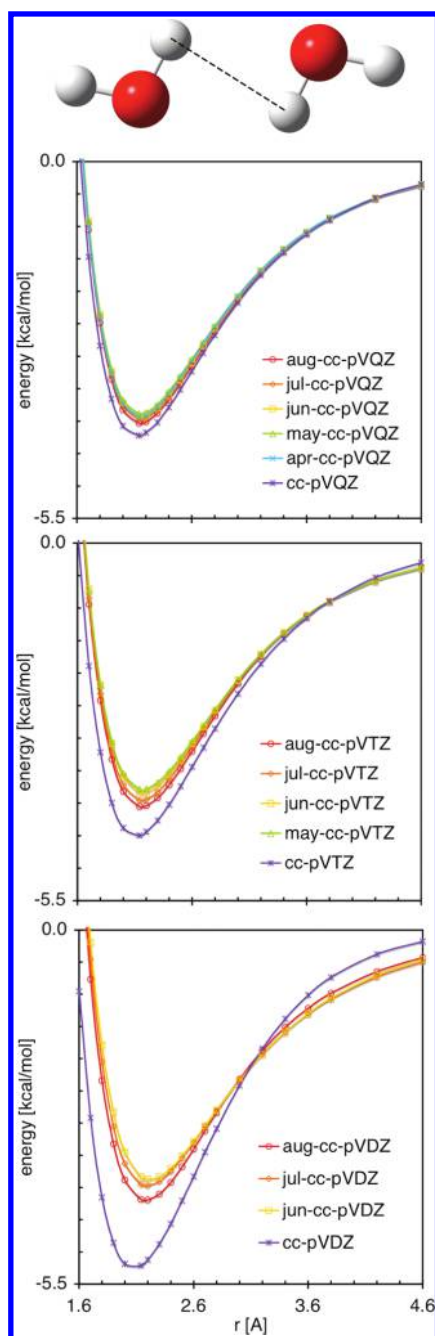


Figure 4. CCSD(T) interaction energy (in kcal/mol) for $(\text{H}_2\text{O})_2$ relative to the energy of $(\text{H}_2\text{O})_2$ molecules at infinite separation. The abscissa r is the distance between two H atoms as indicated by a dashed line (---) in the figure.

In DFT calculations, the savings from using minimal augmentation instead of full augmentation can be significant since DFT is often applied to challenging problems such as large systems or direct dynamics calculations that require a large number of electronic structure calculations on the fly. In such cases, it is very undesirable to increase the cost of the calculations by using more than the minimum number of basis functions required for chemical accuracy. Raising the cost by overaugmentation may make it necessary to adopt compromises such as a reduced number of steps in dynamics calculations, using a smaller than necessary grid, using a lower- ζ basis set, etc. The other extreme is

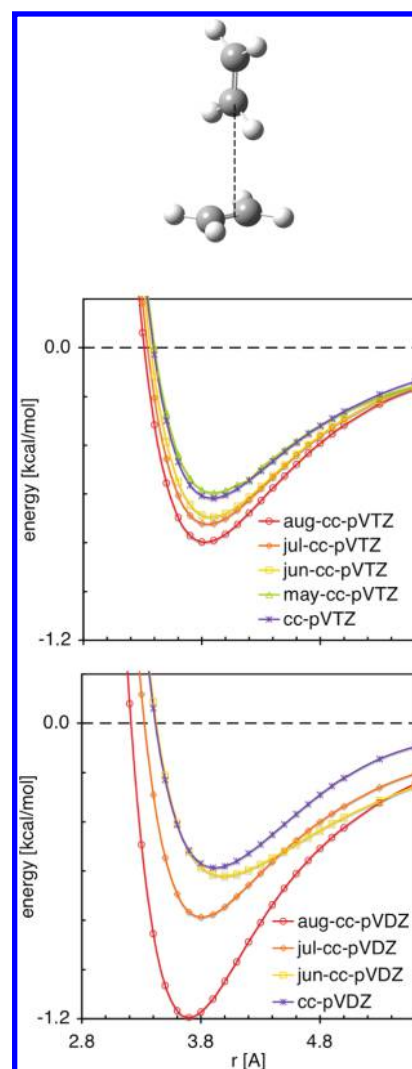


Figure 5. CCSD(T) interaction energy (in kcal/mol) for $(\text{C}_2\text{H}_4)_2$ relative to the energy of C_2H_4 molecules at infinite separation. The abscissa r is the distance indicated by a dashed line (---) in the figure.

the omission of all diffuse functions because of the cost. Our default recommendation for DFT is to use minimally augmented basis sets instead of full augmentation for properties such as barrier heights, electron affinities, polarizabilities, and noncovalent interactions; however, no augmentation is needed for ionization potentials and atomization energies.

WAVE FUNCTION THEORY AND CALENDAR BASIS SETS

Correlated wave function calculations are more sensitive to diffuse basis set expansion beyond the minimum than are density functional calculations, and therefore we recommend higher than minimal expansions of the diffuse space but still recommend less than full augmentation in most cases.

The reason for the slower convergence of correlated wave function calculations, when compared to density functional calculations, is easily understood. When including dynamic electron correlation in WFT, one uses the basis functions not only to represent the single-particle density but also to build the cusps into the many-electron wave function in the region where two

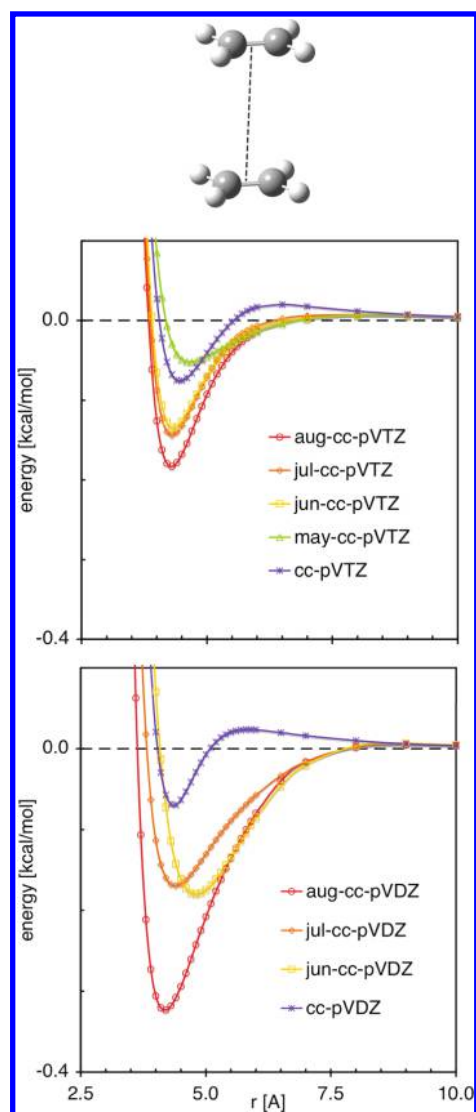


Figure 6. CCSD(T) interaction energy (in kcal/mol) for $(\text{C}_2\text{H}_4)_2$ relative to the energy of C_2H_4 molecules at infinite separation. The abscissa r is the distance indicated by a dashed line (---) in the figure.

electrons come together. The cusp is built up from a large number of configurations, extending to fairly high angular momentum for individual orbitals but with each configuration having a small coefficient, and hence second-order perturbation theory (MP2) already provides realistic predictions of these coefficients^{21,22} (the higher-order correlations are important but are less dominant in the description of the cusps). As a consequence, the basis set requirements may be examined at the MP2 level. We therefore recently studied the effect of diffuse functions on basis set convergence of MP2 calculations, and we developed a convergent sequence of partially augmented convergent correlation-consistent basis sets named after the months.⁶

In these studies, we found that even in WFT, just as in DFT, for all practical purposes one does not need diffuse functions on hydrogen atoms. Thus, our first recommendation is to delete these on all hydrogen and helium atoms; this yields basis sets we label as jul-, a naming convention that we actually introduced in a previous paper.³ (This is sometimes called aug' , but aug' has been used to denote more than one possible

modification to the aug- basis sets, so we believe that a unique notation is useful.)

The other calendar basis sets are obtained by successively deleting diffuse subshells on heavier atoms.⁶ Consider carbon. The aug- and jul- triple- ζ basis sets have diffuse s, p, d, and f subshells. We delete the f subshell to get jun-, then the d subshell to get may-, which is equivalent to maug' for triple- ζ . But at the quadruple- ζ level, carbon has diffuse s, p, d, f, and g diffuse subshells. We delete the g subshell to get jun-, the f subshell to get may-, and the d subset to get apr-, which is equivalent to maug' for quadruple- ζ . Thus, when one increases the basis from jun-cc-pVTZ to jun-cc-pVQZ, one not only increases the valence basis but also adds a diffuse f subshell; in this sense, the diffuse space is convergent when one increases the ζ level but keeps the month constant (unlike the case in going from 6-31+G(d,p) to 6-311+G(2df,2p)). Thus, for some problems, the calendar basis sets can provide a more efficient sequence of basis sets (than unaugmented, minimally augmented, or fully augmented sets) for basis set extrapolation to the complete basis set limit.

The lowest exponents in the unaugmented cc-pV α Z basis sets decrease as α increases. Therefore, even though unaugmented, the underlying cc-pV α Z basis sets become more and more diffuse. This is one of the reasons why—for the previously published data on barrier heights, electron affinities, and hydrogen bonding interactions—the need for the diffuse functions decreases as the valence space increases. For DZ, we recommended jul-, while for QZ we can reduce the basis set to jun- or even may- without a significant loss in accuracy.

In general, the goal of most electronic structure calculations is not to achieve an accuracy on the order of, for example, 0.1 kcal/mol at any cost but rather to achieve as high an accuracy possible at a reasonable cost. The actual accuracy objective is dictated by the system's size and character and by the electronic structure method itself. For example, it is unreasonable to invest an additional $\sim 30\%$ of computational time to improve accuracy by about 0.1 kcal/mol when the expected accuracy of the method is ~ 1 kcal/mol and the additional cost would have been better invested in increasing other features of the basis set or using a different level of theory. The need for this kind of approach is evidenced by the enormous number of calculations in the literature that employ double- ζ and triple- ζ basis sets, even though these calculations are clearly not at the basis set limit.

We recommend jun- basis sets as a better default option than aug- basis sets for correlated WFT calculations. On the basis of our tests, barrier heights, hydrogen bonding, electron affinities, ionization potentials, and atomization energies,⁶ it is generally more beneficial, on a performance versus cost basis, to invest additional computational resources in increasing a basis set from cc-pVDZ to jun-cc-pVTZ rather than to aug-cc-pVDZ or cc-pVTZ. The analogous statement is true for the jun-cc-pVTZ to may-cc-pVQZ transition (skipping the aug-cc-pVTZ and cc-pVQZ). We note that, averaged over a large database, aug- triple- ζ calculations involve about 27% more basis functions than jun- triple- ζ .⁶

An alternative way to converge correlated WFT calculations with respect to the basis set is the F12 method,^{24,25} in which basis functions depending explicitly on interelectronic distances are added to the basis. This is a powerful method because it eliminates much of the need to build up the electron–electron cusps by a succession of higher-angular-momentum basis functions. This method is expected to revolutionize quantum chemistry over the next several years by allowing nearly complete-basis-set calculations with much less effort than was previously required.

Table 1. HF and MP2 Polarizability [\AA^3] Values, Average Unsigned Errors [\AA^3], and Percentage Errors [%] for Basis Sets of Different Degrees of Augmentation^a

	H ₂ O		CH ₄		C ₅ H ₅ N		HO(CH ₂) ₃ SH		mean unsigned error		mean unsigned percentage error	
	HF	MP2	HF	MP2	HF	MP2	HF	MP2	HF	MP2	HF	MP2
aug-cc-pV(Q+d)Z	1.26	1.43	2.37	2.45	9.43	9.61	9.27	9.71	0.00	0.00	0.0	0.0
jul-cc-pV(Q+d)Z	1.24	1.40	2.36	2.43	9.43	9.61	9.26	9.74	0.01	0.02	0.4	0.7
jun-cc-pV(Q+d)Z	1.24	1.40	2.36	2.43	9.43	9.61	9.26	9.73	0.01	0.02	0.4	0.7
may-cc-pV(Q+d)Z	1.24	1.39	2.36	2.43	9.43	9.59	9.26	9.71	0.01	0.02	0.5	0.9
apr-cc-pV(Q+d)Z ^b	1.16	1.28	2.27	2.32	9.31	9.46	9.10	9.52	0.12	0.15	3.8	4.7
cc-pV(Q+d)Z	1.09	1.17	2.27	2.31	9.01	9.13	8.85	9.19	0.28	0.35	6.7	8.4
aug-cc-pV(T+d)Z	1.25	1.41	2.37	2.45	9.43	9.63	9.25	9.76	0.01	0.02	0.3	0.5
jul-cc-pV(T+d)Z	1.20	1.33	2.35	2.42	9.42	9.61	9.22	9.69	0.04	0.03	1.6	1.9
jun-cc-pV(T+d)Z	1.20	1.33	2.35	2.41	9.42	9.60	9.21	9.67	0.04	0.05	1.6	2.2
may-cc-pV(T+d)Z ^b	1.09	1.19	2.19	2.23	9.19	9.34	8.88	9.26	0.24	0.29	7.0	8.2
cc-pV(T+d)Z	0.98	1.03	2.18	2.20	8.59	8.63	8.41	8.66	0.54	0.65	12.2	14.6
aug-cc-pV(D+d)Z	1.20	1.36	2.36	2.44	9.40	9.63	9.17	9.67	0.05	0.03	1.5	1.4
jul-cc-pV(D+d)Z	1.09	1.20	2.28	2.33	9.36	9.57	9.02	9.44	0.15	0.17	5.2	6.1
jun-cc-pV(D+d)Z ^b	0.92	1.02	1.99	2.01	8.91	9.08	8.36	8.67	0.54	0.60	14.6	15.7
cc-pV(D+d)Z	0.74	0.76	1.91	1.91	7.77	7.83	7.53	7.63	1.09	1.27	24.2	27.2

^a The methane (CH₄) and water (H₂O) geometries were optimized at the QCISD/MG3S level, while 3-mercaptopropan-1-ol (HO(CH₂)₃SH) and pyridine (C₅H₅N) geometries were optimized at the MP2-F12/aug-cc-pV(Q+d)Z level with frozen core electrons. Note that for elements up through Mg, cc-pV(X+d)Z is the same as cc-pVXZ. ^b Same as maug-.

For F12 calculations, we recommend may- quadruple- ζ , jun- triple- ζ , and jul- double- ζ .⁶

The next section considers wave function calculations on properties not included in our previous systematic tests, and we will examine whether some of these properties are in greater need of diffuse functions than the properties considered so far.

■ WAVE FUNCTION CALCULATIONS OF VAN DER WAALS POTENTIAL CURVES, POLARIZABILITIES, RAMAN INTENSITIES, AND OPTICAL ROTATION

Accurate calculations of electrical properties of molecules often require large basis sets, including multiple polarization functions and diffuse functions.^{26–36} It is therefore interesting to study the need for higher-angular-momentum diffuse functions, and we present some calculations exploring this need in this section.

The long-range portions of the van der Waals potential curves between nonpolar molecules depend on induced low-order electric multipole moments, and therefore they provide a challenging test of the need for diffuse basis functions. Here we report new calculations designed to examine this issue. Potential energy curves were calculated for Ne₂, (CH₄)₂, (H₂O)₂, and (C₂H₄)₂ by using coupled cluster theory with single and double excitations and a quasiperturbative treatment of connected triple excitations³⁷ [CCSD(T)] using the *Gaussian 09*³⁸ program package; the results are presented in tabular form in the Supporting Information and are plotted in Figures 1–6. The geometries of the methane dimer, two orientations of the water dimer (nonplanar C_s and cyclic C_{2h}), and two orientations of the ethylene dimer (T-shaped and stacked) are shown in the figures. The potential energy curves show that diffuse functions play a larger role for van der Waals potentials than they do for the properties examined earlier in this perspective. Nevertheless, some savings are possible. For example, jun-QZ and jun-TZ agree reasonably well with aug-QZ and aug-TZ, respectively.

Table 2. HF, MP2, and MP2-F12 Finite-Field Approximation to zz Element of the Polarizability Tensor [\AA^3] of H₂O for Basis Sets of Different Degrees of Augmentation

	H ₂ O		
	HF	MP2	MP2-F12
jul-cc-pV(5+d)Z	1.25	1.40	1.41
aug-cc-pV(Q+d)Z	1.26	1.42	1.42
jul-cc-pV(Q+d)Z	1.24	1.38	1.38
jun-cc-pV(Q+d)Z	1.24	1.38	1.38
may-cc-pV(Q+d)Z	1.24	1.37	1.38
apr-cc-pV(Q+d)Z	1.14	1.24	1.29
cc-pV(Q+d)Z	1.11	1.19	1.25
aug-cc-pV(T+d)Z	1.24	1.40	1.41
jul-cc-pV(T+d)Z	1.19	1.31	1.33
jun-cc-pV(T+d)Z	1.18	1.30	1.33
may-cc-pV(T+d)Z	1.06	1.15	1.23
cc-pV(T+d)Z	1.00	1.05	1.16
aug-cc-pV(D+d)Z	1.19	1.34	1.38
jul-cc-pV(D+d)Z	1.06	1.15	1.22
jun-cc-pV(D+d)Z	0.88	0.95	1.13
cc-pV(D+d)Z			

Figure 3 shows particularly rapid convergence with respect to increasing the number of diffuse functions, with all curves from maug- to aug- being closely grouped at all three ζ levels, although they are far from the unaugmented curves. The curves in Figure 4 also converge rapidly.

Another problem for which full augmentation is sometimes recommended is the calculation of electric dipole polarizabilities. First, we used the analytic polarizability algorithm in *MolPro 2009*³⁹ to calculate HF and MP2 spherically averaged polarizabilities

Table 3. HF, MP2, and MP2-F12 Finite-Field Approximation to zz Element of the Polarizability Tensor [\AA^3] of CH_4 for Basis Sets of Different Degrees of Augmentation

	CH_4		
	HF	MP2	MP2-F12
aug-cc-pV(T+d)Z	2.37	2.45	2.44
jul-cc-pV(T+d)Z	2.35	2.42	2.42
jun-cc-pV(T+d)Z	2.35	2.42	2.42
may-cc-pV(T+d)Z	2.19	2.23	2.33
cc-pV(T+d)Z	2.18	2.21	2.31
aug-cc-pV(D+d)Z	2.36	2.44	2.44
jul-cc-pV(D+d)Z	2.28	2.33	2.37
jun-cc-pV(D+d)Z	1.99	2.01	2.24
cc-pV(D+d)Z	1.91	1.91	2.18

for four molecules: methane (CH_4), water (H_2O), 3-mercapto-propan-1-ol ($\text{HO}(\text{CH}_2)_3\text{SH}$), and pyridine ($\text{C}_5\text{H}_5\text{N}$). These results are in Table 1. Then, we used the finite field approximation with a field strength of 0.005 atomic units to calculate the zz element of the polarizability tensor for water and methane by the HF, MP2, and MP2-F12 methods. These results are in Tables 2 and 3.

Table 1 shows that diffuse functions on the H atoms have a significant effect on the results (more than a couple of percent) in the case of double- ζ basis sets for MP2. However, for TZ and QZ, the average difference between aug- and jul- is only 0.04 and 0.02 \AA^3 , respectively. For nonhydrogenic atoms, as in the case of properties considered in earlier sections of this perspective, we find that, at the lower- ζ levels, the higher angular momentum functions affect the results significantly. However, the lower angular momentum functions in jun-TZ and jul-DZ provide most of the difference between augmented and nonaugmented basis sets. At higher ζ levels, full augmentation only adds additional cost and does not affect the accuracy to any significant degree. For example, the results obtained with the may-QZ basis set agree well with those obtained with aug-QZ with the largest percentage error being 2% for H_2O .

Tables 2 and 3 allow us to compare MP2-F12 convergence to MP2 convergence for polarizability elements. For the MP2-F12 calculations of electric dipole polarizabilities, one finds that may-, jun-, and jul- quadruple- ζ basis sets all have mean unsigned percentage errors (with respect to the largest fully augmented basis set) of about 2.7% and less; jul- and jun- triple- ζ basis sets have mean unsigned percentage errors less than 5.7%; and jul- double- ζ has a mean unsigned percentage error (with respect to aug-) of 11%.

We conclude that diffuse functions are especially important in polarizability calculations employing double- ζ basis sets, but partially augmented basis sets of higher- ζ levels can be useful for polarizabilities. This is consistent with findings of Sadlej et al. who introduced one diffuse function for each shell in their single contracted basis sets for electric properties, while their large basis sets are augmented only with a single s on hydrogenic atoms and a single s and a single p on other elements.^{40,41} These basis sets have been found to perform well for Raman intensities by Schlegel,⁴² which is also consistent with our results since Raman intensities depend on the derivatives of polarizabilities with respect to nuclear coordinates. However, Zuber and Hug⁴³ find that full augmentation on heavy atoms and additional p functions on the hydrogen atoms are necessary for the accurate

description of Raman and Raman optical activity (ROA) scattering.

Optical rotation is an example of a property, like van der Waals potential curves, that is especially sensitive to the diffuse space of the basis set. Diffuse functions are necessary for good results; however, the most thorough studies suggest that some savings can be made. For example, in refs 7 and 9, the authors state that their HF and DFT 6-311++G(2d,2p) results, which are augmented but much less than fully augmented, agree well with results obtained by aug-cc-pVTZ.

SUMMARY

Diffuse functions are an important component of basis sets, but in most cases, we recommend less than full augmentation with diffuse functions. For density functional theory, we usually recommend minimal augmentation. The need for diffuse functions is both greater and more variable in correlated wave function theory, so it is harder to make a general recommendation, but a good general starting point is a jun- level of augmentation.

ASSOCIATED CONTENT

S Supporting Information. Tables of the potential energy curves for van der Waals complexes. This material is available free of charge via the Internet at <http://pubs.acs.org/>.

AUTHOR INFORMATION

Corresponding Author

*E-mail: truhlar@umn.edu.

ACKNOWLEDGMENT

This work was supported by the National Science Foundation under grant no. CHE09-56776.

REFERENCES

- (1) Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. v. R. *J. Comput. Chem.* **1983**, *4*, 264.
- (2) Davidson, E. R.; Feller, D. *Chem. Rev.* **1986**, *86*, 681.
- (3) Papajak, E.; Leverentz, H. R.; Zheng, J.; Truhlar, D. G. *J. Chem. Theory Comput.* **2009**, *5*, 1197. **2009**, *5*, 3330(E).
- (4) Papajak, E.; Truhlar, D. G. *J. Chem. Theory Comput.* **2010**, *6*, 597.
- (5) Zheng, J.; Xu, X.; Truhlar, D. G. *Theor. Chem. Acc.* **2011**, *128*, 295.
- (6) Papajak, E.; Truhlar, D. G. *J. Chem. Theory Comput.* **2011**, *7*, 10.
- (7) Stephens, P. J.; Devlin, F. J.; Cheeseman, J. R.; Frisch, M. J. *J. Phys. Chem. A* **2001**, *105*, 5356.
- (8) Baranowska, A.; Laczowski, K. Z.; Sadlej, A. J. *J. Comput. Chem.* **2010**, *31*, 1176.
- (9) Cheeseman, J. R.; Frisch, M. J.; Devlin, F. J.; Stephens, P. J. *J. Phys. Chem. A* **2000**, *104*, 1039.
- (10) Kendall, R. A.; Dunning, T. H., Jr.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6769.
- (11) Frisch, M. J.; Pople, J. A.; Binkley, J. S. *J. Chem. Phys.* **1984**, *80*, 3265.
- (12) Fast, P. L.; Sanchez, M. L.; Truhlar, D. G. *Chem. Phys. Lett.* **1999**, *306*, 407.
- (13) Curtiss, L. A.; Raghavachari, K.; Redfern, C.; Rassolov, V.; Pople, J. A. *J. Chem. Phys.* **1998**, *109*, 7764.
- (14) F. Weigend, F.; Ahlrichs, R. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297.
- (15) Zheng, J.; Truhlar, D. G. *J. Phys. Chem. A* **2009**, *113*, 11919.

- (16) Zheng, J.; Truhlar, D. G. *Phys. Chem. Chem. Phys.* **2010**, *12*, 7782.
- (17) Yang, K.; Zheng, J.; Zhao, Y.; Truhlar, D. G. *J. Chem. Phys.* **2010**, *132*, 164117.
- (18) Meyer, M. M.; Kass, S. R. *J. Org. Chem.* **2010**, *75*, 4274.
- (19) Hermosilla, L.; Catak, S.; Van Speybroeck, V.; Waroquier, M.; Vandenberghe, J.; Motmans, F.; Adriaenssens, P.; Lutsen, L.; Cleij, T.; Vanderzande, D. *Macromolecules* **2010**, *43*, 7424.
- (20) Lynch, B. J.; Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A* **2003**, *107*, 1384.
- (21) Petersson, G. A.; Braunschweig, M. J. *Chem. Phys.* **1985**, *83*, 5129.
- (22) East, A. L. L.; Allen, W. D. *J. Chem. Phys.* **1993**, *99*, 4638.
- (23) Sullivan, M. B.; Iron, M. A.; Redfern, P. C.; Martin, J. M. L.; Curtiss, L. A.; Radom, L. *J. Phys. Chem. A* **2003**, *107*, 5617.
- (24) Ten-no, S. *Chem. Phys. Lett.* **2004**, *398*, 56.
- (25) Werner, H.-J.; Adler, T. B.; Manby, F. R. *J. Chem. Phys.* **2007**, *126*, 164102.
- (26) Werner, H. J.; Meyer, W. *Mol. Phys.* **1976**, *31*, 855.
- (27) Werner, H. J.; Meyer, W. *Phys. Rev. A* **1976**, *13*, 13.
- (28) Morrison, M. A.; Hay, P. J. *J. Phys. B* **1977**, *10*, L647.
- (29) Eades, R. A.; Truhlar, D. G.; Dixon, D. A. *Phys. Rev. A* **1979**, *20*, 867.
- (30) Douglass, C. H., Jr.; Weil, D. A.; Eades, R. A.; Truhlar, D. G.; Dixon, D. A. In *Chemical Applications of Atomic and Molecular Electrostatic Potentials*; Politzer, P., Truhlar, D. G., Eds.; Plenum: New York, 1981; p 173.
- (31) Darling, C. L.; Schlegel, H. B. *J. Phys. Chem.* **1994**, *98*, 5855.
- (32) Peterson, K. A.; Dunning, T. H., Jr. *THEOCHEM* **1997**, *400*, 93.
- (33) Cybulski, A. M. *J. Chem. Phys.* **1999**, *111*, 10520.
- (34) Oliveira, L. N.; Amaral, O. A. V.; Castro, M. A.; Fonseca, T. L. *Chem. Phys.* **2003**, *289*, 221.
- (35) Arruda, P. M.; Neto, A. C.; Jorge, F. E. *Int. J. Quantum Chem.* **2009**, *109*, 1189.
- (36) Camiletti, G. G.; Canal Neto, A.; Jorge, F. E.; Machado, S. F. *THEOCHEM* **2009**, *910*, 122.
- (37) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. *Chem. Phys. Lett.* **1989**, *157*, 479.
- (38) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, Revision A.1; Gaussian, Inc.: Wallingford, CT, 2009.
- (39) Werner, H.-J.; Knowles, P. J.; Lindh, R.; Manby, F. R.; Schutz, M.; Celani, P.; Korona, T.; Mitrushenkov, A.; Rauhut, G.; Adler, T. B.; Amos, R. D.; Bernhardsson, A.; Berning, A.; Cooper, D. L.; Deegan, M. J. O.; Dobbyn, A. J.; Eckert, F.; Goll, E.; Hampel, C.; Hetzer, G.; Hrenar, T.; Knizia, G.; Koppl, C.; Liu, Y.; Lloyd, A. W.; Mata, R. A.; May, A. J.; McNicholas, S. J.; Meyer, W.; Mura, M. E.; Nicklass, A.; Palmieri, P.; Pflüger, K.; Pitzer, R.; Reiher, M.; Schumann, U.; Stoll, H.; Stone, A. J.; Tarroni, R.; Thorsteinsson, T.; Wang, M.; Wolf, A. *MOLPRO*, version 2009.1; University College Consultants Ltd.: Cardiff, Wales, U.K., 2009.
- (40) Benkova, Z.; Sadlej, A. J.; Oakes, R. E.; Bell, S. E. *J. Comput. Chem.* **2005**, *26*, 145.
- (41) Baranowska, A.; Sadlej, A. J. *J. Comput. Chem.* **2010**, *31*, 552.
- (42) Halls, M. D.; Schlegel, H. B. *J. Chem. Phys.* **1999**, *111*, 8819.
- (43) Zuber, G.; Hug, W. *J. Phys. Chem.* **2004**, *108*, 2108.