

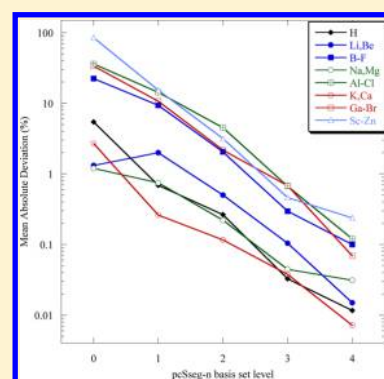
Segmented Contracted Basis Sets Optimized for Nuclear Magnetic Shielding

Frank Jensen*

Department of Chemistry, Aarhus University, DK-8000 Aarhus, Denmark

S Supporting Information

ABSTRACT: A family of segmented contracted basis sets is proposed, denoted pcSseg-*n*, which are optimized for calculating nuclear magnetic shielding constants. For the elements H–Ar, these are computationally more efficient than the previously proposed general contracted pcS-*n* basis sets, and the new basis sets are extended to also include the elements K–Kr. The pcSseg-*n* basis sets are optimized at the density functional level of theory, but it has been shown previously that these property-optimized basis sets are also suitable for calculating shielding constants with correlated wave function methods. The pcSseg-*n* basis sets are available in qualities ranging from (unpolarized) double- ζ to pentuple- ζ quality and should be suitable for both routine and benchmark calculations of nuclear magnetic shielding constants. The ability to rigorously separate basis set and method errors should aid in developing more accurate methods.



INTRODUCTION

Gaussian basis sets have been the *de facto* standard for molecular electronic structure calculations for the last several decades. Several different basis sets are available, and two recent reviews can be consulted for an overview.^{1,2} The modern view is that basis sets should be available in several well-defined quality levels such that the basis set error can be quantified and controlled, and in favorable cases allow the results to be extrapolated to the complete basis set (CBS) limit. The basis set quality can conveniently be classified in terms of the highest angular momentum function included, and this gives rise to the familiar double/triple/quadruple/etc.-zeta (ζ) notation.³ Different basis sets have different performance for different methods; for example, independent particle models (Hartree–Fock (HF) and Density Functional Theory (DFT)) have different basis set requirements than wave function electron correlation models. It is consequently advantageous to employ basis sets that have been specifically optimized for a given method in order to achieve a fast and robust basis set convergence.

The large majority of commonly employed basis sets have been designed by minimizing the basis set parameters with respect to atomic and/or molecular energies. This ensures that the basis sets provide a good representation of the orbitals in the energetically important region in terms of the distance to the nuclei. Molecular properties that depend directly on the energy, such as dissociation energies, equilibrium geometries (first derivatives with respect to nuclear coordinates), and harmonic vibrational frequencies (second derivatives with respect to nuclear coordinates), are similarly well described. Other molecular properties, however, may be less well described, as they may depend on an accurate description of the orbitals in energetically unimportant regions. Electric properties like dipole and higher order multipole moments,

and polarizabilities and hyperpolarizabilities, depend critically on the region far from the nuclei (wave function tail). The basis set convergence can be substantially improved by addition of (diffuse) basis functions with small exponents in order to describe the wave function tail region, and this equivalently ensures that the basis set error at a given ζ level is reduced. Other molecular properties, like nuclear magnetic shielding and spin–spin coupling constants, depend on a good representation of the orbitals near the nucleus, and this requires addition of (tight) basis functions with large exponents in order to describe the wave function core region. Core properties furthermore often require uncontraction of the basis sets relative to the energy-optimized ones in order to provide sufficient flexibility in describing the core changes caused by molecular bonding differences.

Analogous to the employment of basis sets optimized for specific methods, it is advantageous to employ basis sets optimized for specific molecular properties. The optimization criterion for properties, however, is less well-defined, as there in general is no property analog of the variational principle for the energy. Sadlej has derived rules for augmenting basis sets with additional functions for an improved description of electric properties.^{4,5} Provasi et al. have proposed the cc-pVTZ-J basis sets for indirect spin–spin coupling constants.^{6,7} These approaches have identified deficiencies in energy-optimized basis sets and added augmenting functions based on scaling the exponents of existing basis functions. Rappoport and Furche have optimized the exponents for diffuse functions to be used with the Karlsruhe basis sets by maximizing the atomic polarizability.⁸ We have optimized the exponents of tight

Received: October 27, 2014

Published: November 24, 2014



augmenting functions for nuclear magnetic shielding⁹ and spin–spin coupling^{10–12} constants by maximizing the change in the property relative to the un-augmented basis set. Vaara and co-workers have used the concept of completeness-optimized basis functions to cover the important exponent range for the given property.¹³

The ability to systematically approach the CBS limit for a variety of methods and properties is in our opinion important, but this should be combined with the requirement of computational efficiency. Modern basis sets employ the concept of error balance between basis functions of different angular momentum; the basis set error at a given ζ level is therefore implicitly given by the highest angular momentum function included in the basis set, and this also implicitly defines the number of lower angular momentum functions. A computationally efficient basis set should furthermore contract the primitive functions as strongly as possible without losing the inherent accuracy. The contraction of primitive functions to contracted functions can be done in two different ways: a general contraction,¹⁴ where all primitive functions are allowed to contribute to all contracted functions, and a segmented contraction,¹⁵ where each primitive function is allowed to contribute to only one contracted function. While most basis set families contain elements of both types, they can fairly easily be classified according to general and segmented contractions. The cc-pVXZ,^{16–20} nZaP,^{21–23} ANO-XZP,^{24–28} and pc-*n*^{29–34} basis sets are thus of the general contracted type, while the Pople,^{35–46} Sapporo,^{47–50} Jorge,^{51–54} and Karlsruhe⁵⁵ basis sets are of the segmented type.

General contraction is preferred from a construction point of view since it allows strict control of the contraction error and contraction coefficients are obtained simply as orbital coefficients, but general contracted basis sets are computationally less efficient than segmented ones. Construction of segmented contracted basis sets is significantly more difficult since each possible contraction scheme requires a simultaneous optimization of both exponents and contraction coefficients, which is a difficult multiple minima problem and furthermore leads to problems defining the contraction error. We have recently shown how a general contracted basis set can be converted into a segmented contracted version in a unique fashion with full control of the contraction error and identifying the correct local minimum in the exponent/coefficient parameter space.⁵⁶ This allowed conversion of the general contracted pc-*n* basis sets optimized for DFT energetic properties to the segmented contracted pcseg-*n* basis sets. In the present work we employ this method to construct computational efficient segmented basis sets optimized for calculating nuclear magnetic shielding constants. These new pcSseg-*n* basis sets can for the atoms H–Ar be considered as computationally more efficient versions of the previously defined pcS-*n* basis sets,⁹ and with the present work these are extended to also cover the atoms K–Kr. The pcS-*n*, as well as the present pcSseg-*n*, basis sets are optimized at the DFT level, but they have been shown to provide faster basis set convergence than other alternatives also at the MP2 level,⁵⁷ and they would thus appear to be generally useful for the calculation of nuclear magnetic shielding constants.

■ SEGMENTED POLARIZATION CONSISTENT BASIS SETS OPTIMIZED FOR CALCULATING NUCLEAR MAGNETIC SHIELDING CONSTANTS

The nuclear magnetic shielding tensor σ 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 tensor can be written in terms of diamagnetic and paramagnetic contributions, where the former is calculated as an expectation value of the diamagnetic shielding operator (\mathbf{H}^{DS}) while the latter is calculated as a response property of the paramagnetic spin orbit (\mathbf{H}^{PSO}) and orbital Zeeman (\mathbf{L}_G) operators.⁵⁸

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

Here, $\mathbf{r}_{iA/G}$ denotes the position vector between electron *i* and nucleus A or the gauge origin G, μ_N is the nuclear magneton, and g_A is the nuclear *g*-factor. Only the isotropic component, corresponding to one-third of the trace of σ , is observed in solution, and this will be called the nuclear magnetic shielding constant in the present work and given in units of ppm.

Our approach for constructing basis sets optimized for a given property can be divided into four stages:

- (1) An energy-optimized uncontracted pc-*n* basis set is selected as the base which defines the composition in terms of number of functions of each angular momentum. This ensures a balance between functions of different angular momentum and a proper representation of the orbitals at a given ζ level. These basis sets are scanned for deficiencies in terms of diffuse/tight basis functions for the given property at all ζ levels.
- (2) Optimum property augmenting functions are defined by optimizing the exponents of the augmenting functions to give the maximum change in the property relative to the base. The number of augmenting functions is determined by the change in the property compared to the inherent accuracy at a given ζ level. The latter is defined as the basis set error relative to the CBS limit. This leads to definition of a set of standard property augmenting functions.
- (3) The property augmented basis sets from (2) are general contracted using orbital coefficients. The maximum contraction is defined by the contraction error being smaller than, but comparable to, the inherent basis set error. This ensures maximum computational efficiency without losing the accuracy of the underlying set of primitive functions.
- (4) The general contracted basis sets from (3) are converted into computationally more efficient segmented contracted versions by the P-orthogonalization procedure,⁵⁶

Table 1. Compositions of the pcSeg-*n* Basis Sets

atoms	pcSeg-0	pcSeg-1	pcSeg-2	pcSeg-3	pcSeg-4
Primitive Functions					
H,He	3s	4s2p	6s3p1d	9s5p2d1f	11s7p3d2f1g
Li,Be	6s3p	8s4p	11s5p1d	15s7p2d1f	20s9p3d2f1g
B–Ne	6s4p	8s5p1d	11s7p2d1f	15s10p4d2f1g	20s12p6d3f2g1h
Na,Mg	9s6p	12s8p	15s10p1d	21s13p2d1f	25s16p3d2f1g
Al–Ar	9s7p	12s9p1d	15s11p2d1f	21s14p4d2f1g	25s17p6d3f2g1h
K,Ca	11s8p2d	14s10p4d	19s12p5d	28s15p6d1f	32s18p8d2f1g
Sc–Zn	11s8p4d	14s10p6d1f	19s12p8d2f1g	28s15p10d3f2g1h	32s18p12d5f3g2h1i
Ga–Kr	11s9p4d	14s11p7d	19s14p9d1f	28s17p11d2f1g	32s20p13d4f2g1h
Contracted Functions					
H,He	2s	2s1p	3s2p1d	4s4p2d1f	5s6p3d2f1g
Li,Be	3s2p	3s3p	4s4p1d	5s6p2d1f	6s7p3d2f1g
B–Ne	3s2p	3s3p1d	4s5p2d1f	5s8p4d2f1g	6s9p6d3f2g1h
Na,Mg	4s2p	4s4p	5s6p1d	6s8p2d1f	7s10p3d2f1g
Al–Ar	4s3p	4s5p1d	5s7p2d1f	6s9p4d2f1g	7s11p6d3f2g1h
K,Ca	5s4p1d	5s6p3d	6s8p4d	7s10p5d1f	8s12p6d2f1g
Sc–Zn	5s4p2d	5s6p3d1f	6s8p4d2f1g	7s10p6d3f2g1h	8s12p8d5f3g2h1i
Ga–Kr	5s5p1d	5s7p3d	6s10p4d1f	7s12p5d2f1g	8s14p6d4f2g1h

Table 2

Benchmark Molecular Systems
ethanamide, aniline, aziridine, B ₂ H ₆ , B ₂ O ₃ , B ₃ N ₃ H ₆ , benzene, BF ₃ , BH ₃ , B(OH) ₃ , butadiene, C ₂ F ₄ , C ₂ H ₂ , C ₂ H ₄ , C ₂ H ₆ , C ₃ N ₂ , CF ₄ , CH ₂ F ₂ , CH ₃ CHO, CH ₃ CN, CH ₃ F, CH ₃ NH ₂ , CH ₃ NO ₂ , CH ₃ OF, CH ₃ OH, CH ₃ ONO, CH ₄ , CO ₂ , CO, cyclopropane, cyclopropene, ethylamine, ethanol, F ₂ CO, F ₂ O, furan, glyoxal, H ₂ BNH ₂ , H ₂ CCO, H ₂ CO, H ₂ O ₂ , H ₂ O, HCN, HCOOH, HF, HOF, (CH ₃) ₂ O, (CH ₃) ₃ N, CH ₃ COOH, N ₂ H ₂ , N ₂ H ₄ , N ₂ O, NF ₃ , NH ₃ , oxirane, pyridine, pyrrole, C ₂ H ₃ CN, C ₂ H ₃ F, Be ₂ H ₄ , BeCl ₂ , BeF ₂ , BeH ₂ , Be(CH ₃) ₂ , BeS, Be(SH) ₂ , Li ₂ O ₂ , Li ₂ O, Li ₂ S ₂ , LiCl, LiF, LiH, LiCH ₃ , LiNH ₂ , LiOH, LiPH ₂ , LiSH, CH ₃ COCl, Al ₂ Cl ₆ , Al ₂ H ₆ , Al ₂ O ₃ , Al ₂ S ₃ , AlCl ₃ , AlF ₃ , AlH ₃ , C ₂ Cl ₄ , CCl ₄ , CH ₂ Cl ₂ , CH ₃ SH, Cl ₂ SO ₂ , ClF ₃ , ClF, CS ₂ , CSO, CS, (CH ₃) ₂ SO ₂ , (CH ₃) ₂ S, H ₂ CS, H ₂ S, SCl ₂ , HCl, HOCl, P ₄ , PCl ₃ , PCl ₅ , PF ₃ , PF ₅ , PH ₃ , POCl ₃ , S ₂ Cl ₂ , SF ₆ , Si ₂ H ₂ , Si ₂ H ₄ , Si ₂ H ₆ , SiCl ₄ , SiF ₄ , SiH ₄ , SiO, SiS, SO ₂ , SO ₃ , thiirane, thiophene, C ₂ H ₃ Cl, C ₂ H ₃ PH ₂ , C ₂ H ₃ SH, C ₂ H ₃ SiH ₃ , CH ₃ MgCl, Mg ₂ H ₄ , MgCl ₂ , MgF ₂ , MgH ₂ , Mg(CH ₃) ₂ , Mg(OH) ₂ , Mg(SH) ₂ , Na ₂ O ₂ , Na ₂ O, Na ₃ S, NaCl, NaF, NaH, NaCH ₃ , NaNH ₂ , NaPH ₂ , NaOH, NaSH, As ₄ , AsBr ₃ , AsBr ₅ , AsCl ₃ , AsF ₃ , AsH ₃ , BrCl, BrF, BrH, GaBr ₃ , GaCl ₃ , GaF ₃ , GaH ₃ , GeBr ₄ , GeCl ₄ , GeF ₄ , GeH ₄ , Se ₂ Br ₂ , SeBr ₂ , SeCl ₂ , SeF ₂ , SeH ₂ , CaBr ₂ , CaCl ₂ , CaF ₂ , CaH ₂ , Ca(OH) ₂ , Ca(SeH) ₂ , K ₂ O, K ₂ S ₂ , K ₂ Se, K ₂ Se ₂ , K ₂ S, KBr, KCH ₃ , KCl, KF, KH, CoBr ₃ , CoCl ₃ , CrCl ₃ , CrH ₂ , CrH ₃ , CrO, Cu ₂ O, CuBr, CuCl, CuH, CuO, FeH ₃ , FeO, MnBr ₂ , MnCl ₂ , MnCl ₃ , MnH ₃ , MnO ₂ , MnO, NiH ₂ , NiO, ScBr ₃ , ScCl ₃ , ScCl ₄ , ScH ₂ , ScH ₃ , ScO, TiCl ₂ , TiCl ₄ , TiH ₄ , TiO ₂ , TiO, VCl ₃ , VH ₃ , VO, ZnBr ₂ , ZnCl ₂ , ZnH ₂

which ensures full control of the contraction error, and can be considered as removing redundant primitive functions from the general contracted functions.

For nuclear magnetic shielding constants we have previously found that addition of a single tight p-function at stage (2) was required for the atoms H–Ar,⁹ and we have in the present work found that this is also the case for the atoms K–Kr. The standard augmenting p-functions were generated by scaling the inner-most primitive function with a factor of 6.5. No further augmenting functions were found to be required. The optimum contraction at stage (3) was for the atoms H–Ar found to be the same for the s-functions as for the energy-optimized pc-*n* basis sets, but the p-functions should be less strongly contracted. The same is found for the atoms K–Kr, but in addition, the d-functions must also be somewhat less contracted than for the energy-optimized pc-*n* basis sets. The conversion from general to segmented contraction at stage (4) involves a sequence of optimizations, since the exponent of the augmenting tight p-function should not be optimized, but the contraction coefficient should be. The optimum contraction from stage (3) combined with the P-orthogonalization at stage (4) in all cases lead to the conclusion that the p-function space should only contain a single contracted function for all atoms up to Kr. The s-contraction was transferred from the pcseg-*n* basis sets,⁵⁶ and the full s-, p-, and d-parameter space (exponents and contraction coefficients) without the augmenting tight p-function was energy-optimized using atomic energies with the selected level of p-contraction. These basis sets were then augmented with a tight p-function, and the contraction coefficients of the contracted p-function (only)

were optimized to include the augmenting function in the contraction. Diffuse augmenting functions were generated as for the aug-pcseg-*n* basis sets,⁵⁶ and the compositions of the new pcSeg-*n* basis sets are shown in Table 1.

■ COMPUTATIONAL DETAILS

All basis set optimizations have been done using a pseudo-Newton–Raphson algorithm in connection with gradients of the BLYP^{59,60} energy with respect to basis set exponents and contraction coefficients generated by numerical differentiation. Basis sets for third-row transition metals have been optimized with the atoms in their s²dⁿ electronic configuration.³⁴ Benchmark calculations of nuclear magnetic shielding constants have been done using the B97-2 functional⁶¹ and the gauge including atomic orbital procedure using the Gaussian09 program package.⁶² The benchmark molecular systems are listed in Table 2; their geometries have been optimized with the B3LYP^{63,64} or PBE⁶⁵ (transition metal systems) functionals and the 6-31G(d) basis set for the spin state of lowest energy. Several of the transition metal systems are paramagnetic open-shell species, but this is inconsequential for probing the basis set convergence, and the systems have been chosen to probe a wide range of bonding in order to provide a robust test for the basis set convergence. Basis set limiting values have been taken as the results with the uncontracted aug-pc-4 basis set augmented with two tight p-functions and are estimated to be accurate to at least ~0.1%. Geometries for the Ni complexes, shown in Tables 4 and 5, below, were optimized at the PBE/pcseg-1 level within T symmetry.⁵⁶

BENCHMARKS

Quantifying the basis set convergence by statistical methods is complicated by the fact that the magnitude of the isotropic shielding constants vary from close to 0 to ~30 000 ppm for the benchmark systems in Table 2. Quantification in terms of raw shielding constants will be dominated by those that are large in magnitude, while a percent-wise measure will be dominated by those with CBS shielding constants close to zero. We have elected to use a percent-wise measure ($|\Delta\sigma/\sigma_{\text{CBS}}| \times 100\%$, $\Delta\sigma = \sigma - \sigma_{\text{CBS}}$), but excluded those with low σ_{CBS} values, taken as less than 10 ppm for nonmetallic atoms and less than 20 ppm for transition metal atoms (nine constants in total excluded).

The mean absolute deviations (MADs) of the percent-wise error relative to the CBS limit (note the logarithmic axis) are shown in Figure 1 as a function of the ζ level with the pcSseg- n

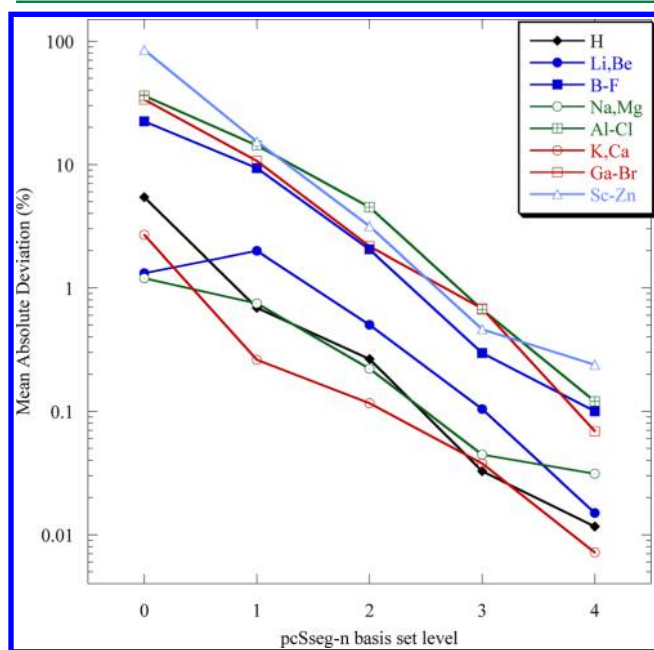


Figure 1. Mean absolute percent-wise deviations of B97-2 nuclear magnetic shielding constants relative to the complete basis set limit ($|\Delta\sigma/\sigma_{\text{CBS}}| \times 100\%$) as a function of basis set quality for the systems in Table 2.

basis sets, partitioned according to rows and periods of the periodic table. It is clear that the calculated shielding constants for all atoms display an exponential convergence toward the CBS limit. It is also evident that s-block atoms (including hydrogen) have significantly smaller basis set errors at a given ζ

level than p- and d-block atoms. Within the s-block elements, the errors decline from the first- through second- and to third-row atoms, with hydrogen resembling second-row elements. Each step up in basis set quality reduces the percent-wise basis set error by roughly a factor of 4. Addition of diffuse functions in general has a small effect, except for ionic systems like Na_2O_2 and K_2O . Tables showing the MAD and maximum AD values for the pcSseg- n and aug-pcSseg- n basis sets are provided as Supporting Information. The pcSseg-1 and pcSseg-2 basis sets will often be sufficiently accurate for routine uses, while the pcSseg-3 and pcSseg-4 basis sets allow strict control of basis set errors, even for pathological systems.

The IGLO basis sets have been used for the calculation of magnetic properties, but they have only been defined for hydrogen and the first- and second-row p-block atoms.^{66,67} In terms of size, the IGLO-II basis set is larger than pcSseg-1, while IGLO-III is similar to pcSseg-2. In a previous publication it was shown that they provide basis set errors that are larger or at par with the general contracted pcS- n basis sets, and as the present pcSseg- n basis sets essentially only represent a computationally more efficient version of the pcS- n for these atoms, that conclusion is still valid. Table 3 shows a comparison of the MAD with two other popular segmented basis set families, the Pople-style 6-31G(d,p) and 6-311G(2df,2pd), and the Karlsruhe Def2 basis sets in DZP, TZP, and QZP quality. None of these have been designed for calculating nuclear magnetic shielding constants, but have been used for that purpose. Table 3 shows that the pcSseg- n basis sets in almost all cases produce significantly lower basis set errors at a given ζ quality level. The three exceptions are the Def2-SVP basis set for third-row transition metal and p-block systems which produces slightly smaller MAD values than pcSseg-1, and the 6-311G(2df,2pd) compared to pcSseg-2 for second-row s-block atoms. Tests showed that for Def2-SVP this is due to fortuitous contraction errors which for the benchmark systems produce slightly better agreement with the CBS limit. Table 1 shows that the pcSseg-1 basis set is 10 and 11 primitive p-functions contracted to 6 and 7 p-functions for third-row transition metals and p-block elements, respectively, where the innermost 5 primitive functions are contracted and the remaining are left uncontracted. The corresponding Def2-SVP basis set is 9 and 10 primitive p-functions contracted to 3 and 4 p-functions, respectively, where the innermost 8 primitive functions are contracted into two functions with the partitioning 5-3. The contraction of the inner-valence 3 primitive functions produces large contraction errors, which for these systems accidentally bring the calculated results in slightly better agreement with the CBS limit on average. For the 6-311G(2df,2pd) for second-row

Table 3. Mean Absolute Percent-wise Deviations of B97-2 Nuclear Magnetic Shielding Constants Relative to the Complete Basis Set Limit ($|\Delta\sigma/\sigma_{\text{CBS}}| \times 100\%$) as a Function of Basis Sets for the Systems in Table 2

atoms	DZP quality			TZP quality			QZP quality	
	pcSseg-1	6-31G(d,p)	Def2-SVP	pcSseg-2	6-311G(2df,2pd)	Def2-TZP	pcSseg-3	Def2-QZP
H	0.7	1.8	3.4	0.26	0.89	1.2	0.03	0.20
Li,Be	2.0	7.4	4.9	0.50	2.5	3.5	0.10	1.11
B-F	9.3	44.4	29.5	2.1	14.5	10.1	0.30	5.65
Na,Mg	0.8	4.0	3.8	0.22	0.14	0.30	0.04	0.19
Al-Cl	14.3	17.5	25.3	4.5	7.3	8.2	0.70	1.55
K,Ca	0.3	6.4	3.7	0.12	0.61	2.0	0.04	0.81
Ga-Br	10.7	24.8	7.3	2.2	2.6	6.7	0.68	3.35
Sc-Zn	15.3	119.5	14.5	3.2	66.8	7.3	0.59	5.56

Table 4. Calculated ^{61}Ni Chemical Shifts (ppm) in $\text{Ni}(\text{PF}_3)_4$, $\text{Ni}(\text{PCl}_3)_4$, and $\text{Ni}(\text{PMe}_3)_4$, Where the Value for $\text{Ni}(\text{CO})_4$ Has Been Defined as Zero

basis set/method	$\text{Ni}(\text{PF}_3)_4$			$\text{Ni}(\text{PCl}_3)_4$			$\text{Ni}(\text{PMe}_3)_4$		
	HF	B97-2	TPSS	HF	B97-2	TPSS	HF	B97-2	TPSS
pcSseg-0	−2605	−1245	−886	−229	864	1013	−489	445	538
pcSseg-1	−1787	−963	−763	−765	242	512	−710	38	150
pcSseg-2	−2159	−999	−766	−1338	25	335	−1118	−66	84
pcSseg-3	−2170	−985	−748	−1317	76	388	−1155	−76	73
pcSseg-4	−2165	−984	−748	−1311	80	393	−1154	−76	76
aug-pcSseg-0	−2611	−1063	−743	−1519	121	451	−1517	−82	146
aug-pcSseg-1	−2139	−1000	−768	−1291	56	366	−1156	−116	38
aug-pcSseg-2	−2178	−964	−729	−1326	63	368	−1166	−78	79
aug-pcSseg-3	−2172	−990	−750	−1317	75	384			
6-31G(d,p)	−2349	−937	−709	−809	668	846	−699	301	322
6-311G(2df,2pd)	−1520	−720	−525	−366	465	710	−479	118	214
exptl ^{70–72}	−929			267			40		

Table 5. Calculated B97-2 ^{61}Ni Chemical Shifts (ppm) in $\text{Ni}(\text{PF}_3)_4$, $\text{Ni}(\text{PCl}_3)_4$, and $\text{Ni}(\text{PMe}_3)_4$, Where the Value for $\text{Ni}(\text{CO})_4$ Has Been Defined as Zero Using Locally Dense Basis Sets

basis set ^a				chemical shift			no. of basis functions			
Ni	A _a	A _b	A _c	$\text{Ni}(\text{PF}_3)_4$	$\text{Ni}(\text{PCl}_3)_4$	$\text{Ni}(\text{PMe}_3)_4$	$\text{Ni}(\text{CO})_4$	$\text{Ni}(\text{PF}_3)_4$	$\text{Ni}(\text{PCl}_3)_4$	$\text{Ni}(\text{PMe}_3)_4$
pcSseg-3	pcSseg-3	pcSseg-3	pcSseg-3	−985	76	−76	693	1285	1333	2473
	pcseg-3	pcseg-3	pcseg-1	−986	75	−90	597	1093	1141	1321
	pcseg-3	pcseg-2	pcseg-0	−982	70	−113	477	733	781	805
	pcseg-2	pcseg-1	pcseg-0	−971	69	−113	293	421	469	493
pcSseg-2	pcSseg-2	pcSseg-2	pcSseg-2	−999	25	−66	361	677	761	1181
	pcseg-2	pcseg-2	pcseg-1	−999	29	−93	313	569	617	785
	pcseg-2	pcseg-1	pcseg-0	−982	72	−105	249	377	425	449
	pcseg-1	pcseg-0	pcseg-0	−882	190	−35	165	253	301	325

^aA_a, basis set on atoms directly coordinated to Ni (C and P); A_b, basis set on atoms one bond removed from Ni (O, F, Cl, C); A_c, basis set on atoms two bonds removed from Ni (H for $\text{Ni}(\text{PMe}_3)_4$).

s-block atoms, the marginally lower MAD value is due to the 6-311G(2df,2pd) basis set containing a 2d1f polarization space for these atoms, compared to only 1d for the pcSseg-2.

ILLUSTRATIVE APPLICATION

A novel feature of the present pcSseg-*n* basis sets is that they are defined for all atoms up to Kr, including third-row transition metals, and in qualities from (unpolarized) DZ to SZP. To illustrate the advantages of being able to separate the basis set and method errors, we have calculated the ^{61}Ni nuclear magnetic shielding constants in four different tetrahedral complexes, $\text{Ni}(\text{CO})_4$, $\text{Ni}(\text{PF}_3)_4$, $\text{Ni}(\text{PCl}_3)_4$, and $\text{Ni}(\text{PMe}_3)_4$,⁶⁸ with three different methods, HF and two DFT functionals, B97-2⁶¹ and revTPSS.⁶⁹ The results with the pcSseg-*n* as well as the 6-31G(d,p) and 6-311G(2df,2pd) basis sets are shown in Table 4. Using the calculated value for $\text{Ni}(\text{CO})_4$ as the reference, it is seen that the B97-2 method tends to give chemical shifts that are slightly too negative in the CBS limit, while TPSS gives values that are too positive compared to experimental values.^{70–72} Both DFT methods, however, perform better than HF. Augmentation with diffuse functions produces results closer to the CBS limit for a given ζ level, but it should be recognized that diffuse augmentation also implies an increased number of basis functions. The basis set convergence in terms of number of functions is very similar for the pcSseg-*n* and aug-pcSseg-*n* basis sets, with the latter providing results of roughly $(n + 1/2)$ ζ quality. The results with the 6-31G(d,p) and 6-311G(2df,2pd) basis sets show that these

contains basis set errors that easily obscure the inherent accuracy of the methods, and the 6-311G(2df,2pd) results are on average further from the CBS values than the 6-31G(d,p) results. The B97-2 functional in combination with the pcSseg-1 basis set gives good agreement with experiments for all three systems, but the results with the larger basis sets show that this is due to a fortuitous cancellation of method and basis set errors for these particular systems. A reliable quantification of the accuracy of different methods requires the ability to rigorously separate method and basis set errors, and the present pcSseg-*n* basis sets should be a valuable resource in that respect.

The pcSseg-*n* basis sets are somewhat larger than the regular pcseg-*n*,⁵⁶ primarily due to the less contraction of the p-functions. The nuclear magnetic shielding, however, is a sufficiently localized property that it is possible to employ the concept of locally dense basis sets if the shielding is only required for a few atoms within the whole molecule.⁷³ Table 5 illustrates this approach for the four Ni complexes, where a high quality basis set (pcSseg-3 or -2) is employed for the Ni atom, while the regular pcseg-*n* basis sets are used for the ligand atoms, and the ζ quality of these can be tapered off as they are farther and farther removed from the atom of interest. The calculated chemical shifts are only mildly affected by employing such mixed quality basis sets, and the last four columns in Table 5 show that the total number of basis functions can be substantially reduced. The full pcSseg-3 basis set on all atoms, for example, amounts to 2473 basis functions for the $\text{Ni}(\text{PMe}_3)_4$ complex, while it is only 493 functions for the

pcSseg-3/pcseg-2,-1,-0 combination, and this translates into a reduction of the computational time by roughly 2 orders of magnitude. In contrast to the findings of Reid et al.,⁷³ there is virtually no difference between using the pcseg-*n* or pcSseg-*n* basis sets for the ligand atoms in the present systems. It is difficult to provide general recommendations in terms of basis set combinations, as the required accuracy will be systems dependent, but the sensitivity can be probed by running a series of calculations where the ζ level of the ligand atoms is gradually increased until the changes are decided to be sufficiently small.

SUMMARY

The present work defines a systematic way of optimizing basis sets toward specific molecular properties that have other basis set requirements than for energetic properties. An energy-optimized uncontracted basis set is augmented with property specific functions to improve the basis set convergence. It is then general contracted based on the property basis set error relative to the complete basis set limit at each ζ level, and finally converted into a computational more efficient segmented contraction by removing the redundancy among the primitive functions in the general contracted functions. The procedure is applied to define a family of basis sets optimized for calculating nuclear magnetic shielding constants, which are shown to have lower basis set errors than other popular alternatives. The pcSseg-*n* basis sets are available⁷⁴ in qualities ranging from (unpolarized) double- ζ to pentuple- ζ quality and should be suitable for both routine and benchmark calculations of nuclear magnetic shielding constants. The ability to separate the method and basis set errors allows probing the inherent method errors and potentially opens the possibility to design more accurate methods. While these basis sets have been developed for use with DFT methods, they may also be suitable for wave function methods including electron correlation. They can furthermore be used as locally dense basis sets in combination with the energy-optimized pcseg-*n* basis sets for large systems, where shielding constants are only required for a (small) subset of atoms.

ASSOCIATED CONTENT

Supporting Information

Tables comparing mean and maximum absolute deviations for the nuclear magnetic shielding constants for the molecules in Table 2, divided according to rows and blocks in the periodic table. This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Author

*E-mail: frj@chem.au.dk

Funding

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

Notes

The authors declare no competing financial interest.

REFERENCES

- Hill, J. G. *Int. J. Quantum Chem.* **2013**, *113*, 21–34.
- Jensen, F. *WIREs* **2013**, *3*, 273–295.
- Weigend, F.; Furche, F.; Ahlrichs, R. *J. Chem. Phys.* **2003**, *119*, 12753–12762.
- Sadlej, A. J. *Collect. Czech. Chem. Commun.* **1988**, *53*, 1995–2016.
- Sadlej, A. J. *Theor. Chem. Acc.* **1991**, *79*, 123–140.
- Provasi, P. F.; Aucar, G. A.; Sauer, S. P. A. *J. Chem. Phys.* **2001**, *115*, 1324–1334.
- Provasi, P. F.; Sauer, S. P. A. *J. Chem. Phys.* **2010**, *133*, No. 054308.
- Rappoport, D.; Furche, F. *J. Chem. Phys.* **2010**, *133*, No. 134105.
- Jensen, F. *J. Chem. Theory Comput.* **2008**, *4*, 719–727.
- Jensen, F. *J. Chem. Theory Comput.* **2006**, *2*, 1360–1369.
- Benedikt, U.; Auer, A. A.; Jensen, F. *J. Chem. Phys.* **2008**, *129*, No. 064111.
- Jensen, F. *Theor. Chem. Acc.* **2010**, *126*, 371–382.
- Manninen, P.; Vaara, J. *J. Comput. Chem.* **2006**, *27*, 434–445.
- Raffenet, R. C. *J. Chem. Phys.* **1973**, *58*, 4452–4458.
- Dunning, T. H. *J. Chem. Phys.* **1970**, *53*, 2823–2833.
- Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007–1023.
- Dunning, T. H.; Peterson, K. A.; Wilson, A. K. *J. Chem. Phys.* **2001**, *114*, 9244–9253.
- Wilson, A. K.; Woon, D. E.; Peterson, K. A.; Dunning, T. H. *J. Chem. Phys.* **1999**, *110*, 7667–7676.
- Prascher, B. P.; Woon, D. E.; Peterson, K. A.; Dunning, T. H., Jr.; Wilson, A. K. *Theor. Chem. Acc.* **2011**, *128*, 69–82.
- Balabanov, N. B.; Peterson, K. A. *J. Chem. Phys.* **2005**, *123*, No. 064107.
- Zhong, S.; Barnes, E. C.; Petersson, G. A. *J. Chem. Phys.* **2008**, *129*, No. 184116.
- Barnes, E. C.; Petersson, G. A.; Feller, D.; Peterson, K. A. *J. Chem. Phys.* **2008**, *129*, No. 194115.
- Barnes, E. C.; Petersson, G. A. *J. Chem. Phys.* **2010**, *132*, No. 114111.
- Widmark, P. O.; Malmqvist, P. A.; Roos, B. O. *Theor. Chem. Acc.* **1990**, *77*, 291–306.
- Widmark, P. O.; Joakim, B.; Persson, B. O. *Theor. Chem. Acc.* **1991**, *79*, 419–432.
- Pouamerigo, R.; Merchan, M.; Nebotgil, I.; Widmark, P. O.; Roos, B. O. *Theor. Chem. Acc.* **1995**, *92*, 149–181.
- Pierloot, K.; Dumez, B.; Widmark, P. O.; Roos, B. O. *Theor. Chem. Acc.* **1995**, *90*, 87–114.
- Roos, B. O.; Lindh, R.; Malmqvist, P. A.; Veryazov, V.; Widmark, P. O. *J. Phys. Chem. A* **2004**, *108*, 2851–2858.
- Jensen, F. *J. Chem. Phys.* **2001**, *115*, 9113–9125.
- Jensen, F. *J. Chem. Phys.* **2002**, *116*, 3502–3502.
- Jensen, F.; Helgaker, T. *J. Chem. Phys.* **2004**, *121*, 3463–3470.
- Jensen, F. *J. Phys. Chem. A* **2007**, *111*, 11198–11204.
- Jensen, F. *J. Chem. Phys.* **2012**, *136*, No. 114107.
- Jensen, F. *J. Chem. Phys.* **2013**, *138*, No. 014107.
- Hehre, W. J.; Ditchfie, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257–2261.
- Dill, J. D.; Pople, J. A. *J. Chem. Phys.* **1975**, *62*, 2921–2923.
- 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*, 3654–3665.
- Rassolov, V. A.; Pople, J. A.; Ratner, M. A.; Windus, T. L. *J. Chem. Phys.* **1998**, *109*, 1223–1229.
- Rassolov, V. A.; Ratner, M. A.; Pople, J. A.; Redfern, P. C.; Curtiss, L. A. *J. Comput. Chem.* **2001**, *22*, 976–984.
- Binning, R. C.; Curtiss, L. A. *J. Comput. Chem.* **1990**, *11*, 1206–1216.
- Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650–654.
- McLean, A. D.; Chandler, G. S. *J. Chem. Phys.* **1980**, *72*, 5639–5648.
- Curtiss, L. A.; McGrath, M. P.; Blaudeau, J. P.; Davis, N. E.; Binning, R. C.; Radom, L. *J. Chem. Phys.* **1995**, *103*, 6104–6113.
- Blaudeau, J. P.; McGrath, M. P.; Curtiss, L. A.; Radom, L. *J. Chem. Phys.* **1997**, *107*, 5016–5021.
- Wachters, A. J. *J. Chem. Phys.* **1970**, *52*, 1033–1036.
- Hay, P. J. *J. Chem. Phys.* **1977**, *66*, 4377–4384.
- Thakkar, A. J.; Koga, T.; Saito, M.; Hoffmeyer, R. E. *Int. J. Quantum Chem.* **1993**, *343*–354.

- (48) Tatewaki, H.; Koga, T.; Takashima, H. *Theor. Chem. Acc.* **1997**, *96*, 243–247.
- (49) Koga, T.; Tatewaki, H.; Matsuyama, H.; Satoh, Y. *Theor. Chem. Acc.* **1999**, *102*, 105–111.
- (50) Noro, T.; Sekiya, M.; Koga, T. *Theor. Chem. Acc.* **2003**, *109*, 85–90.
- (51) Neto, A. C.; Muniz, E. P.; Centoducatte, R.; Jorge, F. E. *THEOCHEM* **2005**, *718*, 219–224.
- (52) Jorge, F. E.; Sagrillo, P. S.; de Oliveira, A. R. *Chem. Phys. Lett.* **2006**, *432*, 558–563.
- (53) Barbieri, P. L.; Fantin, P. A.; Jorge, F. E. *Mol. Phys.* **2006**, *104*, 2945–2954.
- (54) Machado, S. F.; Camiletti, G. G.; Canal Neto, A.; Jorge, F. E.; Jorge, R. S. *Mol. Phys.* **2009**, *107*, 1713–1727.
- (55) Weigend, F.; Ahlrichs, R. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.
- (56) Jensen, F. *J. Chem. Theory Comput.* **2014**, *10*, 1074–1085.
- (57) Flaig, D.; Maurer, M.; Hanni, M.; Braunger, K.; Kick, L.; Thubauville, M.; Ochsenfeld, C. *J. Chem. Theory Comput.* **2014**, *10*, 572–578.
- (58) Helgaker, T.; Jaszunski, M.; Ruud, K. *Chem. Rev.* **1999**, *99*, 293–352.
- (59) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098–3100.
- (60) Lee, C. T.; Yang, W. T.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785–789.
- (61) Wilson, P. J.; Bradley, T. J.; Tozer, D. J. *J. Chem. Phys.* **2001**, *115*, 9233–9242.
- (62) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*; Gaussian, Inc.: Wallingford, CT, 2009.
- (63) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623–11627.
- (64) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
- (65) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
- (66) Schindler, M.; Kutzelnigg, W. *J. Chem. Phys.* **1982**, *76*, 1919–1933.
- (67) Kutzelnigg, W.; Fleischer, U.; Schindler, M. *The IGLO-Method: Ab initio calculation interpretation of NMR chemical shifts magnetic susceptibilities*; NMR Basic Principles and Progress 213; Springer-Verlag: Berlin/Heidelberg, 1991; pp 165–262.
- (68) Buehl, M.; Peters, D.; Herges, R. *Dalton Trans.* **2009**, 6037–6044.
- (69) Perdew, J. P.; Ruzsinszky, A.; Csonka, G. I.; Constantin, L. A.; Sun, J. *Phys. Rev. Lett.* **2009**, *103*, No. 026403.
- (70) Avent, A. G.; Cloke, F. G. N.; Day, J. P.; Seddon, E. A.; Seddon, K. R.; Smedley, S. M. *J. Organomet. Chem.* **1988**, *341*, 535–541.
- (71) Benn, R.; Rufinska, A. *Magn. Reson. Chem.* **1988**, *26*, 895–902.
- (72) Behringer, K. D.; Blumel, J. *Magn. Reson. Chem.* **1995**, *33*, 729–733.
- (73) Reid, D. M.; Kobayashi, R.; Collins, M. A. *J. Chem. Theory Comput.* **2014**, *10*, 146–152.
- (74) Schuchardt, K. L.; Didier, B. T.; Elsethagen, T.; Sun, L.; Gurumoorathi, V.; Chase, J.; Li, J.; Windus, T. L. *J. Chem. Inf. Mod.* **2007**, *47*, 1045–1052.