

# New Gaussian Basis Sets for Highly Accurate Calculations of Isotropic Hyperfine Coupling Constants at Hydrogen

Afaf R. Al Derzi<sup>1</sup>, Stefan Fan<sup>2</sup>

Department of Chemistry, College of Science, Al-Nahrain University, Baghdad, Iraq

<sup>2</sup> Quantum Theory Project, Departments of Physics and of Chemistry, University of Florida, USA

## Abstract

ESR spectroscopy is an important method for characterizing radicals. However, the isotropic hyperfine coupling constants are difficult to calculate, the effects of solvent and molecular vibration are often not negligible. We provide benchmark values near the JULLC complete basis set limit for nonvibrating  $\text{Li}$ ,  $\text{H}_2$ , and  $\text{H}_2^+$  in gas phase. These are used to judge the performance of already existing and newly constructed basis sets. The new (aug)cc-pVQZ-1s basis sets are superior to previous basis sets. They have average errors between -0.3 and +0.1% of the reference values. The rms deviation from the average error is less than ±0.7%. Only extreme contraction significantly increases the errors. Larger valence basis sets improve the accuracy.

## Introduction

Electron spin resonance (ESR) spectroscopy is among the most important methods for the characterization and analysis of radicals. Anisotropic hyperfine coupling tensors are important in solid-state experiments and can be calculated relatively easily. Isotropic hyperfine coupling constants (IHFCs) can be obtained from solutions, but their calculation often is difficult. Intrapole

IHFCs ( $\langle \cdot \rangle$ ) ( $A_{\text{iso}}$ ) are proportional to the spin density ( $\rho_s$ ) at the nucleus of interest,  $N$ :

$$A_{\text{iso}}(N) = \frac{8\pi}{3} \left( \frac{g_s}{g_0} - g_s \beta_s \rho_s(z) \right) \quad (1)$$

In organic radicals, the ratio of the electronic g factors in the radical ( $g_s$ ) and in vacuum ( $g_0$ ) is usually well approximated by unity. The nuclear g factor,  $g_s$ , is isotope dependent. The electronic and nuclear magnetons,  $\beta_s$  and  $\beta_0$ , are constants. For hydrogen, the proportionality factor evaluates to 1898.0 G Bohr [2,3]. The spin density at a nucleus is usually calculated by the  $\delta$ -function formalism [4,5-7], which determines the difference between the  $\alpha$ - and  $\beta$ -electron densities ( $\rho_\alpha, \rho_\beta$ ) at the location of the nucleus. In the equation above,  $n$  is the number of unpaired electrons,  $\psi$  is a basis function,  $\hat{S}_z$  is the  $z$  component of the spin operator, and  $D_{\rho_s}$  is an element of the density matrix. This procedure is computationally easy but also sensitive to errors in the spin density at the nucleus. A nonlocal operator has been developed by Hwang [8] on the basis of the work of Miller, Sander, and Feinberg [9]. Rasanayagam and Chipman developed improved operators [10] that combine local aspects of both the  $\delta$ - and JBLF operators. At the 8s-1s-2p<sup>2</sup>-3p<sup>2</sup>-2d<sup>2</sup>-1f level, the rms basis set error in computed spin densities was reduced to ~2%. Both solvent and vibrational effects change the IHFCs of a molecule by a few percent with respect to the non-vibrating values. In extreme cases, solvent effects may reach 10%, gas-phase [8] and vibrational effects (large amplitude motions)

may be as large as 37% [11]. These effects may be quantified with great difficulty using gas-phase spectroscopy and isotropic solvation. Computationally, solvent effects can be approximated, but vibrational effects are expensive to calculate. The resulting lack of data makes it hard to judge the quality of single-point JULLC calculations by comparing them to experimental coupling constants. It is well known [4-7, 11] that the computed IHFCs of a gas-phase molecule depend on the quantity, method, size-vibrating and basis set. Because the errors due to method and basis set cannot easily be distinguished, systematic studies improving the treatment of correlation and increasing the completeness of the basis set are needed. Two problems are related to the basis set. First, Gaussian functions (primitives) have zero slope at the origin, which is usually placed at the nucleus of an atom. The absence of a sharp derivative of  $\psi$  at the nucleus compared to those of hydrogenic orbitals, the introduction of very light primitives [12] strongly reduces this deficiency by moving the turning point [13] of the wave function closer to the nucleus. The second problem is that the contraction coefficients of primitives are usually optimized for the computation of energy differences (relative energies, geometry optimizations, frequency calculations). These basis sets are least flexible in the outer valence region where the influence of neighboring atoms is strongest. However, the electron and spin density at the nucleus is determined by both the occupation of the basis functions and the contraction coefficient and exponents of the constituting primitives. Because light functions are more important for the calculation of IHFCs than for energy differences, basis sets for the calculation of IHFCs have more flexibility in the inner valence and core regions than were used to calculate energy differences.

Correlation is important, too. At the Hartree-Fock level, IHFCs are often wrong by 100% [13,14] and vary strongly from small to

100%. [13]. Because tight functions are more important for iHFCCs than for energy differences, core correlation is necessary for even the qualitative computation of iHFCCs of atoms heavier than helium. This also explains why the convergence of iHFCCs from multireference CI calculations with energy-selected "excited configurations" is slower than the convergence of the total energy. The fact that the spin density is the difference of the comparably sized  $\alpha$  and  $\beta$  densities may increase problems caused by both overly contracted basis sets and an insufficient treatment of correlation. Various groups have attempted the quantitative calculation of isotropic hyperfine coupling constants with ab initio methods. Semiquantitative agreement with experiment may be reached with Chipman's method, 5 MR-CISD [17], MRD-CI/B<sub>K</sub> [4], B3LYP/LPR-n [6], quadratic CI [14-16], and coupled-cluster methods. MR-CISD calculations with energy-selected excited suffer from the slow convergence of the iHFCCs, configurations. This problem is strongly reduced in MRD-CI/B<sub>K</sub>. DFT is the cheapest method and is used in the calculation of both solvent and vibrational effects, but no systematic way of improving the results is known. Both quadratic CI (QCI) and coupled cluster (CC) give good results, especially when triple excitations are included via perturbation theory. CC is more complete than QCI with little more computational effort. [17] Basis sets developed by Chipman [18], Heller et al. [19], and Barone [20] as well as large, uncontracted, even-tempered basis sets [21] perform better than the usual, energy-optimized basis sets. However, for the iHFCCs of hydrogen atoms, errors of 20% or more are common. It is our goal to find basis sets for hydrogen that reduce this error. To avoid the problems related to geometry, solvent, and vibrational effects, we determined reference values from either the analytical solution of the Schrödinger equation (11) or from calculations close to the full CI/complete basis set limit. These calculations are described in the appendix. H, H<sub>2</sub><sup>+</sup>, and H<sub>2</sub><sup>2+</sup> are experimentally known. The isotropic hyperfine coupling constant (IIC) of H<sub>2</sub><sup>+</sup> in noble gases varies between 501.0 and 512.3 G [22]. Experimental iHFCCs for H<sub>2</sub><sup>+</sup> in various vibrational and rotational states are available. [23,24] The value for v = 0, N = 1 is 329.3 G. Calculations suggest 333.9 G for v = 0, N = 0 and 334.3 G at an internuclear distance of 1.0384 Å [25]. H<sub>2</sub><sup>2+</sup> is experimentally known [26,27], but its iHFCCs are not.

#### Computational Details

All calculations were performed using the ACES II [28] program system. Spherical d, f, and g functions were used. On the basis of the results of Pereira, Watts, and Bartlett [3], we include correlation effects by UIMP-based CCSD

calculations. The spin densities at the nuclei were computed with the  $\delta$ -function formalism from the CC relaxed density matrix. [3,29] All on H<sub>2</sub><sup>+</sup> use a bond length of 1.057 Å, which calculations is intermediate between the cc-pVQZ and cc-pV5Z optimized values of 1.0571 and 1.0569 Å. The calculations on H<sub>2</sub><sup>2+</sup> use the UIMP-CCSD/cc-pVQZ optimized D<sub>2</sub><sup>+</sup> geometry, a saddle point chosen for its symmetry.

#### Basis Set Development

The isotropic HFCC is proportional to the spin density at the nucleus. In isolated atoms, the isotropic HFCC is directly related to electrons in s orbitals. Electrons in other orbitals have nodal planes at the nucleus and hence only indirectly influence the HFCC via the electron-electron interaction. In basis sets composed of Gaussian functions, very tight s primitives are needed to describe the electron density at the nucleus. Ordinary basis sets, with exponents optimized for atomic energies, do not necessarily contain these very tight s primitives. This, together with insufficient flexibility in the inner valence and core regions, leads to inexact electron densities at the nuclei. The problem is particularly severe for hydrogen.

In molecular radicals, the spin density at a nucleus is described by the s functions of that atom and valence and diffuse functions of surrounding atoms. Because the values of these functions at the nucleus of a neighbouring atom are relatively small, we expect their influence to be secondary to the tight s primitives. Bond formation will also increase the anisotropy of the electron density at the atom of interest. Therefore, we expect higher angular momentum functions to be more important than in the isolated atom case.

Because a good treatment of correlation effects is necessary to calculate isotropic HFCCs accurately, we derive the new basis sets from the correlation-consistent basis sets of Dunning and co-workers [30]. First, tight s functions are added, and the basis sets are tested in SCF calculations of the hydrogen atom. In this way, correlation effects are not present, and the exact result is known. In a second step, we test the basis sets on the H<sub>2</sub><sup>+</sup> molecule. There, polarization (and, because of the large bond distance, diffuse) functions become important. The contraction pattern is of importance, too. We then calculate isotropic HFCCs in D<sub>2</sub>(H<sub>2</sub><sup>+</sup>), a complex between H<sub>2</sub> and H<sub>2</sub><sup>+</sup>. Here, correlation becomes important. We compare to well-converged isotropic HFCCs calculated for the isolated, nonvibrating molecules to avoid the problems and uncertainties related to solvent and vibrational averaging effects. We will examine how consistent the new basis sets reproduce the reference iHFCCs and compare to other basis sets. The fifth step is the contraction of the basis sets to

reduce the cost of calculations. The contraction pattern should be flexible enough to describe all three radicals well, not only atomic hydrogen in its ground state. We will also examine the influence of basis sets at neighboring atoms on the calculated isotropic HFCCs. This will be important in cases where only the HFCCs of the hydrogen atoms are of interest because it might be possible to describe the non-hydrogen atoms by regular cc-basis sets.

### Hydrogen Atom

The exact solution of the Schrödinger equation for the hydrogen atom gives the electron density (and hence the spin density) at the nucleus as  $\chi^2$ . This corresponds to an isotropic HFC of 307.7 G. The experimental value in liquid methane [31] is 303.8 G, less than 1% lower. Because the wave function of hydrogen is an s orbital, only s functions contribute. Table 1 shows that the energies and isotropic HFCCs calculated with cc- and cc-pcc-basis sets converge smoothly

toward the exact results, yet the HFCCs are not good. The cc-pVQZ basis sets underestimate the HFCC by ~16%, and even the cc-pV5Z basis sets are 3% too low.

Additional tight s functions improve the situation. Upon adding the tightest s primitive of the cc-pVQZ basis set to (cc-pVDZ-m1), the HFCC improves by 13%. Adding cc-pV12 the tightest two s primitives of cc-pV5Z (cc-pV12-m2) improves the HFCC by 14%. A more systematic way of adding tight functions is to multiply the largest s exponent of the previous basis set by a constant (4 in our case). These basis sets are named cc-pVXZ-nm. Helgaker et al. [32] showed that this type of expansion of cc-basis sets in conjunction with uncontracting the s functions leads to a smoothly convergent series of NMR coupling constants that, in their case, can be determined by the Fermi contact term.

Table (1): Percentage of the Reference Energy and Isotropic HFCCs ( $\Delta_{\text{FC}}$ ) of the Nuclei.

Basis set	$\text{H}^1$		$\text{H}_2^1$		$\text{H}_2^2$	
	% energy	$\Delta_{\text{FC}}$	% energy	$\Delta_{\text{FC}}$	% energy	$\Delta_{\text{FC}}$
cc-pVQZ	99.86	-8.43	99.61	-55.88	99.2	-20.9
cc-pV12	99.96	90.3	99.94	90.9	99.82	91.8
cc-pVQZ	99.99	55.78	99.98	93.26	99.93	93.57
cc-pV5Z	100	57.17	100	57.06		
aug-cc-pVDZ	99.87	-87.54	99.77	-81.96	99.36	-81.34
aug-cc-pV12	99.95	90.18	99.95	90.98	99.56	91.67
aug-cc-pVQZ	99.99	65.76	99.98	93.33	99.96	93.91
aug-cc-pV5Z	100	57.16	100	57.06		
cc-pVQZ-m1	99.9	55.89	99.65	95.22		
cc-pVQZ-m2	99.9	58.32	99.66	99.84		
cc-pVQZ-m3	99.91	57.02	99.66	98.51		
cc-pVQZ-s3	99.91	57.92	99.65	99.31		
cc-pVQZ-s3p1	99.91	57.92	99.67	99.39		
cc-pVQZ-s4	99.91	69.71	99.66	101.1		
cc-pVQZ-s5	99.91	99.8	99.66	101.28		
aug-cc-pVQZ-s6	99.97	59.85	99.83	99.92		
cc-pVQZ-s6	99.91	100.16	99.66	101.56		
cc-pVQZ-s68	99.91	69.66	99.66	100.2		
cc-pVQZ-s1	99.97	54.32	99.94	94.73		
cc-pVQZ-s2	99.97	54.14	99.95	94.77		
cc-pVQZ-s3	99.97	58.56	99.95	99.21		
cc-pVQZ-s4	99.97	99.75	99.95	100.36		
cc-pVQZ-s5s-u	99.91	99.7	99.68	100.15	99.37	99.87
cc-pVQZ-s5s-u	99.98	59.83	99.95	99.5	99.86	99.93
cc-pVQZ-s5s-u	99.99	49.87	99.9	99.93	99.96	99.9
aug-cc-pVQZ-s5s-u	99.95	59.48	99.84	99.65	99.53	99.63
aug-cc-pVQZ-s5s-u	99.95	69.77	99.93	99.81	99.9	99.91
aug-cc-pVQZ-s5s-u	99.99	59.56	99.79	99.93	99.36	99.96
DZP	100.51	52.51	99.79	96.31	99.33	94.25
F7ZP	99.53	104.85	99.88	92.5	99.6	96.19
C-pvman	99.95	55.34	99.67	94.7	99.36	96.21
U-R-2	99.53	100.53	99.67	98.97	99.38	98.61
EPB-3	99.53	100.53	99.84	98.69	99.64	98.3
iplo2	99.96	90.27	99.8	91.2	99.43	92
iplo3	99.99	53.15	99.92	92.7	99.7	93.8
iplo4	99.99	53.15	99.97	93.29	99.85	93

HFCC, Pople-CCSD, percentage with respect to extrapolated UH-CCSD values.

From the cc-pVDZ-sn data, we conclude that one or two tight s functions improve the energy of the hydrogen atom about half of the way to cc-pVTZ results. Further tight s functions have little influence on the energy but increase the calculated HFCC. The effect of tight s functions on the HFCC weakens after the exponents reach the range between 3 000 and 13 000. Finally, with eight tight s functions, the HFCC exceeds the exact value. We attribute this to a slight shift in the electron distribution caused by the additional tight functions, which leads to a slight imperfection of the contraction coefficients. This claim is supported by the HFCC calculated with cc-pVDZ-su8, a completely uncontracted basis set with exponents identical to those of cc-pVDZ-s8. The HFCC calculated with the -su8 basis set is in better agreement with the exact value, and it is smaller than the exact value, as expected from theoretical arguments. In the cc-pVXZ-sn series, the tight s functions have a much smaller effect on the energy of the hydrogen atom. The increase in the spin density tapers off with increasing numbers of tight s functions, but we did not follow the cc-pVXZ-sn series far enough to determine if it exceeds the exact value. The cc-pVXZ-15s-u basis sets are constructed differently. In them, the parent cc-basis set is uncontracted, and five s functions form an even-tempered extension from the tightest regular s function to 999 999. These basis sets give HFCCs that are a little bit lower than those from the corresponding cc-sn basis sets, but the energies are extremely similar. This is also true when basis sets with similar largest exponents are compared. The largest exponent in cc-pVDZ-s8 is ~ 850 000, slightly smaller than in cc-pVDZ-15s-u. The energies are practically identical, but the HFCC calculated with 15s-u is 0.30% too low instead of 0.46% too high. On a related note, the HFCC calculated with cc-pVDZ-15s-u is 0.04% larger than the su8 value, consistent with the slightly larger tightest cc-pVDZ-expo used in cc-pVDZ-15s-u.

#### Testing Basis Sets on $H_2^+$

In the  $H_2^+$  radical cation, polarization functions and the flexibility of the s functions become important. In the hydrogen atom, the s functions did not need to be flexible because the contraction coefficients were optimized for that case. Therefore,  $H_2^+$  is a test for the valence and polarization parts of the basis sets, in addition to the nuclear cusp problem. The wave function of  $H_2^+$  cannot be determined analytically, but energy and spin densities at the nuclei have been determined with high accuracy.

As with the hydrogen atom, the cc-basis sets converge toward limiting values, but the isotropic HFCC's are too low. The sn basis sets give HFCCs between 98.5 and 102% cc-pVDZ of the reference value (%A<sub>iso</sub>, Table 1). The overestimating is stronger than with the hydrogen atom. Adding diffuse functions to the cc-pVDZ-sn basis set improves the calculated HFCCs. The tight p

functions in cc-pVDZ-snpl have very little effect. It is noteworthy that the difference between % A<sub>iso</sub> of  $H_2^+$  and H<sub>2</sub> for cc-pVDZ and the cc-pVDZ-sn and cc-pVYDZ-sn basis sets is about 1.5%. Again, we attribute the convergence toward too high a value to the contraction of the valence functions. Comparing the results obtained using the cc-pVYDZ-s8 and the (uncontracted) cc-pVYDZ-su8 basis sets supports this. The energy is 0.06 mhartree more negative when conclusion: using the -su8 basis set. The HFCC computed from this (energetically) better wave function is only 0.1% higher than the reference value, but the HFCC computed with the -s8 basis set is 2% too high. The cc-pVTZ-sn basis sets behave similarly: increasing the number of tight s functions increases the computed HFCCs at the nuclei. As with the cc-pVDZ-sn basis sets, the sequence of HFCCs converges to a value higher than the reference. The differences between % A<sub>iso</sub> of  $H_2^+$  and H<sub>2</sub> for cc-pVTZ and cc-pVTZ-su8 is about 0.55%. The cc-b-s basis sets show better accuracy. The calculated HFCCs vary between 99.7 and 100.3% of the reference value, and better valence basis sets lead to improved results. As with the aug-cc-pVDZ-s5 basis set, diffuse functions decrease the computed spin density, but the effect is much smaller in the 15s-u basis sets. This lends additional support to the conclusion that contracting the valence part of the hydrogen basis sets is undesirable.

#### H4+

In  $H_4^+$ , electron correlation has to be included. Therefore, correlation functions become important. The energies and HFCCs computed at the UHF-CCSD level with regular cc-basis sets (%A<sub>iso</sub>, Table 1) converge toward the reference values, but the spin densities show deviations between 19 and -6%. The cc-pVXZ-15s-u basis sets improve both energies and HFCCs. The calculated HFCCs are at most 0.4% lower than the reference value and improve with larger valence basis sets. Diffuse functions reduce the calculated HFCCs slightly.  $H_4^+$  has three electrons; therefore, CCSD is not equal to full CI. Engels [4,33] reported that triple excitations are important in ROHF-MRCI calculations of the nitrogen atom. However, Ferreira, Watts, and Bartlett<sup>1990</sup> presented data suggesting that connected triple excitations have little importance in UHF-based coupled-cluster calculations of isotropic HFCCs.

Table (2): Isotropic HFCC in G at the Nuclei of  $H_4^+$

Basis set	UHF-CCSD	UHF-CCSD(T)
cc-pVDZ-15s-u	230.2	230.1
cc-pVTZ-15s-u	230.4	230.3
cc-pVQZ-15s-u	230.5	230.2
aug-cc-pVDZ-15s-u	229.7	229.4
aug-cc-pVXZ-15s-u	230.5	230.1
aug-cc-pVQZ-15s-u	230.5	230.2
reference	230.5	230.5

**Table Q3: Size and Accuracy of Frequently Used Small Basis Sets and Uncontracted Correlation-Consistent Basis Sets with Additional Tight-s Functions**

Basis set	BF/H	av error <sup>a</sup> % ARI	rmsd <sup>b</sup> % ARI	av error <sup>c</sup> %	rmsd <sup>d</sup> %
DZP	5	-5.61	1.55	-21.2	1.9
TZP	9	-1.3	4.43	-10.5	18.1
Chgmn 31111	7	-4.7	0.65	-17.3	6.7
LPR-2	7	-0.6	0.87	-1.4	3.6
EPR-3	10	-0.9	0.91	-2.1	2.6
lglo2	6	-8.8	0.74	-32.4	12.9
lglo3	10	-6.9	0.56	-24.9	12.8
lglo4	19	-6.8	0.08	-21.1	1.8
cc-pVQZ-t5s-u	2	-6.1	0.16	-0.5	0.8
cc-pVTZ-t5s-u	21	-0.1	0.05	0.8	0.3
cc-pVQZ-t5s-u	37	-0.1	0.05	-0.4	0.3
aug-cc-pVQZ-t5s-u	16	-0.4	0.05	-1.6	0.7
aug-cc-pVTZ-t5s-u	30	-0.2	0.05	0.7	0.4
aug-cc-pVQZ-t5s-u	53	-0.1	0.02	-0.4	0.3
cc-pVDZ	5	-26.3	2.02	56.9	15.8
cc-pVTZ	14	-9.0	0.61	-32.9	12.5
cc-pVQZ	30	-6.0	0.77	-22.5	7.1
aug-cc-pVDZ	9	-16.7	1.41	-58.3	14.9
aug-cc-pVTZ	13	-8.1	0.59	-33.2	12.6
aug-cc-pVQZ	46	-6.5	0.22	-22.6	7.1

<sup>a</sup>Average over the unique HFCCs of H<sub>2</sub>, H<sub>2</sub>O, and He<sup>+</sup>. <sup>b</sup>Root-mean-square deviation from the average error.

Note that triple excitations in CI methods are not equivalent to connected triplets in CC theory. The viewpoint of Peraire et al. seems to be supported by the size of the triplets (max 11 = 0.02, max 12 = 0.06) vs 114%. To get an impression, we computed FCI-*pVQZ-15s-cc* spin densities and compared UHF CCSD (them to the corresponding UHF-QCCD results (Table 2). The UHF-CCSD/UHFCCs are about 6.1% lower than the corresponding UCCSD values, indicating that connected triple excitations of triplets provide for the spin density at the nuclei of H<sub>2</sub>H.

Consequently, the Coulomb Hyperfine Coupling Constants in *cc*, *DZP*, and *TZP*

to have confidence in the results of predictive calculations, the accuracy of a given basis set combination needs to be consistent over a broad variety of compounds. To judge how consistent the HFCCs calculated with various basis sets are, we determined the average error and the root-mean-square deviation (rmsd) for the average error. Because a constant error can be avoided by scaling the computed values, the rmsd from the average error is the more important quantity. To provide a comparison, we did the same calculations with some basis sets that have been used in previous studies of isotropic HFCCs (Table 3). Most basis sets have a rmsd from the average error of between 0.6 and 1.1% of the reference value. DZP has a slightly larger rmsd, but it also is the smallest basis set in the field. TZP is larger

and has a small average error, but the rmsd is the largest in the field. The LPR basis sets shine with small average errors and average rmsds. The Chgmn, lglo2, and lglo3 basis sets have slightly smaller rmsds but larger average errors. The lglo basis set has an exceptionally small rmsd, but it is a large basis set. The cc-pVQZ-t5s-u basis sets perform better than the used basis sets. They combine a *unit* from the average error that is comparable to that of the lglo basis set with average errors of -0.4% or less. Two of the cc-pVQZ-t5s-u basis sets are smaller, and one has a size similar to that of lglo4.

**Table (4): Influence of Basis Sets at Neighboring Atoms: Percentage of the Reference Isotropic HFCC at the Nuclei of H<sub>2</sub><sup>+</sup>, and H<sub>4</sub><sup>+</sup>.**

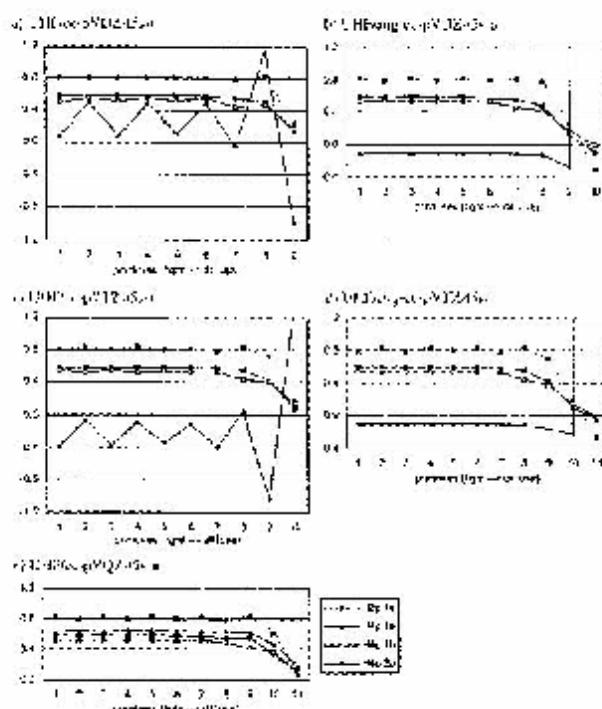
Basis sets	All cc-t5s-u	mixed
Nucleus of interest		
cc-pVDZ-t5s-u	H <sub>2</sub>	H <sub>4</sub> <sup>+</sup>
cc-pVTZ-t5s-u	100.15	99.85
cc-pVQZ-t5s-u	99.90	99.93
aug-cc-pVDZ-t5s-u	99.93	99.97
aug-cc-pVTZ-t5s-u	99.83	99.89
aug-cc-pVQZ-t5s-u	99.92	99.94

### Influence of the Basis Set at Neighboring Atoms

Earlier, we said that the influence of valence and diffuse functions of neighboring atoms would be secondary to the effect of tight s primitives. The data in Table 4 show that this is true for at least the inner-valence functions. Replacing the cc-t5s-u basis sets at the neighboring atoms with regular cc- basis sets changes the calculated iHFCCs by 0.3% or less of the reference values. The effect is stronger with the cc-pVDZ basis set than with the more flexible cc-pVTZ and cc-pVQZ basis sets. The performance of cc-pVXZ-t5s-u basis sets for hydrogen together with regular cc- basis sets for heavier atoms will be explored by using the example of small organic radicals in the following article [34].

### Contracting the Basis Sets

We feel that it is important to choose the contraction pattern such that the contracted basis sets describe several situations equally well. In general, these could be ionized and/or (valence) excited atomic states as in the MEFIT procedure used by the Söhl and Prouss group in the development of LCP basis sets. In the case of hydrogen, however, the ions are closed-shell species and therefore ESR-inactive. We will instead analyze how the individual basis functions contribute to the MOs in H<sub>2</sub>, H<sub>3</sub><sup>+</sup>, and H<sub>4</sub><sup>+</sup>. Figure (1) presents the MO coefficients of s primitives in the occupied orbitals in H<sub>2</sub><sup>+</sup>, and H<sub>4</sub><sup>+</sup> relative to those of the isolated atom. The coefficients of the tighter functions exhibit a quite regular oscillating pattern. This suggests an alternating contraction for the tightest primitives (e.g., primitives 1, 3, and 5 in one and primitives 2, 4, and 6 in a second contraction). Higher angular momentum functions remain as in the parent cc basis sets. Basis sets employing this alternating contraction scheme will have a suffix of t5s-an, with n being the number of s functions. However, in the presence of diffuse functions, the oscillations are not very strong. Therefore, we will also test a conventional, continuous contraction scheme. Those basis sets will have suffixes of t5s-cr, indicating a s functions. The pattern begins to break down at primitive 7 in the DZ basis sets and at primitive 8 in the TZ and QZ basis set.



**Figure (1): MO coefficients of the primitives at hydrogen relative to those of the isolated atom.** (a) UHF/cc-pVDZ-t5s-u, (b) UHF/aug-cc-pVDZ-t5s-u, (c) UHF/cc-pVTZ-t5s-u, (d) UHF/aug-cc-pVTZ-t5s-u, and (e) UHF/cc-pVQZ-t5s-u.

The energies and HFCCs calculated with the different contracted cc-pVXZ-t5s basis sets are collected in Table 5. It also contains the average error (as a percentage of the reference value) and the rmsd from the average error. The t5s-a(v+3) basis sets give results that are virtually identical to those of the uncontracted basis sets ( $v = \text{number of valence (s) functions}$ ,  $v = 2$  for cc-pVDZ, and  $v = 3$  for aug-cc-pVDZ and cc-pVTZ). Contracting one more s primitive causes a slight increase in the rmsd. Here, the alternating contraction scheme is better for the DZ basis sets, and the single contraction scheme is better for the TZ basis sets. Contracting enough s primitives to regain the size of the original cc- basis sets produces basis sets that have significantly lower average errors but somewhat larger rms deviations from the average error than the regular cc- basis sets. Because of the

strong variations in the valence part of the basis sets, at this degree of contraction the continuous contraction scheme is superior.

### **Summary and Conclusions**

Adding tight's functions to one-centered easily sets the tightly exercises the average error in the isotropic coupling constants (CC<sub>II</sub>) at the nuclei of H<sub>1</sub>, H<sub>2</sub>, and H<sub>3</sub><sup>1</sup> calculated by TCF or LIP-CCSD. More importantly, it reduces the error from the average error to a 2% or less. Contracting the inner six to eight p-orbitals is an alternative, whereby creating basis sets with identical  $\alpha$ , but slightly reduced accuracy, increasing the size of the valence part of the basis set improves the accuracy of the results, thereby allowing extrapolation to the complete basis set limit. The best size/accuracy ratios are achieved by cc-pVQZ (1.34), cc-pV5Z (1.45), and cc-pVQZ (1.45), and cc-pVQZ (1.36).

Small- to medium-size basis sets used in previous studies of tetrapic hyperfine coupling constants could have rms deviations from the

average error of  $\pm 2\%$  between 0.6 and 11.6. The somewhat large Iglo1 basis set is very consistent with a rmsd of 0.1%. The LPR basis sets combine a rmsd of 1% with an average error of  $\pm 1\%$ . Chipman's 3111+1 as well as the Iglo2 and Iglo3 basis sets have smaller rms deviations of 0.6 or 0.7%, but the average errors are between  $\pm 5$  and  $\pm 9\%$ . The DZP basis set has a relatively large rms deviation of 1.8%, but it is also the smallest basis set considered. The TZ2P basis set has a small average error, but the rmsd is large. The influence of the basis sets on the glisting atoms is relatively small. Using regular cc-pVQZ basis sets for the surrounding atoms changes the calculated HFCCs by less than  $\pm 3\%$ . The differences are smaller than 0.1% when TZ basis sets are used. We hope that the superior accuracy of the cc-pVQZ basis sets is not significantly reduced when they are used in conjunction with regular cc-p basis sets on first row atoms. This will be investigated in the following article on small organic radicals, together with the influences of the optimization method and correlation methods used in calculating HFCCs [34].

Table S5: Percentage of the Reference Energy and Isotropic HFCC ( $\Delta_{\text{HF}}$ ) of the Nuclei of  $\text{H}_2$ ,  $\text{H}_2^+$ , and  $\text{H}_2^+$  Reconstructed Correlation-Consistent Basis Sets with Additional Tight's Functions.

We use series of well-tempered basis sets to establish trends in energy and EJKCs and extrapolate from them. We use the following

non-terminal; or  $\beta$ ;  $\gamma$  denotes a non-terminal basis set;  $\Gamma^0$  indicates the subsystems are separated by a fixed cut-off;  $\omega$ ;  $\omega_1$  at c.  $\pm 160$ – $205$  rad/s;  $\zeta$  that one of the s-

exponents is 1.00, with 20 tighter and 5 more diffuse functions. The first step is creating a set of  $s$  functions at the hydrogen atom. As a first try, we created the harm-f2-s100 basis sets. As seen in Table 6, they give excellent results. A comparison shows that adding extremely tight functions improves the HFCC but not the energy. Adding extremely diffuse functions improves the energy but has no effect on the coupling constant. We crudely optimized the exponents of seven-membered harmonic basis sets with factors of 2 and 3 by multiplying the exponents with the respective factor to the power of 3/4 or 5/4. Because the optimal base exponent for the factor of 2 basis set was found to be 0.95, we did not create a basis set from it but kept the harm-f2-s100 basis sets. The optimal base exponent for the factor of 3 basis set was found to be 2.50. The factor of 3 basis set was expanded by adding tighter and more diffuse functions so that either a diffuse function or a group of tight functions results in a similar energy gain. Table 6 shows that the s250-14-8 energy is very good but not as close to perfection harm-f3 as the harm-f2-s'00 basis sets. The calculated isotropic HFCC is very good. A comparison between the large harm-f3 basis sets shows that extremely tight functions are necessary to achieve high accuracy in the HFCC. The harm-f3-209 basis sets do not deviate from the factor of 3 for the tightest basis function and are slightly lower in energy than the corresponding harm-f2-250 basis sets. The best harm-f3 basis is 0.000017 hartree or 0.0035% higher than the exact result, but the best harm-f2 basis set is practically perfect.

**Table (6): Percentage of the Exact Energy and Isotropic Hyperfine Coupling Constant at the Nucleus of Hydrogen: Harmonic Basis Sets**

Isotropic HFCC	Exact exp.	HFCC exp.	% energy	% HFCC
Harm-f2-s100-20-5	0.999999	0.999999	99.99	99.99
Harm-f2-s100-20	0.999999	0.999999	99.99	99.99
Harm-f2-s100-20-5	0.999999	0.999999	99.99	99.99
Harm-f2-s100-20-10	0.999999	0.999999	99.99	99.99
Harm-f2-s250-14-5	0.999999	0.999999	99.99	99.99
Harm-f2-s250-14-10	0.999999	0.999999	99.99	99.99
Harm-f2-s250-14-15	0.999999	0.999999	99.99	99.99
Harm-f2-s250-14-20	0.999999	0.999999	99.99	99.99
Harm-f2-s250-14-25	0.999999	0.999999	99.99	99.99
Harm-f2-s250-14-30	0.999999	0.999999	99.99	99.99
Harm-f2-s250-14-35	0.999999	0.999999	99.99	99.99
Harm-f2-s250-14-40	0.999999	0.999999	99.99	99.99
Harm-f2-s250-14-45	0.999999	0.999999	99.99	99.99
Harm-f2-s250-14-50	0.999999	0.999999	99.99	99.99
Harm-f2-s250-14-55	0.999999	0.999999	99.99	99.99
Harm-f2-s250-14-60	0.999999	0.999999	99.99	99.99
Harm-f2-s250-14-65	0.999999	0.999999	99.99	99.99
Harm-f2-s250-14-70	0.999999	0.999999	99.99	99.99
Harm-f2-s250-14-75	0.999999	0.999999	99.99	99.99
Harm-f2-s250-14-80	0.999999	0.999999	99.99	99.99
Harm-f2-s250-14-85	0.999999	0.999999	99.99	99.99
Harm-f2-s250-14-90	0.999999	0.999999	99.99	99.99

\*Deviation from a factor of 3. Deviation of 1.64N 5.6%.

Deviation from a factor of 3 should be 11.987 442.5.

**Table (7): Energy (hartree) of H<sub>2</sub>: Harmonic Basis Sets.**

Basis set	Energy
ang-co-pV5Z	-1.1721
extrapolated (ref 17-co-pVXZ, XYZ, Q, S)	-1.17447
extrapolated (ref 35, ang-co-pVXZ, XYZ, Q, S)	-1.17449
harm-f3-s209-14-9	-1.15482
harm-f3-s209-14-3	-1.15464
harm-f3-s209-14-1-1-p171-11-1	-1.15451
harm-f3-s209-14-3-p062-1-1-1-17195	-1.15450
harm-f3-s209-3-3-p0711-2-1	-1.17209
harm-f3-s209-3-3-p053-3-1	-1.17202
harm-f3-s209-3-3-p063-3-1-d209-0-1	-1.17265
harm-f3-s209-3-3-p063-3-1-d121-1-1	-1.17311
harm-f3-s209-3-3-p061-3-1-d053-2-1	-1.17373
harm-f3-s209-3-3-p053-3-1-d053-2-1-d275-0-1	-1.17361
harm-f3-s209-3-3-p053-3-1-d053-2-1-d159-1-1	-1.17404
harm-f3-s209-14-9-p053-3-1-d053-2-1-d159-1-1	-1.17419
harm-f2-s100-20-5	-1.17439
harm-f2-s100-20-5-p100-10-5	-1.17229
harm-f2-s100-20-5-p100-4-5-d100-2-4	-1.17337

**Table (8): Isotropic HFCC of H<sub>2</sub><sup>14</sup>: Harmonic Basis Sets**

Basis set	Energy
harm-f3-s209-3-3	310.6
harm-f3-s209-3-3-p053-3-1	308.0
harm-f3-s209-3-3-p053-3-1-d053-2-1	307.9
harm-f3-s209-3-3-p053-3-1-d053-2-1-d159-1-1	307.9
harm-f3-s209-14-9	332.0
harm-f3-s209-14-9-p053-3-1-d053-2-1-d159-1-1	334.7
harm-f3-s209-11-9-p053-3-2-d053-2-2-d159-1-1	334.7
harm-f3-s209-14-9-p053-3-3-d053-2-3-d159-1-1	334.7
harm-f2-s100-20-5	334.9
harm-f2-s100-20-5-p100-10-5	334.6
harm-f2-s100-20-5-p100-4-5-d100-2-4	334.6
exact**	334.8

\*Ref 25, interpolated in MOLSTA.

The next step is the optimization of polarization/ correlation functions. This is done with CCSD (here equal to full CI) calculations on H<sub>2</sub>. To keep the size manageable, we augment the harm-f3-209-3-3 basis set and will later add the tight  $s$  functions for ESR calculations. First, two  $p$  functions are added with exponents equal to the  $s$  exponents with the highest coefficients. During the optimization, the exponents are multiplied by 3<sup>3/4</sup> or 3<sup>5/4</sup>. Then, a tighter  $p$  function is repeatedly added, and the  $p$  exponents are optimized again. When additional  $p$  functions fail to reduce the energy significantly, we continue with  $d$  functions and so on. As can be seen from Table 7, the largest harm-f3 basis set is nearly as good as aug-co-pV5Z. For comparison, we created harm-f2 basis sets with polarization functions starting approximately as those in the optimized basis sets and continuing beyond the diffuse limit of the harm-f3 basis sets. Their energies are a little bit lower than those harm-f3s of the corresponding harm-f2 basis sets, indicating that both increasing the density of functions and adding more diffuse polarization functions improve the energy a little bit.

Table 9: Isotropic HFCCs of  $H_2^+$  in G.

Basis set	$H^-$		$H_2^+$		$H_2^+$		av.
	cc-pV	ang-cc-pVQZ	cc-pV	ang-cc-pV	cc-pV	ang-cc-pV	
DZ-15s-n	505.19	505.09	334.75	334.59	230.23	230.57	
TZ-15s-n	506.87	506.57	334.73	334.09	230.41	230.32	
QZ-15s-n	507.29	506.96	334.39	334.57	230.50	230.15	
SZ-15s-n			334.46	334.66			
extr. 1'	507.79	507.30	334.46	334.67	230.59	230.57	230.58
extr. 2 (D,T,Q)	507.26	506.58	334.00	334.67	230.55	230.5	230.53
extr. 2 (T,Q,S)			334.62	334.63			
DZ-15s-n <sup>a</sup>					230.55	230.44	
TZ-15s-n <sup>a</sup>					230.6	230.05	
QZ-15s-n <sup>a</sup>					230.23	230.17	
SZ-15s-n <sup>a</sup>					230.30	230.25	230.30
extr. 1'					230.36	230.23	230.28
extr. 2'							
reference	504.71		334.8 ± 0.3				
CCSD(cc-pVQZ) + (cc-pVQZ, 15s-n, D, T, Q, S, TZ)							

The next step is to establish trends in the calculation of HFCCs at the level of  $H_2^+$  (Table 8). The right-s functions have a large effect; they increase the coupling constant by 76.46 (~ 8%). Adding p functions to the basis-15s-200-5-3 basis set reduces the coupling constant by 2.6 G (~ 1%). The effect of d and f basis sets on the HFCC is extremely small. Adding polarization functions to the basis-15s-200-14-9 or basis-15s-100-20-5 basis set reduces the HFCCs by ~ 3 G. The first value for the basis-15s-200-14-9 basis set is 0.1 G larger than the first value for the basis-15s-80-20-5 basis set. This is similar to the results with hydrogen where the right-s functions of the basis-15s-100-14-9 basis sets have slightly larger and better coupling constants than do basis-15s-80-20-5 basis sets. We conclude that the isotropic hyperfine coupling constant in  $H_2^+$  is 334.8 (± 0.3) G. We can compare this with the results of Stephens and Aufrey [25], they calculated an HFCC of 334.8 G at the nuclear distance of two Bohr. Interpolation [26] yields a value of 334.8 G for 1.057 Å.

Because of the size of the harmonic constants, we cannot use the largest ones for  $H_2^+$ . Conveniently, for both  $H^-$  and  $H_2^+$ , the spin densities calculated with the cc-pVQZ-15s-n basis sets converge rapidly toward the reference values. In both cases, adding the difference from TZ to QZ to the QZ spin densities yields results that are within 0.1% of the reference values. Doing so for  $H_2^+$  gives hyperfine coupling constants of 230.58 G when using CCSD(CCSD) results and 230.39 G when using D spin densities. Using the extrapolation scheme (III-CCSD(T)) of Pernaldo, Wenz, and Domke [37] results in isotropic HFCCs of 230.63 and 230.28 G for CCSD and CCSD(CCSD(T)), respectively. These numbers are the averages of extrapolations using cc-pVQZ-15s-n and ang-cc-pVQZ-15s-n basis sets because

they should become equivalent in the limit of infinite basis sets. The difference between the CCSD and CCSD(CCSD(T)) values is very small. Because all basis set tests on  $H_2^+$  were performed at the CCSD level, we used the CCSD values as references. We consider the extrapolated CCSD(CCSD) results to be nearly exact, because the CCSD(CCSD) energies with the cc-pVQZ (X) D, T, Q basis sets are just 70 to 80 µhartree higher than the FG results.

#### Acknowledgment:

Abd R. A. Deniz is indebted to the William Fulbright Foreign Scholarship Board (grant no. 24319).

#### References

1. Sekine, H.; Barfield, K. J. J. Chem. Phys. 1985, 82, 4225.
2. Veltman, W., Jr. In: *Magnetic Atoms and Molecules*; Dover Publications, New York, 1989; p. 72, 1. p. 30, 1. (CH<sub>3</sub>Cl<sub>2</sub>) p. 4, 14, appendix A, A. A. (H<sub>2</sub>O) = -2.823<sup>8</sup> Δ<sub>0</sub>(G) and (C) = -285.367 μCN<sub>2</sub>10N<sub>2</sub>μsibeh<sup>8</sup>.
3. Pernaldo, S. A.; Wenz, J. D.; Deniz, R. J. J. Chem. Phys. 1994, 100, 1425.
4. Imre, B.; Pernaldo, S. A.; Lutz, S. In: *Advances in Quantum Chemistry*; Lovell, P. O.; Schatz, J. R.; Zutty, M. C., Eds.; Academic Press, San Diego, CA, 1996, Vol. 27, p 277.
5. Chapman, D. M. *Understanding Chem. React.* 1995, 17, 69.
6. Savone, V. In: *Recent Advances in Density Functional Methods*; Cheng, D. Y., Ed.; World Scientific: Singapore, 1993; Part I, p 287.
7. Heller, D.; Davidson, E. R. In: *Molecular Spectroscopy, Electronic Structure, and Intermolecular Interactions*; Springer-Verlag: New York, 1991; p 429.

7. Hirshfeld, E. Z. Int. J. Quantum Chem. 1993, 42, 689.
8. Heller, J.; Rehrer, J.; Feinberg, G. Phys. Rev. A 1998, 58, 2399.
9. Bassolov, V. A.; Chipman, D. M. J. Chem. Phys. 1996, 105, 1470.
10. Plessner, R. W. J. Phys. Chem. 1867, 21, 74.
11. Kondo, H.; Morikawa, K. Chem. Phys. Lett. 1971, 12, 403; Kondo, H.; Morikawa, K. J. Am. Chem. Soc. 1922, 44, 3623; Tarterelli, L. G.; Fairbank, J. H. J. Chem. Phys. 1928, 63, 343.
12. More exactly, the sphere where the second derivative of the wave function goes through zero (the "harpoon sphere") decreases in radius.
13. Perea, S. A.; Salen, J. M.; Viteri, J. J. Chem. Phys. 1997, 106, 7001.
14. Chapman, D. M.; Curchael, L.; Chipman, D. M. J. Phys. Chem. A 1997, 101, 4253.
15. Bartlett, R. J. In Modern Electronic Structure Theory; Marques, D. R., Ed.; World Scientific Publishing Singapore: 1995, Part I, p 13–7.
16. Chipman, D. M. J. Phys. Chem. A 1999, 103, 5435; Chipman, D. M. Phys. Chem. Acta 1969, 56, 73.
17. Heller, J. In *Modern Electronic Structure Theory*; Marques, D. R., Ed.; World Scientific Publishing Singapore: 1995, Part I, p 13–7.
18. Chipman, D. M. J. Phys. Chem. A 1999, 103, 5435; Chipman, D. M. Phys. Chem. Acta 1969, 56, 73.
19. Heller, J. J. Chem. Phys. 1995, 93, 579; (b) Heller, J.; Glendinning, F. D.; McCullough, E. A.; Miller, R. J. J. Chem. Phys. 1995, 99, 1329.
20. Louie, S. M. In *Local Additive Local Density Functional Methods*; Chang, T. M., Ed.; World Scientific: Singapore, 1995, Part I, p 287; Pergamon Press: Oxford, U. K. Chem. Phys. 1995, 105, 1103.
21. Heller, J.; De Jong, F. R. J. Chem. Phys. 1998, 108, 7530.
22. Kiskin, A. B.; Ing-Rhee, M. T.; Mavrikis, I.; Plessner, R. W.; Bartlett, R. J. J. Chem. Phys. 1998, 109, 403.
23. Leibler, K. J. Phys. Soc. Jpn. 1969, 25, 1705.
24. Li, Z. W.; Hevesi, L.; Zeng, L. J. R. Phys. Soc. Jpn. 1992, 61, 1551.
25. Neophytou, M. J.; Anthony, J. P. J. Chem. Phys. 1995, 103, 1329; Neophytou, M. J.; Anthony, J. P. J. Chem. Phys. 1995, 103, 1335.
26. El-Shabani, N. J.; Gilbert, J. R.; Powers, M. J. Chem. Phys. Lett. 1994, 105, 7.
27. Huang, Y.; Shi, Wang, Y.; Guo, Y. L.; Gong, A. S.; Yang, Y. Yu Xue Za Zhi 2001, 18, 133.
28. ADF2001 is a program product of the Quantum Theory Project, University of Florida, authors: Seminario, J. F.; Gross, E.; Wall, J. D.; Kosanovic, M.; Olfert, S.; Peter, S. A.; van der Goot, G.; Forderdak, W. J.; Kuchan, M.; S. A.; Cowalney, S. R.; Beck, S.; Balkov, A.; Tsandolo, D. L.; Brune, L. R.; Rodriguez, P.; Sekula, H.; Faber,
29. Bartlett, R. J. Integrated packages included are VMD (Amber), J. J. Taylor, P. R.), MMDB (Taylor, P. R.); AIM3L (Hegelich, L.; Bauer, L. J. A.; Jorgensen, P.; Olsen, J.; Taylor, P. R.).
30. Gross, E.; Stanton, J. F.; Bartlett, R. J. J. Chem. Phys. 1991, 95, 2659; Gross, E.; Forderdak, W. J.; Stanton, J. F.; Weller, J. D.; Bartlett, R. J. Chem. Phys. Lett. 1991, 172, 207; Wall, J. D.; Gross, E.; Bartlett, R. J. Chem. Phys. Lett. 1992, 213, 1; Weller, J. D.; Stanton, J. F.; Bartlett, R. J. J. Chem. Phys. 1992, 98, 87–8.
31. Dantzig, T. L.; Jr. J. Chem. Phys. 1999, 93, 1367; Kewell, R. A.; Dantzig, T. L.; Jr.; Hansen, K. J. J. Chem. Phys. 1997, 96, 5769.
32. Plessner, R. W.; Schley, P. H. J. Chem. Phys. 1963, 39, 217.
33. Higuchi, T.; Jusaki, M.; Kudo, T.; Goto, K.; Hori, Chem. Phys. 1998, 92, 175.
34. Engholm, B. *Theor. Chim. Acta* 1992, 86, 429.
35. Al-Derzi, A. R.; Louie, S. M.; Bartlett, R. J. J. Phys. Chem. A 2001, 105.
36. We calculated a 41-term polynomial in  $\Gamma$  to the value of  $\gamma_1^{\text{eff}}$  of 15 (1, 2, 3, 4, 5 bonds) in  $\Gamma$ . We calculated the variance in  $\gamma_1^{\text{eff}}$  for 2 bonds at 1.027 Å and added it to the value at 2 bonds.
37. Plessner, R. W.; Woods, D. R.; Chatterjee, T. J. J. Chem. Phys. 1991, 96, 2115 (eq. 2).