Polarization consistent basis sets. III. The importance of diffuse functions

Frank Jensen
Department of Chemistry, University of Southern Denmark, DK-5230 Odense, Denmark

A sequence of diffuse functions to be used in connections with the previously defined polarization consistent basis sets are proposed based on energetic criteria and results for molecular properties. At the Hartree–Fock level the addition of a single set of diffuse *s*- and *p*-functions significantly improves the convergence of calculated electron affinities. A corresponding analysis at the density functional level indicates that only systems with high electron affinities have well-defined basis set limits with common exchange-correlation functionals that have electron self-interaction errors. The majority of reported density functional calculations of electron affinities appear to be artifacts of the limited basis set used. The good agreement with experiments for such calculations is most likely due to a reasonable modeling of the physics of the anionic species, rather than being a theoretically sound procedure. For molecular properties like dipole and quadrupole moments, and static polarizabilities, the addition of diffuse functions up to *d*-functions is required to reach the basis set limit in a consistent fashion, but higher order angular momentum functions are significantly less important. © *2002 American Institute of Physics.* [DOI: 10.1063/1.1515484]

I. INTRODUCTION

In the first paper in this series we proposed a sequence of polarization consistent (pc) basis sets, which is suitable for systematically approaching the basis set limit for Hartree–Fock (HF) and density functional (DF) methods.[1] We denote these basis sets pc-*n*, with *n* indicating the level of angular momentum functions included beyond those necessary for the isolated atom. A pc-1 basis set is thus of double zeta plus polarization (DZP) level, a pc-2 of triple zeta plus polarization (TZP) level, etc. In the second paper it was shown that the largest of these basis set, pc-4, is capable of achieving micro-Hartree accuracy in absolute energies, corresponding to an accuracy of better than 0.01 kJ/mol per atom for total atomization energies.[2] These results were obtained and calibrated for neutral and cationic systems. For anionic systems it is well-known that wave function based methods must employ basis sets which are augmented by additional functions with small exponents.[3] Such diffuse functions are also required for an accurate calculation of molecular properties that depend on the wave function tail. Anionic systems represent a formal problem for most current DF methods due to the improper cancellation of the self-interaction energy.[4] The present paper examines the question of augmenting the pc-*n* basis sets with diffuse functions in order to improve the basis set convergence for certain properties, and examines the problem of describing anionic systems by DF methods.

II. COMPUTATIONAL DETAILS

All calculations have been done using the GAUSSIAN 98 program package[5] with restricted open shell wave functions for open shell species. Density functional calculations have used the BLYP functional[6] and a grid consisting of 99 radial and 590 angular point for calculating the exchange-correlation contribution.[7] Basis set optimization have been done using a pseudo Newton–Raphson procedure, as de-

scribed earlier.[8] The previously defined pc-*n* basis sets have been used in their uncontracted forms. Contraction slightly deteriorates the results, analogous to the situation for the energy,[2] but trends mirror the present results.

III. ELECTRON AFFINITIES

In analogy with the previously strategy,[1] we rely on results from numerical Hartree–Fock calculations as an absolute reference. At the HF level the importance of each additional diffuse function can be evaluated from its energetic contribution, based on functions that have been optimized with respect to the exponents. In contrast to the previous analysis which lead to the pc-*n* basis sets,[1] however, the exponents of the additional diffuse functions cannot be fully optimized for all systems. For a system like $^-$CN, for example, the highest occupied molecular orbital (HOMO) is a σ-type orbital located mainly on carbon. Addition of a set of *s*- and *p*-functions on both carbon and nitrogen, followed by minimization of the total energy with respect to the exponents, leads to collapse of the nitrogen *p*-function, i.e., the exponent converges to the same value as the outer *p*-exponent of the underlying pc-*n* basis set. Even when all exponents can be optimized, as for example for the F$^-$ atom, the anionic character is often strongly localized on one type of basis function, i.e., the *p*-function for the F$^-$ atom. This means that the optimum *p*-exponent value is well-defined, but the *s*-exponent depends only indirectly on the energy change due to the additional electron. Using isolated atoms for determining the optimum exponents will thus only provide well-defined exponent values for one type of function. In order to provide a general procedure, molecular systems must be employed, but this often suffers from a variational collapse (*vide supra*). In the present case we have thus employed a strategy where the exponents of the additional diffuse basis functions are generated by scaling the outer pc-*n*

TABLE I. Energy contribution for augmenting functions for the $^-$CN and CN molecules. Energies are in atomic units, with the contribution for augmenting functions relative to entry immediately above. EA is the electron affinity in kJ/mol.

Basis	Aug	HF ΔE ($^-$CN)	ΔE (CN)	EA	BLYP ΔE ($^-$CN)	ΔE (CN)	EA
pc-0	none	−91.950 359	−91.886 886	166.65	−92.482 773	−92.424 661	152.58
	s1p1	−0.107 531	−0.042 336	337.82	−0.122 791	−0.038 132	374.85
pc-1	none	−92.283 453	−92.169 839	298.30	−92.810 374	−92.693 417	307.07
	s1p1	−0.014 847	−0.001 594	333.09	−0.025 800	−0.002 252	368.90
	s1p1d1	−0.002 738	−0.004 497	328.47	−0.001 143	−0.001 744	367.32
pc-2	none	−92.342 076	−92.220 540	319.09	−92.875 605	−92.742 380	349.79
	s1p1	−0.002 731	−0.000 235	325.65	−0.006 501	−0.000 253	366.19
	s1p1d1	−0.000 349	−0.000 436	325.42	−0.000 084	−0.000 089	366.18
	s1p1d1f1	−0.000 029	−0.000 022	325.44	−0.000 048	−0.000 059	366.15
pc-3	none	−92.348 435	−92.224 976	324.14	−92.884 522	−92.746 382	362.69
	s1p1	−0.000 349	−0.000 010	325.03	−0.001 274	−0.000 016	365.99
	s1p1d1	−0.000 014	−0.000 010	325.04	−0.000 008	−0.000 003	366.00
	s1p1d1f1	−0.000 004	−0.000 006	325.04	−0.000 003	−0.000 001	366.01
	s1p1d1f1g1	−0.000 002	−0.000 002	325.04	−0.000 001	−0.000 001	366.01
pc-4	none	−92.348 885	−92.225 123	324.94	−92.885 668	−92.746 532	365.30
	s1p1	−0.000 033	−0.000 001	325.02	−0.000 268	−0.000 000	366.01
	s1p1d1	−0.000 001	−0.000 000	325.02	−0.000 002	−0.000 000	366.01
	s1p1d1f1	−0.000 000	−0.000 000	325.02	−0.000 001	−0.000 000	366.01
	s1p1d1f1g1	−0.000 000	−0.000 000	325.02	−0.000 000	−0.000 000	366.01
	s1p1d1f1g1h1	−0.000 000	−0.000 000	325.02	−0.000 000	−0.000 000	366.01
	Limit[a]	−92.348 951	−92.225 134	325.08			
	Residual error[b]	−0.000 032	−0.000 009				

[a]Numerical HF result, Refs. 9, 17.
[b]Error for the fully augmented pc-4 basis set relative to the HF limit.

exponents by a factor common for all elements, but different for each angular momentum. The optimum scale factors are determined by minimizing the total energy. This prevents a variational collapse for atoms having little anionic character in the molecule, but allows an optimum exponent value for those atoms making a significant contribution to the orbital describing the additional electron. At the DF level of theory, however, this approach will in many cases lead to exponent divergence, due to inadequacies in the exchange-correlation potential, as discussed in more detail below.

A. Hartree–Fock results

The energetic contributions for adding diffuse functions to each of the pc-n basis set for the $^-$CN and CN molecules are shown in Table I, very similar results have been obtained for other systems. For the neutral CN radical the pc-4 basis set gives an energy 11 micro-Hartree higher than the numerical HF limit,[9] but the corresponding error for the anionic system is 66 micro-Hartree. Half of this error (33 micro-Hartree) is recovered by adding a set of diffuse s- and p-functions, with the higher angular momentum functions contributing less than 2 micro-Hartree. Addition of multiple diffuse s- and p-functions also has a very small effect, e.g., the energy obtained by adding two sets of s- and p-functions with full optimization of the scale factors is only 0.1 micro-Hartree lower than for adding just one set of diffuse functions. For the smaller pc-n basis sets the underlying basis set is less complete, and addition of diffuse polarization functions has a larger effect. The effect is very similar for the

anionic and neutral systems, indicating that the primary function of diffuse polarization functions for the smaller pc-n basis sets is to make up for deficiencies in the underlying basis set. There is a clear trend that each step up in angular momentum reduces the energetic importance by approximately an order of magnitude. The importance of a specific type of angular momentum functions is similarly reduced by approximately an order of magnitude by increasing the size of the underlying pc-n basis set.

The energetic importance of the diffuse s- and p-functions is much less for the neutral CN system than for $^-$CN, and changes in the calculated electron affinity (EA) for the $s1p1$ augmented basis sets consequently mirror the changes in the total energy for the anionic system. The calculated EA with the pc-1 basis set is improved from 298 kJ/mol to 333 kJ/mol by adding a set of diffuse s- and p-functions, but the change by adding a diffuse d-function is only 5 kJ/mol. This should be compared to the absolute error of 3 kJ/mol from the fully augmented pc-1 basis set compared to the numerical HF result. With the pc-2 basis set, the addition of diffuse s- and p-functions reduce the error in the calculated EA from 6.0 kJ/mol to 0.6 kJ/mol, while adding diffuse d- and f-functions further improves the result by 0.2 kJ/mol. For the pc-3 and pc-4 basis sets the contribution to the EA from all the diffuse polarization functions is negligible. Using an $s1p1$ augmented pc-4 basis set the numerical limit for the EA is reproduced to within 0.06 kJ/mol.

TABLE II. Electron affinities (kJ/mol) calculated with $s1p1$ augmented pc-n basis sets and the BLYP functional

System	pc-0	pc-1	pc-2	pc-3	pc-4	Expt.[a]
CN	389.4	368.8	366.2	366.0	366.0	372.6
F	380.1	365.6	361.2	360.9	360.9	328.2
OH	200.6	193.6	194.8	196.9	**197.5**	176.3
C	139.4	133.1	**133.7**	**135.4**	**136.0**	121.8
H	81.6	79.4	82.2	**83.5**	**84.0**	72.8
CF$_2$	29.4	**−10.6**	**−8.9**	**4.4**	**11.7**	17.3
NO	21.4	**−12.4**	**−9.3**	**1.0**	**7.7**	2.5

[a]Reference 10. Entries in **bold** indicate that attempts of exponent optimization lead to divergence.

B. Density functional results

For neutral systems we have shown that there is only a small difference between the basis set convergence for HF and DF wave functions, and the same basis set can thus be used for estimating both the HF and KS limits.[2] For anionic systems, however, the situation is somewhat different, since most of the commonly employed exchange-correlation functionals do not have a proper cancellation of the Coulomb self-interaction energy by the (approximate) exchange functional.[4] This is turn leads to an improper behavior of the corresponding potential far from the nucleus. Since anionic systems have loosely bound electrons, this incorrect behavior is problematic. Nevertheless, many DF methods have been used in combination with standard augmented basis sets for calculating EAs, and the results in general display a good agreement with experiments. Reference 10 gives a recent review of the literature, with references to the majority of EA calculations by DF methods.

A central question is whether the apparent success for describing anionic systems by DF methods is just an artifact of using a finite basis set, a problem closely related to the HOMO energy. If this is positive, the resulting orbital cannot be normalized and the corresponding electron is not in a bound state, implying that the calculated EA is merely an artifact of the limited basis set used. A negative value, on the other hand, indicates a bound state and thus a well-defined EA. A detailed investigation of the F$^−$ atom by Jarecki and Davidson[11] showed that the HOMO has a negative orbital energy with the BLYP functional when a large diffuse basis set is used, but care must be taken to ensure that the numerical integration of the exchange-correlation functional is accurate. With an aug-cc-pV5Z basis set augmented with 2 additional diffuse p-functions, the orbital energy appears to be converged with a value of 0.05 a.u., but addition of 16 diffuse p-functions is necessary to reach the actual converged value of −0.06 a.u., with the most diffuse function having an exponent of 10^{-7}.

It has been argued that the problem of incomplete cancellation of the self-interaction energy should be largest for small localized anions,[12] and the demonstration of a negative HOMO energy for a "problematic" system like F$^−$ should thus make DF methods valid for calculating EAs for most systems of interest. We believe that a much more important factor is the *magnitude* of the EA, which is closely related to the HOMO energy. A large EA corresponds to a low (negative) HOMO energy and consequently a spatially "tight" orbital. A small value of the EA, on the other hand, corre-

sponds to a loosely bound electron occupying a diffuse orbital with a high energy. Since the latter to a larger extend samples the problematic region far from the nucleus where the exchange-correlation potential has the wrong behavior, this indicates that DF methods with self-interaction errors will only be able to describe systems with high EA values. Since this implies (*vide infra*) that most EAs calculated by DF methods are just artifacts of the incomplete basis set used, we will elaborate on this in more detail.

In this context it is useful to consider the experimental EAs for the systems shown in Table II. The CN radical has a large electron affinity (373 kJ/mol), and the HOMO is consequently quite tight spatially, resembling the other valence orbitals. The NO radical, on the other hand, has a near zero EA (3 kJ/mol), and the (exact) HOMO is (will be) very diffuse. A radical like OH represents an intermediate case with an EA of 176 kJ/mol. Table I shows that with the BLYP exchange-correlation functional it is possible to optimize scale factors for the augmented pc-n basis sets for the $^−$CN molecule, with a behavior very similar to that for the HF method. The HOMO energy is calculated to be essentially zero with the $s1p1$ augmented pc-4 basis set, but the results for F$^−$ suggest that addition of more diffuse functions is likely to make the HOMO energy negative.[11]

For the F$^−$ system (experimental EA of 328 kJ/mol) it is also possible to optimize the exponents for $s1p1$ augmented pc-n basis sets (polarization functions do not contribute to the energy for F$^−$). The total energy obtained at the HF level with an $s1p1$ augmented pc-4 basis set is 10 micro-Hartree above the numerical HF limit. With the same basis set the BLYP method gives a HOMO energy of 0.048 a.u., which is very similar to the results obtained by Jarecki and Davidson with a similar sized basis set.[11] Their results, however, indicate that the total energy is lowered by 0.008 a.u. when multiple diffuse p-functions are added, which also reverse the sign of the HOMO energy. The procedure by optimizing $s1p1$ augmented pc-n basis sets for this system leads to an apparent EA basis set limit of 361 kJ/mol, but the results by Jarecki and Davidson[11] indicate that further enlargement of the basis set will increase this value by ~20 kJ/mol.

The observed basis set convergence for $^−$CN and F$^−$ is *not* representative of the general case. For $^−$OH a complete set of optimized scale factors can be determined at the HF level, analogous to the results shown in Table I, resulting in an EA basis set limit value of −13 kJ/mol. For the pc-0, pc-1, and pc-2 basis sets it is also possible to optimized scale factors for the augmenting functions with the BLYP method.

Optimization of scale factors for a *s1p1* augmented pc-3 basis set is also possible, however, with a *s1p1d1* augmentation the exponent optimization leads to a scale factor for the *p*-functions diverging towards a very large value, corresponding to the basis function exponents approaching zero. We have checked that this is not a consequence of keeping the exponents of the underlying pc-*n* basis set fixed, attempts of full optimization of all basis set exponents also leads to divergence. We have also checked that this is not due to the use of an insufficient grid for calculating the exchange-correlation energy. The results for the augmented pc-0, pc-1 and pc-2 basis sets indicate a positive HOMO energy of 0.07 a.u., i.e., an unbound electron, despite the fact that the difference in total energy between OH and $^-$OH corresponds to a sizable EA of 198 kJ/mol, which can be compared to the experimental value of 176 kJ/mol. The fact that it is possible to optimize the outer exponents for the pc-0, pc-1, and pc-2 basis sets is a consequence of the incompleteness of the underlying basis set, i.e., the additional augmenting function will to a certain extend also improve the description of the orbitals not being the HOMO. Once the basis set approach completeness for the neutral system, however, the augmenting functions primarily describe the HOMO, and consequently attempt to describe an unbound electron by letting the exponent approach zero. In agreement with these arguments the exponent divergence starts earlier for systems with a lower EA. For the carbon atom with an experimental EA of 122 kJ/mol, it is possible to optimize *s1p1* augmented pc-0 and pc-1 basis sets, but not an augmented pc-2 basis set. For the NO system with an experimental EA of 3 kJ/mol, only an *s1p1* augmented pc-0 basis set can be optimized, already the *s1p1* augmentation of the pc-1 basis set leads to divergence. The same level of divergence is observed for the CF_2 system with an experimental EA of 17 kJ/mol, showing that the more extended delocalization of the HOMO is less important than the magnitude of the EA for describing whether DF methods are appropriate for calculating EAs.

The use of hybrid DF methods, which include a fraction of exact HF exchange, improves the situation to a certain extend, since the HF method is completely self-interaction free. It does not remove the problem, however, but merely shifts the onset of the divergence to (slightly) lower values of the (experimental) EA. With the B3LYP hybrid,[13] which includes 20% HF exchange, a complete set (*s1p1d1f1g1*) of augmentation functions can be optimized with the pc-3 basis set for $^-$OH, but divergence is observed when augmentation functions are added to the pc-4 basis set. For NO$^-$ divergence is again observed already with an *s1p1* augmented pc-1 basis set, i.e., the addition of exact exchange in this case does not improve the situation.

The divergence in the basis set exponent optimization is a clear indication that the additional electron is not bound with the employed exchange-correlation functional. When using standard basis sets, however, this is not a useful criterion. A positive HOMO energy is a similar warning, but the results by Jarecki and Davidson[11] show that a (slightly) positive value for a limited basis set may become negative when sufficiently diffuse functions are added. A positive HOMO energy, however, is a clear indication that a more extensive basis set investigation is required. The present results indicate that only systems with high EAs have theoretically well-defined values with many common exchange-correlation functionals. For the BLYP method this limit appears to be close to 320 kJ/mol, which exclude most species of interest (only 13 of the 110 systems in Ref. 10 have EAs larger than 320 kJ/mol). Nevertheless, a sequence of calculations with standard basis sets augmented with a single set of diffuse functions may show an *apparent* basis sets convergence. This is illustrated in Table II, where the pc-*n* basis sets have been augmented with *s1p1* diffuse functions, with exponents determined by scaling the outer exponent in the regular basis set, as described below. It is clear that these results could easily lead to the conclusion that a DZP or TZP type basis set (like pc-1 and pc-2) augmented with diffuse functions is capable of giving results in respectable agreement with experiments, and that the remaining error can be attributed to the approximate nature of the exchange-correlation functional, since little change is observed upon going from the pc-1 to the pc-2 basis set. Note, however, that for the low EA systems the pc-3 and pc-4 results clearly show that the basis set is not converged.

The inability to establish a well-defined scale factor by optimization of the basis set exponents means that the energy of the anion is not well-defined. In a basis set with fixed exponents the addition of more and more diffuse functions will continue to lower the energy. In practice such calculations are troublesome, since it becomes difficult to achieve wave function convergence for basis sets containing very diffuse functions. The calculated EA with functionals having an incomplete cancellation of the self-interaction energy is therefore essentially a measure of the basis set incompleteness, and not a measure of the inherent accuracy of the functional. This result has implications for a large number of reported EAs using DF methods.[10] These calculations have, with a few exceptions, been done using standard basis sets, typical of DZP or TZP type, augmented with a single set of diffuse functions.

Although the present results show that only systems with high EAs have theoretically well-defined values when common exchange-correlation functionals are employed, a large body of reported calculations has shown that DF methods *operationally* give very respectable agreement with experimental EAs when used in combination with an augmented DZP or TZP type basis set.[10] We believe that this is an example of a methodology giving the right answer for the right *physical* reason, but not for the right *theoretical* reason. An augmented DZP or TZP type basis set effectively constrains the additional electron to occupy the physically correct region in space, and DFT provides a good estimate of the energy derived from the corresponding orbital/density. Provided that the augmenting functions are not too diffuse, the procedure thus *models* the real physical system, and can be used as a "parameterized" method for estimating EAs, where the parameterization consists of selecting a proper combination of exchange/correlation functional and basis set.

IV. MOLECULAR PROPERTIES

The EA is mainly related to the total energy of the anionic species, and the above analysis shows that diffuse s- and p-functions are the most important for systems with elements from the first row in the Periodic Table. There are, however, many other molecular properties that are known to be sensitive to an accurate description of the wave function tail, the electric dipole and quadrupole moments being some of the simplest. For the pc-n basis sets we have shown that the dipole moment converges monotomic towards the limiting value, but the convergence is rather slow and not smooth.[1]

The HF energy for anions provides a well-defined criterium for selecting diffuse basis function exponents, but no similar measure is available for properties. At the DF level of theory the total energy is not a useful criterium, since most anions have unbound electrons, and alternative strategies must be employed to assign suitable diffuse functions. To get an initial estimate of exponents suitable for diffuse s- and p-functions, we have determined the optimum scale factors for five molecular test cases ($^-$CN, $^-$OH, $^-$NH, NO$^-$, $^-$OF) at the HF level by minimizing the total energy. The scale factors for the s- and p-functions are well-defined, since they make a large contribution to the HOMO, as illustrated in Table I. The ratio between the optimized exponents of the diffuse functions (ζ_{dif}) and the outer exponents of the pc-n basis set (ζ_1) is in general slightly larger than the ratio between the two outermost exponents in the pc-n basis set (ζ_2/ζ_1). We have thus adapted a procedure where the diffuse s- and p-exponents are generated according to the following formula, with $K=0.20$:

$$\zeta_{dif} = \frac{\zeta_1}{(\zeta_2/\zeta_1 + K)}.$$

An exception is the exponent for the diffuse s-function for the pc-0 basis set, since ζ_2/ζ_1 is very large in this case, due to the exponent gap between the $1s$- and $2s$-orbitals. For this special case we have chosen a ratio slightly smaller than for the corresponding p-functions.

The choice of exponents for the diffuse polarization functions is more problematic. For the pc-3 and pc-4 basis sets the total energy is insensitive to the diffuse polarization functions (Table I), and the HF energy is therefore not a good criterium for choosing basis set exponents. The pc-1 and pc-2 basis sets are less complete than pc-3 and pc-4, and the energy optimized exponents for the diffuse polarization functions become rather tight to compensate for inadequacies in the underlying basis set. An energetic criterium is therefore questionably for choosing diffuse polarization exponents.

An alternative criterium is that the diffuse polarization functions should span a similar physical space as the diffuse s- and p-functions. Since the basis functions describe continuous distributions, several choices are possible for measuring the "similarity".[14] Requiring that a basis function of angular momentum L should have a maximum at the same radial distance as the diffuse p-function gives the formula $\zeta_L = L\zeta_{L=1}$. If the basis function is weighted with an r^2 term from the volume element, the corresponding formula becomes $\zeta_L = \frac{1}{2}(L+1)\zeta_{L=1}$. Other criteria like maximizing the overlap, or having the same expectation value of r, give similar formula, with slightly different multiplicative factors. We have examined several such choices for the diffuse polarization exponents, and settled on the formula $\zeta_L = (L+1)\zeta_{L=1}$. The pc-3 and pc-4 basis sets already contain fairly diffuse polarization functions, and the results are consequently insensitive to the exact choice of exponents. For the pc-1 and pc-2 basis sets the present choice appears to give slightly better agreement with the basis set limit than the other alternatives examined for a selection of molecular systems. In a sense we have chosen a parameterized formula for determining the diffuse polarization exponents to ensure that results with the augmented pc-1 and pc-2 basis sets model the basis set limit. The optimum exponents will of course depend on the specific molecular system and the specific property, but the present choice appears suitable for electric moments and polarizabilities.[15] The diffuse function exponents for the pc-n basis sets are somewhat smaller than for the aug-cc-pVxZ basis sets,[3] especially for the high angular momentum functions.

The results for the dipole moment of CO at the HF and BLYP levels are shown in Table III, very similar results have been obtained for other systems. The addition of diffuse s- and p-functions gives only a marginal improvement relative to the pc-n basis sets, but addition of a diffuse d-function significantly improves the basis set convergence. Higher angular momentum diffuse functions are significantly less important. Figure 1 shows the logarithmic error in the BLYP dipole moment for H$_2$O as a function of the number of basis functions. For this case diffuse s- and p-functions actually slightly deteriorates the calculated results for the pc-0, -1, and -2 basis sets, while inclusion of diffuse d-functions gives a significant improvement.

The quadrupole moment [defined as $\Theta = Q_{zz} - \frac{1}{2}(Q_{xx} + Q_{yy})$[16]] similarly measures the quality of the wave function tail, although at slightly larger distances from the nuclei than the dipole moment. The calculated results for the N$_2$ molecule are shown in Table III, and displays very similar characteristics as the dipole moment. Diffuse s-, p- and d-functions are important, but higher angular momentum diffuse functions have little effect, and an $s1p1d1$ augmented pc-2 basis set provides results close to the basis set limit. Figure 2 shows the logarithmic error in the BLYP quadrupole moment for NH$_3$ as a function of the number of basis functions. Addition of diffuse s- and p-functions provides a more smooth convergence than the unaugmented pc-n basis sets, but the calculated values are further removed from the basis set limit. Diffuse d-functions significantly improve the results, and diffuse f-functions have a somewhat larger effect than for the dipole moment.

The (static) dipole polarizability (defined as 1/3 of the trace of the polarizability tensor) is an example of a property, which might be expected to be sensitive to the presence of higher angular momentum diffuse functions. The results for FH in Table III, however, show very similar convergence characteristics as the dipole and quadrupole moments, i.e., diffuse s-, p- and d-functions are essential, but higher angular momentum functions are significantly less important. Figure

TABLE III. Dipole moment (Debye) for CO, quadrupole moment (Debye Å) for N_2, and static polarizability (a.u.) for FH calculated at the HF and BLYP levels of theory.

Basis	Aug	μ(CO)		$\Theta(N_2)$		α(FH)	
		HF	BLYP	HF	BLYP	HF	BLYP
pc-0	none	−0.8676	−0.2074	−2.2129	−2.2652	1.481	1.596
	s1p1	−0.7543	−0.1281	−2.0275	−2.0573	2.357	3.021
pc-1	none	−0.3835	0.1262	−1.5700	−1.8248	2.647	2.878
	s1p1	−0.3619	0.1314	−1.5010	−1.7081	3.199	3.923
	s1p1d1	−0.2549	0.1935	−1.1348	−1.4738	4.836	6.110
pc-2	none	−0.2910	0.1645	−1.4407	−1.7599	3.528	4.026
	s1p1	−0.2812	0.1790	−1.3953	−1.6663	3.768	4.553
	s1p1d1	−0.2638	0.1878	−1.2518	−1.5638	4.894	6.231
	s1p1d1f1	−0.2648	0.1870	−1.2631	−1.5742	4.891	6.280
pc-3	none	−0.2665	0.1828	−1.2842	−1.6178	4.536	5.483
	s1p1	−0.2658	0.1847	−1.2722	−1.5846	4.596	5.639
	s1p1d1	−0.2651	0.1850	−1.2622	−1.5745	4.905	6.304
	s1p1d1f1	−0.2651	0.1850	−1.2629	−1.5755	4.908	6.318
	s1p1d1f1g1	−0.2651	0.1850	−1.2628	−1.5755	4.907	6.319
pc-4	none	−0.2651	0.1847	−1.2649	−1.5826	4.783	5.942
	s1p1	−0.2651	0.1848	−1.2635	−1.5757	4.795	5.989
	s1p1d1	−0.2650	0.1848	−1.2639	−1.5757	4.909	6.316
	s1p1d1f1	−0.2650	0.1848	−1.2641	−1.5758	4.909	6.323
	s1p1d1f1g1	−0.2650	0.1848	−1.2641	−1.5757	4.910	6.324
	s1p1d1f1g1h1	−0.2650	0.1849	−1.2640	−1.5757	4.910	6.324
	Limit	−0.2650[a,c]	0.122[b,d]	−1.2642[a,c]	−1.52[b,e]	4.910[a,f]	5.60[b,g]

[a]Numerical HF result.
[b]Experimental value.
[c]Reference 16.
[d]Reference 19.

[e]Reference 20.
[f]Reference 18.
[g]Reference 21.

3 shows the logarithmic error in the BLYP polarizability for CO as a function of the number of basis functions. Diffuse *s*- and *p*-functions give only a marginal improvement, but addition of diffuse *d*-functions has a large effect.
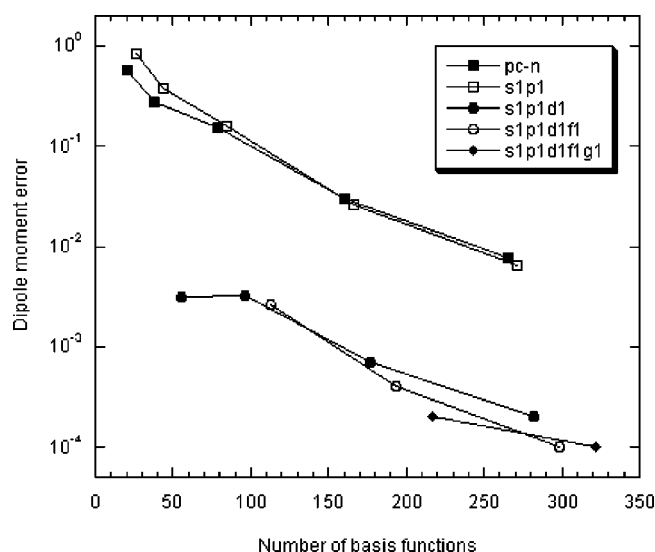
For the aug-cc-pVxZ basis sets[3] a full set of augmenting functions is included for a given value of *x*, but other basis sets often include only diffuse *s*- and *p*-functions. Given the results in Tables I and III, it is relevant to consider how many
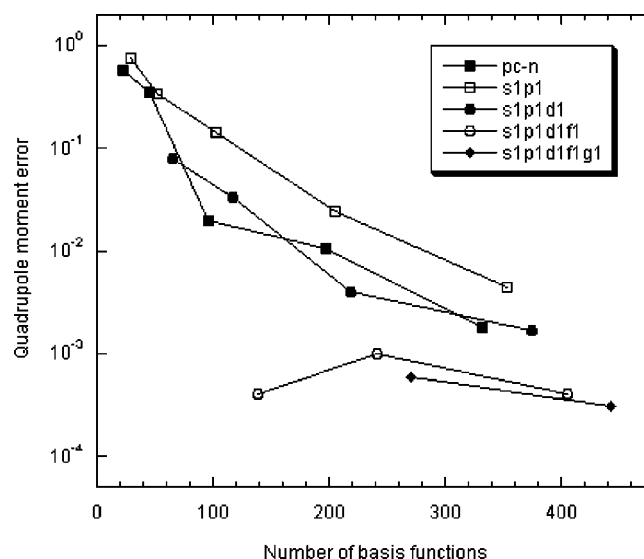

FIG. 1. Logarithmic error relative to the basis set limit for the BLYP dipole moment for H_2O as a function of the number of basis functions. Legend indicates the level of diffuse functions added to the pc-*n* basis sets.

FIG. 2. Logarithmic error relative to the basis set limit for the BLYP quadrupole moment for NH_3 as a function of the number of basis functions. Legend indicates the level of diffuse functions added to the pc-*n* basis sets.
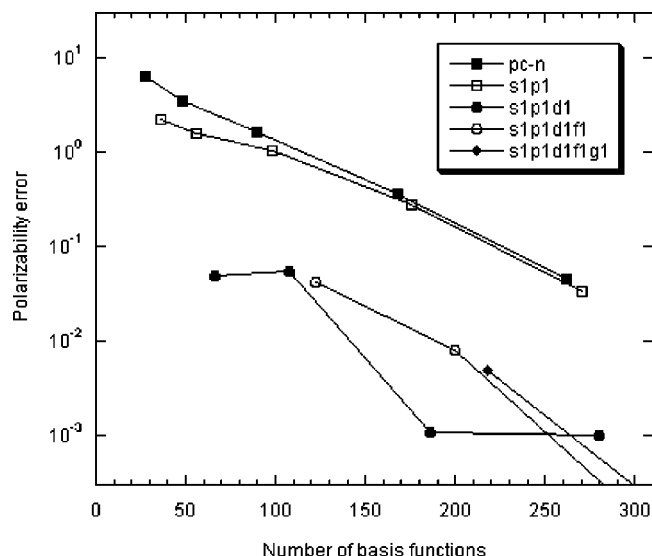
FIG. 3. Logarithmic error relative to the basis set limit for the BLYP polarizability for CO as a function of the number of basis functions. Legend indicates the level of diffuse functions added to the pc-*n* basis sets.

types of diffuse functions that should be included for a given pc-*n* basis set. For a property like EA, the diffuse *s*- and *p*-functions are by far the most important, while *d*-functions are essential for properties like electric moments and polarizability. One may thus consider using a sequence of pc-*n* basis sets augmented with only a subset of diffuse functions, e.g., only *s*- and *p*-functions for calculation of EAs. For the pc-1 and pc-2 basis sets, however, addition of diffuse *d*-functions gives a non-negligible improvement of the results (Table I). Since the size of a fully augmented pc-*n* basis set is always smaller than the unaugmented pc-$(n+1)$ basis set, the most cost-efficient way of achieving basis set convergence is probably to employ fully augmented pc-*n* basis sets, although for the pc-4 (and to some extend also pc-3) basis set, a removal of the high angular momentum diffuse functions will in most cases not affect the accuracy of the calculated property.

For all the above three properties does the addition of diffuse functions significantly improve the basis set convergence, and the augmented pc-4 results reproduce the HF limits very accurately. The similarity of the HF and BLYP basis set convergences leaves little doubts that the corresponding BLYP results are also essentially at the basis set limit. This in effect allows the difference between the BLYP/aug-pc-4 results and the experimental data to be attributed primarily to deficiencies in the exchange-correlation potential, and secondary to the neglect of vibrational averaging.

V. SUMMARY

Based on the previously proposed polarization consistent basis sets we suggest a scheme for adding diffuse functions. The exponents for the *s*- and *p*-functions are assigned by scaling the outer exponents for the regular pc-*n* basis set with a factor determined by explicit optimization at the Hartree–

Fock level. Exponents for polarization functions are assigned by scaling the corresponding exponents for the diffuse *p*-functions. Calculations for electron affinities show that diffuse *s*- and *p*-functions are significantly more important than diffuse polarization functions. With density functional methods, however, it appears that only systems with high electron affinity values have a well-defined basis set limit, due to the incomplete cancellation of the self-interaction energy with most common exchange-correlation functionals. These results strongly suggest that the majority of previous DF calculations of electron affinities are just artifacts of the limited basis set used. For molecular properties like dipole and quadrupole moments, and static polarizabilities, it is necessary to also include diffuse *d*-functions to achieve a smooth basis set convergence, but diffuse functions with higher angular momentum are less important.

[1] F. Jensen, J. Chem. Phys. **115**, 9113 (2001).
[2] F. Jensen, J. Chem. Phys. **116**, 7372 (2002).
[3] R. A. Kendall, T. H. Dunning, Jr., and R. J. Harrison, J. Chem. Phys. **96**, 6796 (1992); T. van Mourik and T. H. Dunning, Jr., Int. J. Quantum Chem. **76**, 205 (2000).
[4] N. Rüsch and S. B. Trickey, J. Chem. Phys. **106**, 8940 (1997); J. P. Perdew, Chem. Phys. Lett. **64**, 127 (1979); A. Zunger, J. P. Perdew, and G. L. Oliver, Solid State Commun. **34**, 933 (1980); N. A. Cordero and J. A. Alonso, Int. J. Quantum Chem. **61**, 253 (1997); L. A. Cole and J. P. Perdew, Phys. Rev. A **25**, 1265 (1982).
[5] M. J. Frisch, G. W. Trucks, H. B. Schlegel *et al.*, GAUSSIAN 98, Gaussian, Inc., Pittsburgh, PA, 1998.
[6] A. D. Becke, Phys. Rev. A **38**, 3098 (1988); C. Lee, W. Yang, and R. G. Parr, Phys. Rev. B **37**, 785 (1988).
[7] P. M. W. Gill, B. G. Johnson, and J. A. Pople, Chem. Phys. Lett. **209**, 506 (1993).
[8] F. Jensen, J. Chem. Phys. **110**, 6601 (1999).
[9] J. Kobus (private communication). The calculations were carried out using the finite difference HF method; J. Kobus, L. Laaksonen, and D. Sundholm, Comput. Phys. Commun. **98**, 346 (1996).
[10] J. C. Rienstra-Kiracofe, G. S. Tschumper, H. F. Schaefer III, S. Nanadi, and G. B. Ellison, Chem. Rev. **102**, 231 (2002).
[11] A. A. Jarecki and E. R. Davidson, Chem. Phys. Lett. **300**, 44 (1999).
[12] G. de Oliveria, J. M. L. Martin, F. de Proft, and P. Geerlings, Phys. Rev. A **60**, 1034 (1999).
[13] A. D. Becke, J. Chem. Phys. **98**, 5648 (1993); P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch, J. Phys. Chem. **98**, 11623 (1994).
[14] T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular Electronic-Structure Theory* (Wiley, Chichester, 2000).
[15] See EPAPS Document No. E-JCPSA6-117-304244 for supplementary tables. A direct link to this document may be found in the online article's HTML reference section. The document may also be reached via the EPAPS homepage (http://www.aip.org/pubservs/epaps.html) or from ftp.aip.org in the directory /epaps/. See the EPAPS homepage for more information.
[16] J. Kobus, D. Moncrieff, and S. Wilson, Phys. Rev. A **62**, 062503 (2000).
[17] J. Kobus, D. Moncrieff, and S. Wilson, Mol. Phys. **96**, 1559 (1999).
[18] J. Kobus, D. Moncrieff, and S. Wilson, J. Phys. B **34**, 5127 (2001).
[19] J. S. Muenter, J. Mol. Spectrosc. **55**, 490 (1975).
[20] J. M. M. Roco, A. C. Hernandez, S. Velasco, and A. Medina, J. Chem. Phys. **110**, 5218 (1999).
[21] A. Kumar and W. J. Meath, Can. J. Chem. **63**, 1616 (1985).