

## Basis Set Convergence of Nuclear Magnetic Shielding Constants Calculated by Density Functional Methods

Frank Jensen

*Department of Chemistry, University of Aarhus, Langelandsgade 140,  
DK-8000 Aarhus, Denmark*

Received January 11, 2008

**Abstract:** The previously proposed polarization consistent basis sets, optimized for density functional calculations, are evaluated for calculating nuclear magnetic shielding constants. It is shown that the basis set convergence can be improved by adding a single p-type function with a large exponent and allowing for a slight decontraction of the p functions. The resulting pcS-*n* basis sets should be suitable for calculating nuclear magnetic shielding constants with density functional methods and are shown to perform significantly better than existing alternatives for a comparable computational cost.

### I. Introduction

The use of nuclear magnetic resonance methods for probing molecular structures in solution-phase environments is well established, and technical and methodology improvements continue to push the limits for the size of molecules that can be handled. An increasingly important element for interpreting the experimental data is the simultaneous calculation of spectral information, which allows a direct correlation between molecular structure and quantities such as nuclear magnetic shielding and spin–spin coupling constants.<sup>1</sup> In order for this to become a routine procedure, it is necessary that there exist computational procedures that both are fast and have well-defined accuracies. While sophisticated methods such as coupled cluster can provide very accurate results, they are limited to relatively small systems.<sup>2</sup> Density functional methods,<sup>3</sup> on the other hand, are applicable to systems with hundreds or even thousands of atoms.<sup>4</sup> The main drawback of density functional methods is the inability to systematically improve the results, but Keal and Tozer have recently proposed new exchange–correlation functionals aimed at calculating nuclear magnetic shielding constants.<sup>5</sup>

The second component in performing electronic structure calculations is the use of a basis set for expanding the molecular orbitals. A large basis set will enable the full potential of the chosen method for calculating the wave function to be realized but also requires a large computational cost. A small basis set, on the other hand, is computationally efficient but introduces errors in the results. It is therefore

desirable to have a sequence of basis sets such that the accuracy can be controlled and assessed at each level, and at the same time being as compact as possible. For wave function based methods including electron correlation, the correlation consistent (cc-pVXZ)<sup>6</sup> basis sets developed by Dunning and co-workers represent such a hierarchy for energies and structural properties. For independent particle models, such as density functional theory, we have developed the polarization consistent (pc-*n*)<sup>7</sup> basis sets for providing a fast and controlled convergence toward the basis set limit.

In discussions of basis set convergence, it is important to realize that different methods and properties have different basis set requirements and convergence rates. Both the cc-pVXZ and pc-*n* basis sets have been constructed using energetic criteria, such that functions which contribute similar amounts of energy are included at the same stage, and this leads to the maximum angular momentum function included in the basis set as the natural expansion parameter. The differences between the cc-pVXZ and pc-*n* basis sets are related to the fact that the electron correlation energy converges as an inverse polynomial in the maximum angular momentum function,<sup>8</sup> while the density functional energy displays an exponential convergence.<sup>9</sup> By virtue of construction, the cc-pVXZ and pc-*n* basis sets are therefore expected to provide the fastest convergence toward the limiting value for correlation and density functional energies, respectively. Energetically related properties, such as equilibrium geometries and vibrational frequencies, usually also display a smooth convergence toward the basis set limit.<sup>10</sup>

For molecular properties that depend on the energetically unimportant region far from the nuclei, such as electric multipole moments and polarizabilities, the basis set convergence can be substantially improved by adding diffuse functions, leading to the aug-cc-pVXZ<sup>11</sup> and aug-pc-*n*<sup>12</sup> basis sets. In recent work we have shown how the basis set convergence for calculating nuclear spin–spin coupling constants can be improved by adding tight functions, leading to the definition of the pcJ-*n* basis sets.<sup>13</sup> The spin–spin coupling constant has four independent contributions, and an interesting observation was that the different operators have different basis set requirements. The Fermi-contact operator is only sensitive to the presence of (tight) s-type functions and the paramagnetic spin–orbit (PSO) operator is only sensitive to the presence of p-type functions, while the spin-dipole operator is sensitive to p-, d-, and f-type functions. In order to ensure a fast basis set convergence of the spin–spin coupling constant, it was therefore necessary to add tight s-, p-, d-, and f-type basis functions.

The nuclear magnetic shielding constant  $\sigma$  can be defined as the second derivative of the energy with respect to an external magnetic field  $\mathbf{B}$  and a nuclear magnetic moment  $\mathbf{I}$ . In a perturbation formulation, the shielding constant can be written in terms of a diamagnetic and paramagnetic contribution, where the former is calculated as an expectation value of the diamagnetic shielding operator ( $\mathbf{H}^{\text{DS}}$ ) while the latter is calculated as a response property of the paramagnetic spin–orbit ( $\mathbf{H}^{\text{PSO}}$ ) and orbital Zeeman ( $\mathbf{L}_G$ ) operators.<sup>14,15</sup>

$$\begin{aligned} \sigma &= \frac{\partial^2 E}{\partial \mathbf{B} \partial \mathbf{I}} \\ &= \langle \Psi_0 | \mathbf{H}^{\text{DS}} | \Psi_0 \rangle - 2 \sum_{n \neq 0} \frac{\langle \Psi_0 | \mathbf{H}^{\text{PSO}} | \Psi_n \rangle \langle \Psi_n | \mathbf{L}_G | \Psi_0 \rangle}{E_0 - E_n} \\ \mathbf{H}^{\text{DS}} &= \frac{g_A \mu_N}{2c^2} \frac{\mathbf{r}_{iG}^t \mathbf{r}_{iA} - \mathbf{r}_{iA} \mathbf{r}_{iG}^t}{r_{iA}^3} \\ \mathbf{H}^{\text{PSO}} &= \frac{g_A \mu_N}{c^2} \frac{\mathbf{r}_{iA} \times \mathbf{p}_i}{r_{iA}^3} \\ \mathbf{L}_G &= \frac{1}{2} \mathbf{r}_{iG} \times \mathbf{p}_i \end{aligned} \quad (1)$$

Here  $\mathbf{r}_{iA/G}$  denoted the position vector between electron *i* and nucleus *A* or the gauge origin *G*,  $\mu_N$  is the nuclear magneton, and  $g_A$  is the nuclear *g* factor. The nuclear magnetic shielding constant is a  $3 \times 3$  tensor, but only the average isotropic component corresponding to one-third of the trace of  $\sigma$  is observed in solution, and we will consequently focus on this. It is customary to use the units of ppm, and this will also be the case here.

Given our findings for the basis set requirements of the PSO operator, it follows that the basis set convergence for nuclear magnetic shielding constants potentially could be improved by adding tight p functions. Furthermore, to our knowledge detailed basis set requirements of the diamagnetic shielding and orbital Zeeman terms have not been investigated. The present paper examines the basis set convergence of the nuclear magnetic shielding constant using density functional methods in more detail.

A large body of previous work on calculating nuclear magnetic shielding constants exists, with the majority employing standard basis sets, such as the Pople style *k*-*lm*G basis sets,<sup>16</sup> the Ahlrichs basis sets of double, triple, and quadruple quality,<sup>17</sup> and the cc-pVXZ<sup>18</sup> and pc-*n*<sup>19</sup> families of basis sets. The IGLO basis sets have been proposed explicitly for magnetic properties,<sup>20</sup> and Manninen and Vaara have proposed to use basis sets complete to within a given threshold in a given exponent range for calculating magnetic properties, but only a single basis set suitable for first-row elements and hydrogen has been defined.<sup>21</sup>

## II. Computational Details

All calculations have been performed with the Dalton<sup>22</sup> and Gaussian-03<sup>23</sup> program packages using the KT3 and B3LYP<sup>24</sup> exchange-correlation functionals. The GIAO technique has been employed to ensure gauge independence of the calculated results.<sup>25</sup> Molecular geometries have been taken from the G3 data set<sup>26</sup> or optimized at the B3LYP/6-31G(d,p) level. We emphasize that only the convergence with respect to the basis set is investigated in the present case, and no attempt is made to compare with experimental results. The latter requires attention to the molecular geometry and the quality of the exchange-correlation functional as well as vibrational and environmental effects. It is demonstrated that the basis set convergence is very similar for the two employed functionals, and the basis set convergence should therefore be representative for Hartree–Fock and density functional methods in general.

## III. Improved Basis Sets for Calculating Nuclear Magnetic Shielding Constants

The notation for the polarization consistent basis sets is pc-*n*, where *n* indicates the level of polarization beyond the atomic system, i.e. pc-0 is unpolarized, pc-1 is of double- $\zeta$  quality with a single polarization function, pc-2 is of triple- $\zeta$  quality with d- and f-type polarization functions, etc. An initial exploration using the uncontracted pc-*n* basis sets showed that only p-type tight functions had any significant influence on the calculated nuclear magnetic shielding constants, which is consistent with the findings for the PSO operator in the previous study. Diffuse functions in some cases also had a significant effect, which may be related to the orbital Zeeman operator or simply to the fact that polar systems with lone pairs in general require diffuse functions for an adequate description. The diamagnetic shielding operator was not found to have additional basis set requirements beyond those already included for representing the electron density.

In order to determine the optimum exponents for the tight p functions, we employed an optimization procedure analogous to that used for defining the pcJ basis sets where the optimization criterion is to maximize the change in the nuclear shielding constant relative to the regular pc-*n* basis set. The optimum exponents determined for a selection of molecular systems showed a high degree of regularity with a near-constant ratio of 6.5 relative to the highest exponent already included in the basis set. Addition of a second tight

**Table 1.** Basis Set Composition in Terms of Uncontracted Functions

| basis                        | H           | Li–Be                      | B–Ne           | Na–Mg                       | Al–Ar          |
|------------------------------|-------------|----------------------------|----------------|-----------------------------|----------------|
| pc-0                         | 3s          | 5s1p                       | 5s3p           | 8s5p                        | 8s6p           |
| pc-1                         | 4s1p        | 7s3p                       | 7s4p1d         | 11s7p                       | 11s8p1d        |
| pc-2                         | 6s2p1d      | 10s4p1d                    | 10s6p2d1f      | 13s9p1d                     | 13s10p2d1f     |
| pc-3                         | 9s4p2d1f    | 14s6p2d1f                  | 14s9p4d2f1g    | 17s12p2d1f                  | 17s13p4d2f1g   |
| pc-4                         | 11s6p3d2f1g | 19s8p3d2f1g                | 18s11p6d3f2g1h | 21s15p3d2f1g                | 21s16p6d3f2g1h |
| pcS-0                        | 3s          | 5s2p                       | 5s4p           | 8s6p                        | 8s7p           |
| pcS-1                        | 4s2p        | 7s4p                       | 7s5p1d         | 11s8p                       | 11s9p1d        |
| pcS-2                        | 6s3p1d      | 10s5p1d                    | 10s7p2d1f      | 13s10p1d                    | 13s11p2d1f     |
| pcS-3                        | 9s5p2d1f    | 14s7p2d1f                  | 14s10p4d2f1g   | 17s13p2d1f                  | 17s14p4d2f1g   |
| pcS-4                        | 11s7p3d2f1g | 19s9p3d2f1g                | 18s12p6d3f2g1h | 21s16p3d2f1g                | 21s17p6d3f2g1h |
| STO-3G                       | 3s          | 6s3p                       | 6s3p           | 9s6p                        | 9s6p           |
| 6-31G(d,p)                   | 4s1p        | 10s4p1d                    | 10s4p1d        | 16s10p1d                    | 16s10p1d       |
| 6-311G(2df,2pd) <sup>a</sup> | 5s2p1d      | 11s5p2d1f                  | 11s5p2d1f      | 13s9p2d1f                   | 13s9p2d1f      |
| cc-pVDZ                      | 4s1p        | 9s4p1d                     | 9s4p1d         | 12s8p1d                     | 12s8p2d        |
| cc-pVTZ                      | 5s2p1d      | 11s5p2d1f                  | 10s5p2d1f      | 15s10p2d1f <sup>f</sup>     | 15s9p3d1f      |
| cc-pVQZ                      | 6s3p2d1f    | 12s6p3d2f1g                | 12s6p3d2f1g    | 16s12p3d2f1g <sup>g</sup>   | 16s11p4d2f1g   |
| cc-pV5Z                      | 8s4p3d2f1g  | 14s8p4d3f2g1h <sup>b</sup> | 14s8p4d3f2g1h  | 20s14p4d3f2g1h <sup>h</sup> | 20s12p5d3f2g1h |
| cc-pCVDZ                     | 4s1p        | 9s4p1d                     | 10s5p1d        | 13s9p2d                     | 13s9p2d        |
| cc-pCVTZ                     | 5s2p1d      | 11s5p2d1f                  | 12s7p3d1f      | 17s12p4d2f <sup>i</sup>     | 17s11p4d2f     |
| cc-pCVQZ                     | 6s3p2d1f    | 12s6p3d2f1g                | 15s9p5d3f1g    | 19s15p6d4f2g <sup>j</sup>   | 19s14p6d4f2g   |
| SVP                          | 4s1p        | 7s4p <sup>c</sup>          | 7s4p1d         | 10s6p                       | 10s7p1d        |
| TZV                          | 5s2p1d      | 11s4p <sup>d</sup>         | 11s6p2d1f      | 14s8p1d                     | 14s9p2d1f      |
| QZV                          | 7s3p2d1f    | 15s7p2d1f <sup>e</sup>     | 15s7p2d1f      | 20s12p3d1f                  | 20s14p4d2f1g   |
| IGLO-II                      | 5s1p        | 9s5p1d                     | 9s5p1d         | 11s7p2d                     | 11s7p2d        |
| IGLO-III                     | 6s2p        |                            | 11s7p2d        |                             | 12s8p3d        |

<sup>a</sup> McLean-Chandler basis set<sup>30</sup> for Na–Ar. <sup>b</sup> 7p for Li. <sup>c</sup> 1p for Li. <sup>d</sup> 3p for Li. <sup>e</sup> 6p for Li. <sup>f</sup> 16s for Na. <sup>g</sup> 19s for Na. <sup>h</sup> 19s12p for Na. <sup>i</sup> 18s for Na. <sup>j</sup> 22s for Na.

p function was in all cases found to give almost negligible changes. These findings are in complete agreement with the previous study for the PSO operator and suggest that a faster basis set convergence can be obtained by adding a single tight p-type function to the regular pc-*n* basis sets.

The lighter s-block elements (H, He, Li, Be) do not have occupied p orbitals, and it is therefore not a priori clear whether the results for these elements will be sensitive to the presence of tight p functions. In test calculations we found that the results for these elements are less affected by tight p functions than for the p-block elements but that a systematic improvement does occur, and we have therefore elected to include a tight p function also for these elements. We thus define a sequence of pc-type basis sets optimized for calculating nuclear magnetic shielding constants by adding a single tight p-type function to the regular pc-*n* basis sets to produce a basis set denoted pcS-*n*, where S indicates shielding. The size of the basis sets are shown in Table 1.

The pc-*n* basis sets employ a general contraction scheme using atomic orbital coefficients, where the degree of contraction is determined by the criterion that the contraction error should be smaller than the error of the uncontracted basis set relative to the basis set limit. For the pc-*n* basis sets this leads automatically to the pc-1 basis set being of double- $\zeta$  quality in the valence region, the pc-2 basis set being of triple- $\zeta$  quality, etc. When this criterion is used for the pcJ-*n* basis sets, it allows only a small degree of contraction, as the nuclear spin–spin coupling constant requires substantial flexibility in the core region. We have employed the same strategy in the present case, and it is found that the nuclear magnetic shielding constant requires more flexibility in the inner valence region for the p orbitals than the regular pc-*n* basis sets, and the recommended contractions are shown in Table 2. The pcS-0 basis set is of

the same size as the pc-0 basis set, while the pcS-1 and pcS-2 basis sets have one or two more (contracted) p functions. The pcS-3 basis set requires further decontraction of the p functions to ensure that the contraction error remains below the inherent error of the uncontracted basis set. The pcS-4 basis set is sufficiently large that the remaining basis set errors are only fractions of a ppm, and it is difficult to devise a contraction scheme without degrading this accuracy. We have chosen the contraction shown in Table 2, where the s contraction is relaxed for second-row elements relative to the pc-4 basis set and the p functions are almost uncontracted.

#### IV. Basis Set Convergence

We have examined the performance of the pcS-*n* and aug-pcS-*n* families of basis sets, as well as a selection of other popular basis sets, shown in Tables 1 and 2, for the systems in Table 3. As there are some differences between the nuclear magnetic shielding constants, the results will be divided into five groups: Hydrogen (*H*), first- ( $M_1 = \text{Li, Be}$ ) and second-row ( $M_2 = \text{Na, Mg}$ ) metallic elements, and first- ( $A_1 = \text{C–F}$ ) and second -row ( $A_2 = \text{Si–Cl}$ ) main-group elements. The basis set limiting value is in each case taken as the uncontracted aug-pcS-4 value, except for the largest systems where the pcS-4 results were used. The difference in the results with the (uncontracted) aug-pcS-3 and aug-pcS-4 basis sets indicates that the reference values are converged well below 0.001 ppm for hydrogen and 0.1 ppm for the remaining elements. A couple of pathological cases where this is not the case are discussed at the end of this section.

Table 4 shows the (uncontracted) basis set errors relative to the aug-pcS-4 results quantified in terms of the mean absolute deviation (MAD) over the symmetry-unique shielding constants using the KT3 and B3LYP exchange-correla-

**Table 2.** Basis Set Composition in Terms of Contracted Functions

| basis                        | H          | Li–Be                 | B–Ne          | Na–Mg                 | Al–Ar          |
|------------------------------|------------|-----------------------|---------------|-----------------------|----------------|
| pc-0                         | 2s         | 3s1p                  | 3s2p          | 4s2p                  | 4s3p           |
| pc-1                         | 2s1p       | 3s2p                  | 3s2p1d        | 4s2p                  | 4s3p1d         |
| pc-2                         | 3s2p1d     | 4s2p1d                | 4s3p2d1f      | 5s3p1d                | 5s4p2d1f       |
| pc-3                         | 5s4p2d1f   | 6s3p2d1f              | 6s5p4d2fg1    | 6s5p4d2d1f            | 6s5p4d2f1g     |
| pc-4                         | 7s6p3d2f1g | 8s4p3d2f1g            | 8s7p6d3f2g1h  | 7s5p3d2f1g            | 7s6p6d3f2g1h   |
| pcS-0                        | 2s         | 3s1p                  | 3s2p          | 4s2p                  | 4s3p           |
| pcS-1                        | 2s1p       | 3s3p                  | 3s3p1d        | 4s4p                  | 4s4p1d         |
| pcS-2                        | 3s2p1d     | 4s3p1d                | 4s4p2d1f      | 5s6p1d                | 5s6p2d1f       |
| pcS-3                        | 5s4p2d1f   | 6s3p2d1f              | 6s8p4d2fg1    | 7s8p2d1f              | 7s9p4d2f1g     |
| pcS-4                        | 7s6p3d2f1g | 8s8p3d2f1g            | 8s10p6d3f2g1h | 10s11p3d2f1g          | 10s12p6d3f2g1h |
| STO-3G                       | 1s         | 2s1p                  | 2s1p          | 3s2p                  | 3s2p           |
| 6-31G(d,p)                   | 2s1p       | 3s2p1d                | 3s2p1d        | 4s3p1d                | 4s3p1d         |
| 6-311G(2df,2pd) <sup>a</sup> | 3s2p1d     | 4s3p2d1f              | 4s3p2d1f      | 6s5p2d1f              | 6s5p2d1f       |
| cc-pVDZ                      | 2s1p       | 3s2p1d                | 3s2p1d        | 4s3p1d                | 4s3p2d         |
| cc-pVTZ                      | 3s2p1d     | 4s3p2d1f              | 4s3p2d1f      | 5s4p2d1f              | 5s4p3d1f       |
| cc-pVQZ                      | 4s3p2d1f   | 5s4p3d2f1g            | 5s4p3d2f1g    | 6s5p3d2f1g            | 6s5p4d2f1g     |
| cc-pV5Z                      | 5s4p3d2f1g | 6s5p4d3f2g1h          | 6s5p4d3f2g1h  | 7s6p4d3f2g1h          | 7s6p5d3f2g1h   |
| cc-pCVDZ                     | 2s1p       | 4s3p1d                | 4s3p1d        | 5s4p2d                | 5s4p2d         |
| cc-pCVTZ                     | 3s2p1d     | 6s5p3d1f              | 6s5p3d1f      | 7s6p4d2f              | 7s6p4d2f       |
| cc-pCVQZ                     | 4s3p2d1f   | 8s7p5d3f1g            | 8s7p5d3f1g    | 9s8p6d4f2g            | 9s8p6d4f2g     |
| SVP                          | 2s1p       | 3s2p <sup>b</sup>     | 3s2p1d        | 4s2p                  | 4s3p1d         |
| TZV                          | 3s2p1d     | 5s3p <sup>c</sup>     | 5s3p2d1f      | 5s4p1d                | 5s4p2d1f       |
| QZV                          | 4s3p2d1f   | 7s4p2d1f <sup>d</sup> | 7s4p3d2f1g    | 9s5p4d1f <sup>e</sup> | 9s6p4d2f1g     |
| IGLO-II                      | 3s1p       |                       | 5s4p1d        |                       | 7s6p2d         |
| IGLO-III                     | 4s2p       |                       | 7s6p2d        |                       | 8s7p3d         |

<sup>a</sup> McLean-Chandler basis set<sup>30</sup> for Na–Ar. <sup>b</sup> 1p for Li. <sup>c</sup> 2p for Li. <sup>d</sup> 6s for Li. <sup>e</sup> 3d for Na.

**Table 3.** Molecular Systems Used for Calibration

CH<sub>4</sub>, NH<sub>3</sub>, H<sub>2</sub>O, HF, N<sub>2</sub>, F<sub>2</sub>, CO, CO<sub>2</sub>, F<sub>2</sub>O  
 C<sub>2</sub>H<sub>2</sub>, C<sub>2</sub>H<sub>4</sub>, C<sub>2</sub>H<sub>6</sub>, H<sub>2</sub>CO, HCOOH, H<sub>2</sub>CS, N<sub>2</sub>H<sub>2</sub>, N<sub>2</sub>H<sub>4</sub>  
 CH<sub>3</sub>NH<sub>2</sub>, CH<sub>3</sub>NO<sub>2</sub>, CH<sub>3</sub>OH, CH<sub>3</sub>F, CH<sub>3</sub>CN, CH<sub>2</sub>F<sub>2</sub>, CH<sub>3</sub>CHO,  
 H<sub>2</sub>CCHCN  
 CH<sub>3</sub>SiH<sub>3</sub>, CH<sub>3</sub>PH<sub>2</sub>, CH<sub>3</sub>SH, CH<sub>3</sub>Cl, CH<sub>2</sub>Cl<sub>2</sub>, C<sub>2</sub>N<sub>2</sub>, C<sub>2</sub>F<sub>4</sub>, C<sub>2</sub>Cl<sub>4</sub>  
 cyclopropene, butadiene, benzene, furan, pyrrole, thiophene,  
 pyridine, (CH<sub>3</sub>)<sub>2</sub>SO  
 LiH, LiCH<sub>3</sub>, Li<sub>2</sub>O, LiF, Li<sub>2</sub>S<sub>2</sub>, LiCl  
 Be<sub>2</sub>H<sub>4</sub>, Be(CH<sub>3</sub>)<sub>2</sub>, BeF<sub>2</sub>, BeCl<sub>2</sub>  
 SiH<sub>4</sub>, PH<sub>3</sub>, H<sub>2</sub>S, HCl, Si<sub>2</sub>H<sub>2</sub>  
 P<sub>2</sub>, Cl<sub>2</sub>, CS, CS<sub>2</sub>, CSO, PF<sub>5</sub>, PCl<sub>5</sub>, SF<sub>6</sub>, Cl<sub>2</sub>SO<sub>2</sub>  
 NaH, NaCH<sub>3</sub>, Na<sub>2</sub>O, NaF, Na<sub>2</sub>S, NaCl  
 Mg<sub>2</sub>H<sub>4</sub>, Mg(CH<sub>3</sub>)<sub>2</sub>, MgF<sub>2</sub>, MgCl<sub>2</sub>

tion functionals. It is seen that addition of both diffuse functions (aug-pc-*n*) and a tight p function (pcS-*n*) to the pc-*n* basis sets has an effect and, with a few exceptions, leads to lower basis set errors. It is noticeable that in some cases there is a cooperative effect, where addition of both tight and diffuse functions leads to a larger improvement than the sum of the two individual effects. At the pc-1 level with the KT3 functional for first-row nonmetallic elements ( $A_1 = C-F$ ), for example, the addition of diffuse functions improves the MAD by 2.7 ppm, the addition of a tight p function lowers the MAD by 10.1 ppm, but the combined effect is 15.2 ppm. These values should be compared to the MAD value of 18.0 ppm for the pc-1 basis set, which consequently is lowered to 2.8 ppm with the aug-pcS-1 basis set.

The basis set errors are much smaller for hydrogen shieldings than for the other elements, and the metallic s-block elements (Li, Be, Na, Mg) have lower errors at a given level than the p-block elements (C–F, Si–Cl). This is consistent with the importance of p functions for the PSO operator and the dominance of s-type bonding for hydrogen and the metallic elements. The basis set error for the second-

row elements Si–Cl tend to be somewhat larger than for the first-row elements C–F, and the improvement by adding diffuse and tight functions is smaller. Table 4 shows that addition of a single tight p function can improve the basis set convergence, although the improvement is not as spectacular as for the spin–spin coupling constants. It can also be noted that the basis set convergence is very similar for the two employed exchange–correlation functionals.

Table 5 compares the performance of the (contracted) pcS-*n* and aug-pcS-*n* basis sets with a selection of other commonly used basis sets for the set of systems in Table 3. The Pople type STO-3G,<sup>27</sup> 6-31G(d,p),<sup>28</sup> and 6-311G-(2df,2pd)<sup>29,30</sup> basis sets represent minimum and double- and triple- $\zeta$  quality, and the last two can be augmented with diffuse functions (6-31++G(d,p) and 6-311++G(2df,2pd)). The Dunning family of correlation consistent basis sets cc-pVXZ ( $X = D, T, Q, 5$ )<sup>31</sup> can be augmented with both diffuse (aug-cc-pVXZ)<sup>32</sup> and tight functions (cc-pCVXZ).<sup>33</sup> The latter has been designed for recovering core and core-valence correlation and adds tight functions of all types as well as multiple sets of tight functions for the T and Q basis sets (Table 2). The Ahlrichs-type basis sets SVP,<sup>34</sup> TZP,<sup>35</sup> and QZP<sup>36</sup> are of double-, triple-, and quadruple- $\zeta$  quality, but no standard sets of diffuse and tight functions have been defined. The IGLO basis sets have been designed for magnetic properties but are not defined for s-block elements. They are furthermore somewhat difficult to classify in terms of quality, as the number of s and p functions (Table 2) indicates at least quadruple- $\zeta$  quality, but the lack of high angular momentum functions suggests that they are at best of triple- $\zeta$  quality. A comparison of the results in Table 5 for the KT3 and B3LYP functionals indicates a very similar basis set behavior, indicating that the conclusions discussed in the following should be valid for density functional methods in general.



**Table 4.** Mean Absolute Deviations (ppm) Relative to the Basis Set Limit for the Symmetry-Unique Nuclear Magnetic Shielding Constants for the Systems in Table 3<sup>a</sup>

| basis set | KT3      |                       |                       |                       |                       | B3LYP    |                       |                       |                       |                       |
|-----------|----------|-----------------------|-----------------------|-----------------------|-----------------------|----------|-----------------------|-----------------------|-----------------------|-----------------------|
|           | <i>H</i> | <i>M</i> <sub>1</sub> | <i>A</i> <sub>1</sub> | <i>M</i> <sub>2</sub> | <i>A</i> <sub>2</sub> | <i>H</i> | <i>M</i> <sub>1</sub> | <i>A</i> <sub>1</sub> | <i>M</i> <sub>2</sub> | <i>A</i> <sub>2</sub> |
| pc-0      | 1.5      | 14.9                  | 63.5                  | 7.4                   | 81.0                  | 1.5      | 15.1                  | 64.2                  | 8.8                   | 90.6                  |
| aug-pc-0  | 1.2      | 12.9                  | 32.8                  | 7.9                   | 54.0                  | 1.1      | 13.8                  | 30.8                  | 8.6                   | 60.5                  |
| pcS-0     | 1.5      | 9.3                   | 67.0                  | 6.8                   | 88.9                  | 1.5      | 8.1                   | 65.0                  | 8.0                   | 99.8                  |
| aug-pcS-0 | 1.1      | 2.8                   | 12.5                  | 3.5                   | 61.2                  | 1.1      | 3.0                   | 14.5                  | 3.8                   | 66.7                  |
| pc-1      | 0.38     | 2.8                   | 18.0                  | 2.7                   | 18.3                  | 0.37     | 2.8                   | 17.8                  | 2.9                   | 19.7                  |
| aug-pc-1  | 0.25     | 2.9                   | 15.3                  | 2.5                   | 16.8                  | 0.25     | 2.9                   | 15.9                  | 2.6                   | 19.3                  |
| pcS-1     | 0.25     | 1.1                   | 7.9                   | 2.6                   | 19.1                  | 0.24     | 1.2                   | 7.7                   | 2.8                   | 20.4                  |
| aug-pcS-1 | 0.11     | 1.0                   | 2.8                   | 1.8                   | 15.4                  | 0.11     | 1.1                   | 3.3                   | 1.8                   | 17.6                  |
| pc-2      | 0.14     | 0.94                  | 4.27                  | 0.95                  | 6.65                  | 0.14     | 0.94                  | 3.92                  | 1.05                  | 7.19                  |
| aug-pc-2  | 0.10     | 0.87                  | 2.65                  | 1.17                  | 3.96                  | 0.10     | 0.95                  | 2.83                  | 1.17                  | 3.88                  |
| pcS-2     | 0.073    | 0.50                  | 2.88                  | 0.75                  | 6.50                  | 0.074    | 0.45                  | 2.28                  | 0.84                  | 6.92                  |
| aug-pcS-2 | 0.032    | 0.39                  | 0.61                  | 0.81                  | 3.65                  | 0.034    | 0.42                  | 0.55                  | 0.79                  | 3.59                  |
| pc-3      | 0.016    | 0.16                  | 0.49                  | 0.19                  | 0.95                  | 0.017    | 0.16                  | 0.40                  | 0.19                  | 0.85                  |
| aug-pc-3  | 0.013    | 0.16                  | 0.25                  | 0.19                  | 0.44                  | 0.014    | 0.16                  | 0.29                  | 0.18                  | 0.43                  |
| pcS-3     | 0.008    | 0.07                  | 0.32                  | 0.18                  | 0.96                  | 0.008    | 0.07                  | 0.23                  | 0.18                  | 0.84                  |
| aug-pcS-3 | 0.005    | 0.06                  | 0.05                  | 0.16                  | 0.41                  | 0.004    | 0.07                  | 0.05                  | 0.15                  | 0.39                  |
| pc-4      | 0.001    | 0.052                 | 0.108                 | 0.023                 | 0.118                 | 0.001    | 0.048                 | 0.087                 | 0.017                 | 0.110                 |
| pcS-4     | 0.001    | 0.005                 | 0.059                 | 0.020                 | 0.121                 | 0.000    | 0.005                 | 0.036                 | 0.013                 | 0.104                 |

<sup>a</sup> All results have been generated using completely uncontracted basis sets. The basis set limits have been taken as the aug-pcS-4 results. **KT3** and **B3LYP** denote the employed exchange-correlation functionals. **H** indicates hydrogen shielding constants (75 data points), **M**<sub>1</sub> indicates Li and Be shielding constants (10 data points), **A**<sub>1</sub> indicates C, N, O, and F shielding constants (92 data points), **M**<sub>2</sub> indicates Na and Mg shielding constants (10 data points), and **A**<sub>2</sub> indicates Si, P, S, and Cl shielding constants (32 data points).

The results in Table 5 have been grouped according to a qualitative classification corresponding to subdouble-, double-, triple-, quadruple-, and pentuple- $\zeta$  quality, which in most cases also indicates the highest angular momentum functions included in the basis sets (Table 2). From a computational point of view the total number of (contracted) functions is an important factor, and we have in Table 5 included the average number of basis functions per atom ( $N_{\text{basis}}$ ) over the whole data set in Table 3 as an indicator of the computational cost. Since addition of both diffuse and tight functions rapidly increases the size of a basis set, the highest angular momentum functions included are not necessarily good indicators of the basis set size. The aug-cc-pCVTZ basis set, for example, contains more functions than the cc-pVQZ basis set, despite the latter formally being of higher  $\zeta$  quality.

The unpolarized pcS-0 and aug-pcS-0 basis sets are not expected to be able to generate useful results, and basis sets of double- $\zeta$  quality are normally considered as the first level where semiquantitative results can be expected, with the 6-31G(d,p) basis set being widely used for routine applications. The performance of the 6-31G(d,p), 6-31++G(d,p), cc-pVDZ, and aug-pVDZ basis sets are very similar, with typical deviations of  $\sim 20$  and  $\sim 30$  ppm for first- and second-row elements, respectively. Addition of tight functions (aug-cc-pCVDZ) reduces the errors slightly but deteriorates the results for hydrogen. The Ahlrichs SVP basis set has somewhat larger errors. The aug-pcS-1 basis set clearly outperforms these standard basis sets and reduces the average error to  $\sim 3$  ppm for first-row elements. The results for second-row elements are also improved to a value of  $\sim 15$  ppm. The IGLO-II results formally compare favorably with the pcS-1 results, but it should be noted that the former results do not include the metallic systems, where many of the large deviations are observed. Without these systems, the MAD

value of 8.2 ppm for the **A**<sub>1</sub> systems is reduced to 5.1 ppm, which can be compared with the IGLO-II value of 8.8 ppm.

At the triple- $\zeta$  level, the standard 6-311G(2df,2pd), 6-311++G(2df,2pd), cc-pVTZ, aug-cc-pVTZ, and Ahlrichs TZV basis sets perform roughly at par, with typical errors of  $\sim 0.1$  ppm for hydrogen and  $\sim 10$  ppm for the nonmetal atoms. The aug-pcS-2 basis set for comparison has errors of 0.03 ppm for hydrogen and less than 2 and 4 ppm for first- and second-row elements, respectively. The aug-cc-pCVTZ basis set improves the aug-cc-pVTZ results, showing that tight functions are important but also increases the size of the basis set by nine functions per atom on the average. Despite having significantly fewer functions, the aug-pcS-2 basis set performs better than the aug-cc-pCVTZ basis set. The IGLO-III basis set belongs to the triple- $\zeta$  family in terms of number of functions per atom but displays larger errors than the aug-pcS-2 basis set, despite the fact that the results do not include some of the more difficult systems.

The cc-pVQZ, aug-cc-pVQZ, and aug-cc-pCVQZ results are only marginally improved over those for the corresponding triple- $\zeta$  basis sets, and the Ahlrichs QZV basis set also only provides a small reduction in the basis set error compared to the TVZ basis set. In contrast, aug-pcS-3 reduces the basis set error by almost an order of magnitude relative to aug-pcS-2, and the mean error is now below 0.002 ppm for hydrogen and below 0.5 ppm for all the remaining elements.

At the pentuple- $\zeta$  level, the cc-pV5Z basis set only provides a minor improvement relative to the cc-pVQZ results, and there are still errors of  $\sim 15$ – $20$  ppm for the Na and Mg systems. The pcS-4 basis set, on the other hand, reduces the basis set errors to below 0.001 ppm for hydrogen and below  $\sim 0.1$  ppm for all the remaining elements.

The grouping in Table 5 displays a significant variation in the number of basis functions per atom within each quality

**Table 5.** Mean Absolute Deviations (ppm) Relative to the Basis Set Limit for the Symmetry-Unique Nuclear Magnetic Shielding Constants for the Systems in Table 3<sup>a</sup>

| basis set         | $\langle N_{\text{basis}} \rangle$ | KT3      |                       |                       |                       |                       | B3LYP    |                       |                       |                       |                       |
|-------------------|------------------------------------|----------|-----------------------|-----------------------|-----------------------|-----------------------|----------|-----------------------|-----------------------|-----------------------|-----------------------|
|                   |                                    | <i>H</i> | <i>M</i> <sub>1</sub> | <i>A</i> <sub>1</sub> | <i>M</i> <sub>2</sub> | <i>A</i> <sub>2</sub> | <i>H</i> | <i>M</i> <sub>1</sub> | <i>A</i> <sub>1</sub> | <i>M</i> <sub>2</sub> | <i>A</i> <sub>2</sub> |
| pcS-0             | 6.4                                | 1.5      | 10.8                  | 63.9                  | 6.6                   | 92.9                  | 1.5      | 10.4                  | 63.6                  | 9.9                   | 100.0                 |
| aug-pcS-0         | 9.1                                | 1.1      | 7.6                   | 14.9                  | 15.3                  | 63.1                  | 1.1      | 8.4                   | 16.2                  | 24.5                  | 67.0                  |
| STO-3G            | 3.9                                | 2.4      | 5.0                   | 88.2                  | 19.5                  | 194.1                 | 2.4      | 5.9                   | 96.9                  | 23.2                  | 224.8                 |
| pcS-1             | 12.1                               | 0.20     | 1.4                   | 8.2                   | 2.2                   | 19.5                  | 0.21     | 4.0                   | 8.4                   | 2.6                   | 21.8                  |
| aug-pcS-1         | 18.6                               | 0.08     | 1.4                   | 3.3                   | 1.3                   | 15.9                  | 0.09     | 3.2                   | 3.9                   | 2.9                   | 21.0                  |
| cc-pVDZ           | 11.4                               | 0.35     | 5.6                   | 19.7                  | 20.5                  | 34.1                  | 0.35     | 8.4                   | 26.1                  | 27.6                  | 58.5                  |
| aug-cc-pVDZ       | 18.2                               | 0.19     | 4.3                   | 18.6                  | 19.3                  | 29.4                  | 0.19     | 5.8                   | 24.8                  | 31.7                  | 55.8                  |
| aug-cc-pCVDZ      | 20.7                               | 0.22     | 3.4                   | 12.8                  | 14.4                  | 18.3                  | 0.20     | 3.8                   | 17.7                  | 20.7                  | 25.4                  |
| 6-31G(d,p)        | 11.3                               | 0.47     | 5.0                   | 23.8                  | 16.9                  | 30.4                  | 0.43     | 7.4                   | 29.5                  | 24.1                  | 49.7                  |
| 6-31++G(d,p)      | 14.0                               | 0.25     | 4.1                   | 22.2                  | 19.5                  | 30.6                  | 0.23     | 6.3                   | 28.6                  | 31.3                  | 49.6                  |
| SVP               | 10.2                               | 0.45     | 6.5                   | 32.5                  | 13.4                  | 47.6                  | 0.47     | 8.4                   | 33.0                  | 21.6                  | 73.0                  |
| IGLO-II           | 16.5                               |          |                       | (8.8)                 |                       | (13.1)                |          |                       | (6.5)                 |                       | (14.0)                |
| pcS-2             | 25.0                               | 0.07     | 0.7                   | 3.5                   | 1.0                   | 6.3                   | 0.07     | 2.2                   | 2.3                   | 2.1                   | 6.8                   |
| aug-pcS-2         | 37.5                               | 0.02     | 0.7                   | 1.7                   | 0.8                   | 3.3                   | 0.03     | 0.8                   | 0.9                   | 2.1                   | 4.1                   |
| cc-pVTZ           | 24.4                               | 0.19     | 2.2                   | 10.5                  | 11.2                  | 20.5                  | 0.19     | 2.5                   | 9.5                   | 17.9                  | 26.1                  |
| aug-cc-pVTZ       | 37.3                               | 0.10     | 2.1                   | 10.6                  | 15.7                  | 18.6                  | 0.11     | 2.1                   | 9.5                   | 13.5                  | 22.1                  |
| aug-cc-pCVTZ      | 46.0                               | 0.11     | 0.8                   | 3.8                   | 4.7                   | 3.3                   | 0.11     | 1.0                   | 4.9                   | 4.6                   | 3.8                   |
| 6-311G(2df,2pd)   | 24.4                               | 0.20     | 2.2                   | 10.3                  | 0.8                   | 11.4                  | 0.20     | 2.2                   | 8.8                   | 1.1                   | 14.6                  |
| 6-311++G(2df,2pd) | 27.1                               | 0.10     | 2.2                   | 9.4                   | 0.9                   | 5.4                   | 0.11     | 2.2                   | 7.8                   | 1.6                   | 8.2                   |
| TVZ               | 23.1                               | 0.16     | 3.5                   | 9.1                   | 5.3                   | 9.2                   | 0.15     | 3.7                   | 8.2                   | 2.2                   | 14.9                  |
| IGLO-III          | 24.9                               |          |                       | (1.4)                 |                       | (12.4)                |          |                       | (1.8)                 |                       | (13.5)                |
| pcS-3             | 54.7                               | 0.004    | 0.05                  | 0.37                  | 0.22                  | 0.97                  | 0.005    | 0.05                  | 0.22                  | 0.43                  | 1.02                  |
| aug-pcS-3         | 75.2                               | 0.002    | 0.04                  | 0.12                  | 0.12                  | 0.45                  | 0.002    | 0.04                  | 0.06                  | 0.39                  | 0.65                  |
| cc-pVQZ           | 45.5                               | 0.083    | 0.74                  | 7.31                  | 13.30                 | 26.90                 | 0.082    | 0.42                  | 4.23                  | 15.43                 | 35.99                 |
| aug-cc-pVQZ       | 66.6                               | 0.051    | 0.73                  | 6.71                  | 18.12                 | 25.41                 | 0.052    | 0.41                  | 3.81                  | 5.92                  | 33.41                 |
| aug-cc-pCVQZ      | 85.8                               | 0.046    | 0.21                  | 1.67                  | 1.97                  | 1.29                  | 0.048    | 0.22                  | 1.68                  | 1.72                  | 0.52                  |
| QVZ               | 45.9                               | 0.065    | 1.08                  | 5.92                  | 3.08                  | 3.33                  | 0.066    | 1.28                  | 3.39                  | 0.79                  | 3.20                  |
| pcS-4             | 92.3                               | 0.001    | 0.01                  | 0.08                  | 0.03                  | 0.15                  | 0.001    | 0.02                  | 0.04                  | 0.21                  | 0.13                  |
| cc-pV5Z           | 76.7                               | 0.030    | 0.30                  | 3.79                  | 19.24                 | 3.62                  | 0.029    | 0.05                  | 1.20                  | 15.75                 | 2.73                  |

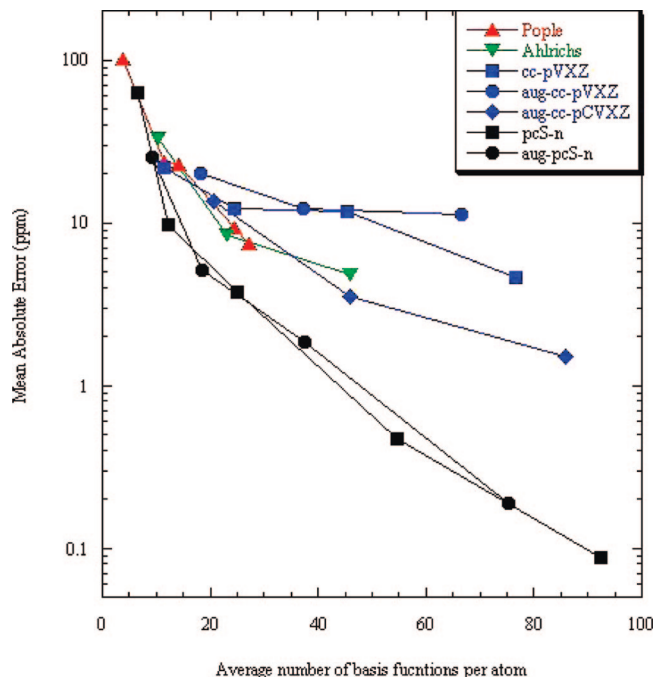
<sup>a</sup> All results using contracted basis sets. The basis set limit has been taken as the uncontracted aug-pcS-4 results. The cc-pVXZ and aug-cc-pVXZ basis sets include an additional tight d function for the elements Si–Cl.  $\langle N_{\text{basis}} \rangle$  denotes the average number of basis functions per atom for the whole data set. KT3 and B3LYP denote the employed exchange-correlation functionals. *H* indicates hydrogen shielding constants (75 data points), *M*<sub>1</sub> indicates Li and Be shielding constants (10 data points), *A*<sub>1</sub> indicates C, N, O, and F shielding constants (92 data points), *M*<sub>2</sub> indicates Na and Mg shielding constants (10 data points), and *A*<sub>2</sub> indicates Si, P, S, and Cl shielding constants (32 data points). The IGLO basis sets are not defined for s-group elements, and the values in parentheses are for only 82 (*A*<sub>1</sub>) and 26 (*A*<sub>2</sub>) data points.

level: the aug-cc-pCVQZ basis set, for example, has 10 more basis functions than the aug-pcS-3 basis set, despite both being of quadruple- $\zeta$  quality augmented with tight and diffuse functions. In order to provide an alternative comparison, we have displayed the mean average deviation for all the non-hydrogen shielding constants as a function of the average number of functions per atom in Figure 1. The pcS-*n* and aug-pcS-*n* families of basis sets clearly display a smooth, controlled, and exponential convergence toward the limiting value. The cc-pVXZ and aug-cc-pVXZ basis sets display little convergence as the basis set is enlarged and have problems reducing the average error below 10 ppm. Augmenting with tight functions (aug-cc-pCVXZ) improves the results, but at a high computational cost, as many tight functions are added. It can be noted that the aug-cc-pCVTZ basis set has a much better performance than the cc-pVQZ set, despite the two basis sets being of almost the same size. The Ahlrichs- and Pople-type basis sets perform roughly at par with the cc-pVXZ basis sets, and augmenting the Pople basis sets with diffuse functions has only a small influence. Considering Figure 1, we find it significant that the aug-pcS-1 basis set, which is only marginally larger in size than

the very popular 6-31++G(d,p), has basis set errors that are almost an order of magnitude smaller. Similarly, the aug-pcS-2 basis set, which is similar in size to the 6-311++G(2df,2pd) set, reduces the basis set errors by almost an order of magnitude.

When the MAD value is displayed as a function of the average number of functions per atom as in Figure 1, the convergence of the pcS-*n* and aug-pcS-*n* curves is seen to be very similar, with the aug-pcS-*n* results for a given *n* being of intermediate quality compared to the corresponding pcS-*n* and pcS-(*n*+1) results. When viewed in this fashion, the effect of augmenting the pcS-*n* basis sets with diffuse functions can be considered as simply being the results of having a more complete basis set. It can also be noted that the error reduction by including diffuse functions in Table 5 mainly arises from the polar systems in Table 3. For the nonpolar systems, which include a large fraction of typical organic molecules, the pcS-*n* basis sets provide results of quality similar to that for the aug-pcS-*n* basis set.

When statistical methods are used for evaluating the performance of various methods and basis sets, there is always a risk of biasing the results by the selection of



**Figure 1.** Mean absolute deviation relative to the basis set limit of nuclear magnetic shielding constants (ppm) for all non-hydrogen atoms as a function of the average number of basis functions per atom.

compounds in the test set. The systems in Table 3 were selected to represent a variety of structural elements found in many applications and covering a range of molecular bonding. In our initial selection of systems, we also had included the compounds BeO, BeS, MgO, and MgS. These systems, however, turned out to display pathological behaviors with respect to basis set convergence, and the errors were so large that they would have completely dominated the statistical measure in Tables 4 and 5 had they been included. The worst of these cases is the MgO system, for which the calculated shielding constants are shown in Table 6. The oxygen shielding constant with the B3LYP functional is calculated to be  $-2440$  ppm with the (uncontracted) aug-pcS-4 basis set, and this value is presumably converged to within  $\sim 10$  ppm, as judged from the aug-pcS-3 result. Using the  $-2440$  ppm value as the reference, it is seen that all the basis sets of double- $\zeta$  quality have errors measured in the hundreds or thousands of ppm, and the SVP basis set marks a spectacular failure with a calculated value of  $+23\,602$  ppm. Part of this discrepancy is due to the fact that this basis set does not have d-type functions on Mg, but the pcS-1 result of  $-6031$  ppm shows that this is not the only reason, as both these basis sets have the same angular momentum functions (Table 2). Basis sets of triple- $\zeta$  quality have typical errors of  $\sim 300$  ppm, while augmentation with diffuse functions reduces the error to  $\sim 150$  ppm. Only at the quadruple- $\zeta$  level, preferably augmented with both tight and diffuse functions, does the error drop to acceptable levels. For this specific system, inclusion of both d- and f-type functions on both atoms, as well as diffuse functions, is required to produce a qualitatively correct description. The aug-pcS- $n$  basis sets here provide less accurate results than the aug-cc-pCVXZ basis set at the same  $\zeta$  level. This is due

**Table 6.** Nuclear Magnetic Shielding Constants (ppm) for MgO

| basis set         | KT3  |       | B3LYP |        |
|-------------------|------|-------|-------|--------|
|                   | Mg   | O     | Mg    | O      |
| STO-3G            | 1686 | -2727 | 4674  | -13630 |
| 6-31G(d,p)        | 890  | -1315 | 1046  | -2966  |
| 6-31++G(d,p)      | 901  | -1420 | 1018  | -2610  |
| 6-311G(2df,2pd)   | 915  | -1368 | 1046  | -2766  |
| 6-311++G(2df,2pd) | 918  | -1359 | 1028  | -2572  |
| SVP               | 1990 | -8587 | -2275 | 23602  |
| TVZ               | 915  | -1356 | 1053  | -2733  |
| QVZ               | 908  | -1344 | 1004  | -2460  |
| cc-pVDZ           | 880  | -1310 | 1044  | -3090  |
| cc-pVTZ           | 887  | -1372 | 1018  | -2876  |
| cc-pVQZ           | 892  | -1402 | 996   | -2712  |
| cc-pV5Z           | 914  | -1349 | 999   | -2533  |
| aug-cc-pVDZ       | 898  | -1488 | 1029  | -2889  |
| aug-cc-pVTZ       | 898  | -1385 | 992   | -2586  |
| aug-cc-pVQZ       | 888  | -1385 | 980   | -2591  |
| aug-cc-pCVDZ      | 904  | -1456 | 1026  | -2773  |
| aug-cc-pCVTZ      | 901  | -1339 | 993   | -2427  |
| aug-cc-pCVQZ      | 903  | -1335 | 996   | -2424  |
| pcS-0             | 1308 | -3862 | 7336  | -51269 |
| pcS-1             | 1100 | -2567 | 1484  | -6031  |
| pcS-2             | 921  | -1436 | 1041  | -2734  |
| pcS-3             | 910  | -1361 | 1008  | -2477  |
| pcS-4             | 906  | -1338 | 1002  | -2440  |
| aug-pcS-0         | 916  | -1476 | 1056  | -2999  |
| aug-pcS-1         | 937  | -1525 | 1073  | -2944  |
| aug-pcS-2         | 922  | -1439 | 1033  | -2660  |
| aug-pcS-3         | 909  | -1354 | 1007  | -2466  |
| aug-pcS-4         | 907  | -1343 | 1002  | -2440  |

to differences in how the higher angular momentum functions are included for Mg. The cc-pVXZ basis sets include d functions for Mg at the DZ level and f functions at the TZ level, analogous to the case for the p-block elements. The pcS- $n$  basis sets, in contrast, only include p-type polarization functions at the DZ level (pcS-1) for s-block elements such as Mg and only up to d functions at the TZ level (pcS-2). It is therefore necessary to go to the pcS-3 level to obtain a qualitatively correct result.

The B3LYP oxygen shielding constant is the most sensitive to the quality of the basis set, but the same trend is seen for the magnesium atom. The KT3 functional provides similar trends, although the changes with respect to basis sets are less dramatic. The calculated shielding constants at the basis set limit differ by  $\sim 1000$  ppm for oxygen and by  $\sim 100$  ppm for magnesium between the two functionals, indicating the importance of selecting a suitable exchange-correlation functional. While double- or triple- $\zeta$  quality basis sets will be sufficient for the large majority of routine applications, it is in our opinion valuable to have a well-defined hierarchy of basis sets for systematically approaching the basis set limit for problematic cases, as for example MgO. It is gratifying to see that both the pcS- $n$  and aug-pcS- $n$  basis sets display a monotonic convergence toward the limiting value for this difficult system.

## V. Summary

On the basis of our previous analysis for nuclear spin-spin coupling constants, we show that an improved basis set convergence for nuclear magnetic shielding constants can be obtained by addition of a single tight p-type basis function.



When used in combination with the previously proposed polarization consistent basis sets, this leads to the definition of a hierarchy of basis sets denoted pcS-*n*. An evaluation of the performance for a selection of typical molecular systems shows that these new basis sets represent an improvement with respect to reducing basis set errors relative to existing basis sets. A typical error at the aug-pcS-1 (double- $\zeta$ ) level is 5 ppm for non-hydrogen atoms, which is reduced to 2 ppm upon going to the aug-pcS-2 (triple- $\zeta$ ) level. The nuclear shielding constant for hydrogen displays a much smaller basis set effect, with typical errors of 0.1 and 0.03 ppm at the aug-pcS-1 and aug-pcS-2 levels.

Basis set limitations are only one possible error component in a comparison with experimental values, as the reference geometry, vibrational averaging, solvent effects,<sup>37</sup> and inadequacies in the exchange-correlation functional will need to be addressed in order to provide a direct comparison with experiments. The present pcS-*n* basis sets, however, should be suitable for controlling the basis set error, and the pcS-1 and pcS-2 basis sets should be suitable for Hartree–Fock and density functional methods in general and allow calculations for large systems.

The present work adds yet another sequence of basis sets to an already large variety, and it is reasonable to ask whether this represents an improvement of the computational capabilities or only serves to further complicate the selection of a basis set for a given problem. Not surprisingly, we favor the first option. Modern computational chemistry should in our opinion be able to control and assess the errors in the calculated results. An essential component for this is a well-defined hierarchy of basis sets which approaches the basis set limiting value in a smooth fashion and preferably is available for a reasonable selection of elements. Basis sets such as 6-31G(d,p) do not have a clear protocol for systematic improvements and must therefore be used as part of a precalibrated procedure, where the error evaluation is done by comparison with external reference data. Such an approach becomes problematic on encountering pathological cases, as illustrated by the MgO system in the present case, and for systems where no experimental data are available for calibration. The availability of a hierarchy of basis sets allows identification of pathological systems and provides the possibility of controlling the basis set errors, albeit at an increased computational cost.

The use of basis sets designed for specific properties has a long history,<sup>38</sup> but the pcJ-*n* and pcS-*n* basis sets in our opinion are the first to allow a systematic and fast convergence toward the basis set limiting value for nuclear magnetic properties. When property-specific basis sets are discussed, it should be recognized that basis sets are always a compromise between accuracy and computational efficiency. A basis set suitable for calculating a range of properties accurately will be so large that it is not computationally efficient. From an application point of view, the interest is usually in a single or narrow range of molecular properties, and having computationally efficient basis sets is necessary for tackling many real-world problems having a large number of atoms. The present pcS-*n* basis sets can be considered as a subset of the pcJ-*n* basis sets, where it is seen that nuclear

magnetic shielding constants do not need as many tight functions as spin–spin coupling constants and can be contracted substantially harder without losing accuracy, thereby improving the computational efficiency. As such, we feel that the pcS-*n* basis sets should be a useful addition to the field of computational chemistry.

**Acknowledgment.** This work was supported by grants from the Danish Center for Scientific Computation and the Danish Natural Science Research Council.

**Supporting Information Available:** Tables giving exponents and contraction coefficients for the pcS-*n* and aug-pcS-*n* basis sets for the elements H–Ar. The information is available free of charge via the Internet at <http://pubs.acs.org>.

## References

- (1) Lutnaes, O. B.; Ruden, T. A.; Helgaker, T. *Magn. Reson. Chem.* **2004**, *42*, S117. Helgaker, T.; Lutnaes, O. B.; Jaszunski, M. *J. Chem. Theory Comput.* **2007**, *3*, 86.
- (2) Gauss, J.; Stanton, J. F. *J. Chem. Phys.* **1995**, *103*, 3561. Gauss, J.; Stanton, J. F. *J. Chem. Phys.* **1996**, *104*, 2574. Stanton, J. F.; Gauss, J.; Siehl, H.-S. *Chem. Phys. Lett.* **1996**, *262*, 183.
- (3) Parr, R. G.; Yang, W. *Density Functional Theory*; Oxford University Press: Oxford, U.K., 1989. Koch, W.; Holthausen, M. C. *A Chemist's Guide to Density Functional Theory*; Wiley-VCH: Weinheim, Germany, 2000.
- (4) Ochsenfeld, C.; Kussmann, J.; Koziol, F. *Angew. Chem., Int. Ed.* **2004**, *43*, 4485. Schaller, T.; Buechele, U. P.; Klaerner, F.-G.; Blaeser, D.; Boese, R.; Brown, S. P.; Spiess, H. W.; Koziol, F.; Kussmann, J.; Ochsenfeld, C. *J. Am. Chem. Soc.* **2007**, *129*, 1293. Zienau, J.; Kussmann, J.; Koziol, F.; Ochsenfeld, C. *Phys. Chem. Chem. Phys.* **2007**, *9*, 4552.
- (5) Keal, T. W.; Tozer, D. *J. Chem. Phys.* **2005**, *121*, 5654.
- (6) Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007. Wilson, A. K.; van Mourik, T., Jr. *J. Mol. Struct.* **1996**, *388*, 339.
- (7) Jensen, F. *J. Chem. Phys.* **2001**, *115*, 9113; **2002**, *116*, 3502; **2002**, *116*, 7372. Jensen, F.; Helgaker, T. *J. Chem. Phys.* **2004**, *121*, 3462.
- (8) Kutzelnigg, W.; Morgan, J. D. *J. Chem. Phys.* **1992**, *96*, 4484; **1992**, *97*, 8821. (E).
- (9) Klopper, W.; Kutzelnigg, W. *J. Mol. Struct.* **1986**, *135*, 339. Christensen, K. A.; Jensen, F. *Chem. Phys. Lett.* **2000**, *317*, 400. Jensen, F. *Theor. Chem. Acc.* **2000**, *104*, 484. Schwenke, D. W. *J. Chem. Phys.* **2005**, *122*, 014107.
- (10) Jensen, F. *J. Chem. Phys.* **2003**, *118*, 2459.
- (11) Kendall, R. A.; Dunning, T. H., Jr.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796.
- (12) Jensen, F. *J. Chem. Phys.* **2002**, *117*, 9234.
- (13) Jensen, F. *J. Chem. Theory Comput.* **2006**, *2*, 1360.
- (14) Helgaker, T.; Jaszunski, M.; Ruud, K. *Chem. Rev.* **1999**, *99*, 293. Karadakov, P. B. *Mod. Magn. Reson.* **2006**, *1*, 59.
- (15) Vaara, J. *Phys. Chem. Chem. Phys.* **2007**, *9*, 5399.
- (16) Cheeseman, J. R.; Trucks, G. W.; Keith, T. A.; Frisch, M. J. *J. Chem. Phys.* **1996**, *104*, 5497. Zhang, Y.; Wu, A.; Xu, X.; Yan, Y. *J. Phys. Chem. A* **2007**, *111*, 9431. Wu, A.; Zhang, Y.; Xu, X.; Yan, Y. *J. Comput. Chem.* **2007**, *28*, 2431. d'Antuono, P.; Botek, E.; Champagne, B.; Spassova, M.;



- Denkova, P. *J. Chem. Phys.* **2006**, *125*, 144309. Chesnut, D. B. *Chem. Phys. Lett.* **2003**, *380*, 251.
- (17) Wu, A.; Cremer, D.; Gauss, J. *J. Phys. Chem. A* **2003**, *107*, 8737. Magyarfalvi, G.; Pulay, P. *J. Chem. Phys.* **2003**, *119*, 1350. Auer, A. A.; Gauss, J.; Stanton, J. F. *J. Chem. Phys.* **2003**, *118*, 10407.
- (18) Sefzik, T. H.; Turco, D.; Iuliucci, R. J.; Facelli, J. C. *J. Phys. Chem. A* **2005**, *109*, 1180. Jaszunski, M. *Chem. Phys. Lett.* **2004**, *385*, 122. Moon, S.; Case, D. A. *J. Comput. Chem.* **2006**, *27*, 825. Kupka, T.; Ruscic, B.; Botto, R. E. *J. Phys. Chem. A* **2002**, *106*, 10396. Kupka, T.; Ruscic, B.; Botto, R. E. *Solid State Nucl. Magn. Reson.* **2003**, *23*, 145.
- (19) Kupka, T.; Lim, C. *J. Phys. Chem. A* **2007**, *111*, 1927.
- (20) Schindler, M.; Kutzelnigg, W. *J. Chem. Phys.* **1982**, *76*, 1919. Kutzelnigg, W.; Fleischer, U.; Schindler, M. *The IGLO-Method: Ab Initio Calculation Interpretation of NMR Chemical Shifts Magnetic Susceptibilities*; Springer-Verlag: Heidelberg, Germany, 1990; 23.
- (21) Manninen, P.; Vaara, J. *J. Comput. Chem.* **2006**, *27*, 434.
- (22) Helgaker, T.; Jensen, H. J. Aa.; Jørgensen, P.; Olsen, J.; Ruud, K.; Ågren, H.; Auer, A. A.; Bak, K. L.; Bakken, V.; Christiansen, O.; Coriani, S.; Dahle, P.; Dalskov, E. K.; Enevoldsen, T.; Fernez, B.; Hättig, C.; Hald, K.; Halkier, A.; Heiberg, H.; Hettema, H.; Jonsson, D.; Kirpekar, S.; Kobayashi, R.; Koch, H.; Mikkelsen, K. V.; Norman, P.; Packer, M. J.; Pedersen, T. B.; Ruden, T. A.; Sanchez, A.; Saue, T.; Sauer, S. P. A.; Schimmelpfening, B.; Sylvester-Hvid, K. O.; Taylor, P. R.; Vahtras O., DALTON, a Molecular Electronic Structure Program, Release 2.0, 2005.
- (23) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; 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.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*; Gaussian, Inc., Wallingford, CT, 2004.
- (24) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648. Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623.
- (25) London, F. *J. Phys. Radium, Paris* **1937**, *8*, 397. McWeeny, R. *Phys. Rev.* **1962**, *126*, 1028. Ditchfield, R. *Mol. Phys.* **1974**, *27*, 789. Wolinski, K.; Hilton, J. F.; Pulay, P. *J. Am. Chem. Soc.* **1990**, *112*, 8251.
- (26) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Rassolov, V.; Pople, J. A. *J. Chem. Phys.* **1998**, *109*, 764. Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Pople, J. A. *J. Chem. Phys.* **2000**, *112*, 7374.
- (27) Hehre, W. J.; Ditchfield, R.; Stewart, R. F.; Pople, J. A. *J. Chem. Phys.* **1970**, *52*, 2769.
- (28) Francl, M. M.; Pietro, W. J.; Hehre, W. J.; Binkley, J. S.; Gordon, M. S.; DeFrees, D. J.; Pople, J. A. *J. Chem. Phys.* **1982**, *77*, 3654.
- (29) Frisch, M. J.; Pople, J. A.; Binkley, J. S. *J. Chem. Phys.* **1984**, *80*, 3265. Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *Theor. Chem. Acc.* **1980**, *72*, 650.
- (30) McLean, A. D.; Chandler, G. S. *J. Chem. Phys.* **1980**, *72*, 5639.
- (31) Woon, D. E.; Dunning, T. H. *J. Chem. Phys.* **1993**, *98*, 1358. Dunning, T. H., Jr.; Peterson, K. A.; Wilson, A. K. *J. Chem. Phys.* **2001**, *114*, 9244.
- (32) Kendall, R. A., Jr.; Harrison, J. R. *J. Chem. Phys.* **1992**, *96*, 6796.
- (33) Woon, D. E. *J. Chem. Phys.* **1995**, *103*, 4572.
- (34) Schafer, A.; Horn, H.; Ahlrichs, R. *J. Chem. Phys.* **1992**, *97*, 2571.
- (35) Schafer, A.; Huber, C.; Ahlrichs, R. *J. Chem. Phys.* **1994**, *100*, 5829.
- (36) Weigend, F.; Furche, F.; Ahlrichs, R. *J. Chem. Phys.* **2003**, *119*, 12753.
- (37) Cybulski, H.; Sadley, J. *J. Chem. Phys.* **2006**, *323*, 218. Aidas, K.; Moegelhoj, A.; Kjaer, H.; Nielsen, C. B.; Mikkelsen, K. V.; Ruud, K.; Christiansen, O.; Kongsted, J. *J. Phys. Chem. A* **2007**, *111*, 4199.
- (38) Sadlej, A. J. *J. Chem. Phys. Lett.* **1977**, *47*, 50. Roos, B. O.; Sadlej, A. J. *J. Chem. Phys.* **1985**, *94*, 43. Wolinski, K.; Roos, B. O.; Sadlej, A. J. *Theor. Chim. Acta* **1985**, *68*, 431. Sadlej, A. J. *Collect. Czech. Chem. Commun.* **1988**, *53*, 1995. Benkova, Z.; Sadlej, A. J.; Oakes, R. E.; Bell, S. E. *J. Comput. Chem.* **2005**, *26*, 145. Baranowska, A.; Siedlecka, M.; Sadlej, A. J. *Theor. Chem. Acc.* **2007**, *118*, 959.

CT800013Z