JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Spin Component-Scaled Second-Order Møller−Plesset Perturbation Theory for Calculating NMR Shieldings

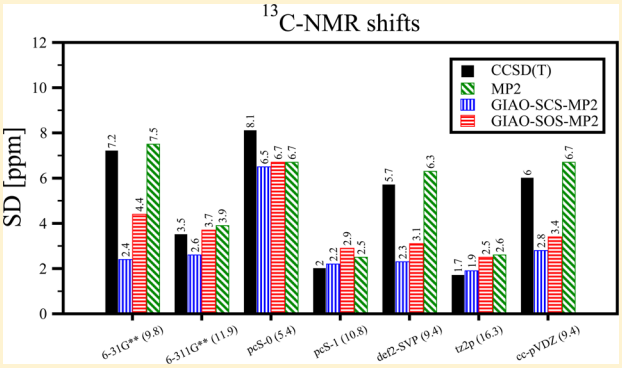
Marina Maurer and Christian Ochsenfeld*

Chair of Theoretical Chemistry, Department of Chemistry, University of Munich (LMU), Butenandtstrasse 7, D-81377 Munich, Germany

Center for Integrated Protein Science (CIPSM) at the Department of Chemistry, University of Munich (LMU), Butenandtstrasse 5−13, D-81377 Munich, Germany

Supporting Information

ABSTRACT: Spin component-scaled and scaled opposite-spin second-order Møller−Plesset perturbation approaches (SCS-MP2 and SOS-MP2) are introduced for calculating NMR chemical shifts in analogy to the well-established scaled approaches for MP2 energies. Gauge-including atomic orbitals (GIAO) are employed throughout this work. The GIAO-SCS-MP2 and GIAO-SOS-MP2 methods typically show superior performance to nonscaled MP2 and are closer to the coupled-cluster singles doubles perturbative triples (CCSD(T))/cc-pVQZ reference values. In addition, the pragmatic use of mixed basis sets for the Hartree−Fock and the correlated part of NMR chemical shift calculations is shown to be beneficial.



1. INTRODUCTION

Quantum-chemical calculations of nuclear-magnetic resonance (NMR) data have become important tools for the often complicated interpretation of experimental NMR spectra and in this way for gaining reliable structural information. While Hartree−Fock (HF)[1−3] or density-functional theory (DFT)[4,5] methods are quite efficient, more accurate and reliable results for NMR shieldings can be obtained by wave function-based correlation methods like Møller−Plesset second-order perturbation theory (MP2)[6−9] or coupled-cluster (CC) theory.[10−12] The importance of using reliable quantum-chemical methods cannot be overemphasized, since fully empirically parametrized approaches typically lack sensitivity to structural changes.[13,14]

In order to improve the applicability of quantum-chemical NMR calculations to large molecular systems, linear- or even sublinear-scaling methods have been developed at HF and DFT levels, so that nowadays NMR spectra for molecules with more than 1000 atoms can be studied on simple workstation computers.[15−17] In contrast, the formally steep $O(M^5)$ or at least $O(M^6)$ scaling with molecular size (M) of MP2 and CC methods, respectively, is still the limiting factor with respect to the size of the molecules which can be treated. Nevertheless, in recent years large efforts have been made for reducing the $O(M^5)$ scaling for calculating NMR shieldings at the MP2 level: For improving efficiency Gauss and Werner[18] introduced a local-correlation approximation for molecular-orbital (MO) based MP2 shieldings that was recently employed by Loibl and Schütz[19] for an efficient implementation and combined with a density fitting scheme. A different pathway was chosen by us in developing a linear- or even sublinear-scaling formulation for MP2-NMR shieldings.[20] Here, we employ the atomic orbital (AO) basis and developed a pilot implementation that also provides evidence for the possibility of sublinear-scaling of MP2 theory for nuclei-selected NMR shieldings.[20] Motivated by our recent AO-based SOS-MP2 energy calculation for a DNA repair system with 2 025 atoms and 20 371 basis functions,[21] we are currently working on an efficient implementation which will allow us to obtain NMR properties of similarly large systems.

These developments show the increasing possibilities for performing MP2 calculations on large molecules, so that the question arises if it is possible to improve the accuracy of MP2 NMR chemical-shift predictions by employing spin-component scaling similar to the scaling schemes introduced for MP2 energies by Grimme[22] and also employed by others[23−25] in scaling the same and opposite spin components of the MP2 energy differently. The original spin component scaled (SCS) MP2 energy method of Grimme, where the opposite spin component is scaled by a factor of 1.20 and the same spin component by a factor of 0.33 shows great benefits for reaction energies, barrier heights, geometries, and harmonic vibrational frequencies.[22] An alternative was introduced by Head-Gordon and co-workers[24] as a simplification of the SCS-MP2 method with the so-called scaled opposite spin (SOS) MP2 method. Here, the same spin part is entirely neglected and only the opposite spin part of the MP2 energy scaled by a factor of 1.30.[24]

Received: March 24, 2014
Published: December 22, 2014

   dx.doi.org/10.1021/ct5007295 | J. Chem. Theory Comput. 2015, 11, 37−44

The gain of accuracy of SOS-MP2 for reaction and atomization energies is similar to that of SCS-MP2, but the great advantage is the huge efficiency improvement. By combining the MP2 method with the resolution of identity approximation[24] and calculating only the opposite spin part of the MP2 energy, the scaling behavior of the computational cost reduces from $O(M^5)$ to $O(M^4)$. While the different scaling of the antiparallel and parallel spin component was also applied to the CIS(D)[26,27] and CC2[28] methods for excited states, no studies have been published for calculating MP2 NMR shieldings so far.

In our present work, we introduce spin component-scaled MP2 schemes for calculating NMR chemical shifts in analogy to the well-established SCS- and SOS-approximations for MP2 energies.[22−25] We always employ the gauge-including AO (GIAO)[1,29,30] formulation that has proven to be most efficient for the calculation of NMR shieldings.[2,31] In analogy to approximations for MP2 energies, we denote our schemes for NMR calculations by GIAO-SOS-MP2 and GIAO-SCS-MP2, respectively. We tested our new methods for MP2 based carbon, phosphorus, nitrogen, oxygen, and fluorine NMR shifts. To evaluate the accuracy of the carbon NMR shifts, we employ our recently introduced benchmark set.[32] For the other nuclei, new benchmark sets are presented. After introducing the equations for optimizing scaling factors for NMR shieldings, we present first results for carbon NMR shifts employing a triple-ζ basis set (tz2p[33,34]). Besides triple-ζ results, we also show the benefits of the scaled MP2 NMR method for relative carbon shifts calculated with medium-sized and also smaller basis sets. Furthermore, we present another pragmatic approach for increasing the accuracy of, e.g., carbon NMR chemical shifts by combining large basis HF results with the electron-correlation part described at MP2 or CCSD(T) levels using smaller basis sets. In the last chapter, the gain of accuracy is presented for phosphorus, nitrogen, oxygen, and fluorine MP2 NMR shifts.

2. OPPOSITE AND SAME SPIN TERMS OF MP2 NMR SHIELDINGS

Similar to ground state energies, MP2 based NMR shieldings can be separated into the HF contribution, and into the opposite spin (OS) and same spin (SS) terms of the perturbative second-order correction:

$$\sigma_{\text{SCS-MP2}} = \sigma_{\text{HF(basis set of SCS-MP2)}} + c_{\text{OS}}\,\sigma_{\text{OS}} + c_{\text{SS}}\,\sigma_{\text{SS}} \quad (1)$$

To obtain the scaling factors c_{OS} and c_{SS}, we fit the (absolute) MP2 NMR shieldings in a least-squares procedure to CCSD(T) results. In addition, we correct for systematic deviations of the (absolute) shieldings and therefore add a constant offset MSD_{opt}:

$$\sigma_{\text{SCS-MP2}} + \text{MSD}_{\text{opt}} \overset{!}{=} \sigma_{\text{CCSD(T)}} \quad (2)$$

By optimizing the absolute shieldings with the systematic deviation MSD_{opt} instead of relative shifts, the resulting scaling factors are independent of the chosen reference like, e.g., TMS. Since relative shifts are the difference of the scaled, e.g., TMS value and a scaled absolute shielding of an other molecule, the MSD_{opt} value vanishes for relative shifts.

Furthermore, we also split the CCSD(T) based shielding tensor into its HF and correlation contributions

$$\sigma_{\text{CCSD(T)}} = \sigma_{\text{HF(basis set of CCSD(T))}} + \sigma_{\text{CCSD(T)-corr}} \quad (3)$$

and insert eqs 1 and 3 into eq 2

$$c_{\text{OS}}\,\sigma_{\text{OS}} + c_{\text{SS}}\,\sigma_{\text{SS}} + \text{MSD}_{\text{opt}} \overset{!}{=} \sigma_{\text{CCSD(T)-corr}} + \underbrace{\sigma_{\text{HF(basis set of CCSD(T))}} - \sigma_{\text{HF(basis set of SCS-MP2)}}}_{\Delta\sigma_{\text{HF}}} \quad (4)$$

Normally, the HF contributions in the CCSD(T) shieldings $\sigma_{\text{HF(basis set of CCSD(T))}}$ and in the MP2 shieldings $\sigma_{\text{HF(basis set of SCS-MP2)}}$ are the same, if the same basis set is used. Here, however, we use different basis sets for the two methods. Therefore, a deviation occurs because of different HF-based NMR shieldings included in the correlation methods, which plays an additional role in our fitting procedure. We abbreviate this HF difference by $\Delta\sigma_{\text{HF}}$:

$$c_{\text{OS}}\,\sigma_{\text{OS}} + c_{\text{SS}}\,\sigma_{\text{SS}} + \text{MSD}_{\text{opt}} \overset{!}{=} \sigma_{\text{CCSD(T)-corr}} + \Delta\sigma_{\text{HF}} \quad (5)$$

Since we aim for scaling factors for the entire chemical shift, we have included the HF difference throughout the fitting procedure. We have found this empirical approach to be pragmatic and useful, as illustrated below. An alternative approach would be to scale only the correlation parts. We have tested this variant for carbon NMR shifts and observe, as expected, a fit of a lower quality. The corresponding data and discussion can be found in section 4.2.

3. METHODOLOGICAL ASPECTS

To optimize the scaling factors for our GIAO-SCS-MP2 and GIAO-SOS-MP2 methods, for carbon NMR shifts we apply our benchmark set (see Figure 1) introduced by Flaig et al.,[32] which

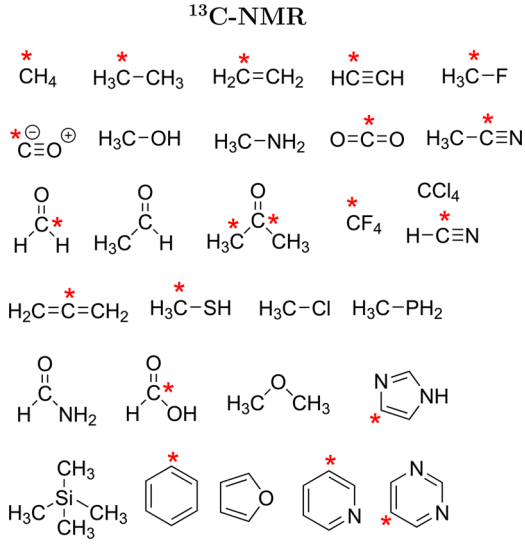


Figure 1. Molecular benchmark set of our study in ref 32. The carbon nuclei of the fitting set are labeled with a red star.

comprises various organic molecules with carbon NMR shifts spanning a broad range of the NMR scale. For the other nuclei, new benchmark sets are presented in Figures 3−6. The structures of these benchmark sets were optimized at the CCSD(T)/cc-pVTZ[35] level by using the program package CFOUR.[36] All geometries are available via the web site http://www.cup.uni-muenchen.de/pc/ochsenfeld/download.html and also in the Supporting Information. The reference data at the CCSD(T)/cc-pVQZ level were calculated with the program package CFOUR.[36]

dx.doi.org/10.1021/ct5007295 | J. Chem. Theory Comput. 2015, 11, 37−44

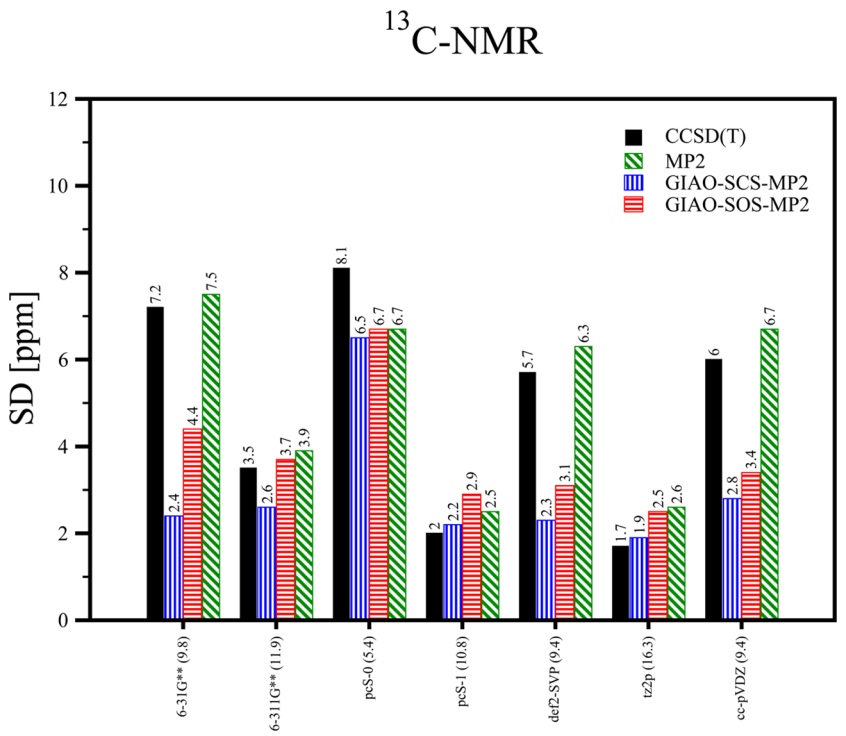


Figure 2. Standard deviation (SD) of ^{13}C chemical shifts for the GIAO-SCS-MP2 and GIAO-SOS-MP2 methods in comparison to nonscaled MP2 and CCSD(T) results (see also the work of Flaig et al.[32]) are listed for various basis sets. As a reference the CCSD(T)/cc-pVQZ data is employed. Data for the total benchmark set are shown, while the fitting coefficients for GIAO-SCS-MP2 and GIAO-SOS-MP2 have been obtained for the 20-carbon shieldings fitting set. As a measure for the different basis set sizes, the average number of basis functions per atom of the molecular benchmark set is listed within parentheses.

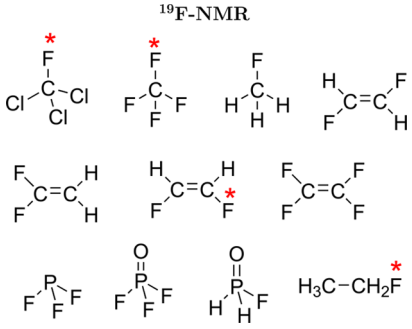


Figure 3. Molecular benchmark set for fluorine NMR shifts. The fluorine nuclei of the fitting set are labeled with a red star.

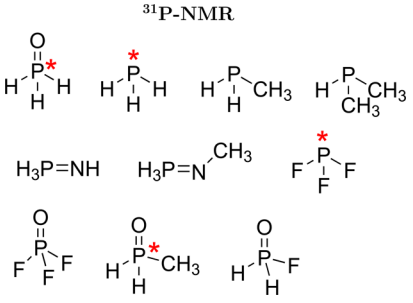


Figure 5. Molecular benchmark set for phosphorus NMR shifts. The phosphorus nuclei of the fitting set are labeled with a red star.

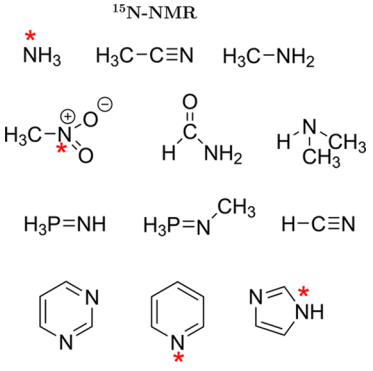


Figure 4. Molecular benchmark set for nitrogen NMR shifts. The nitrogen nuclei of the fitting set are labeled with a red star.

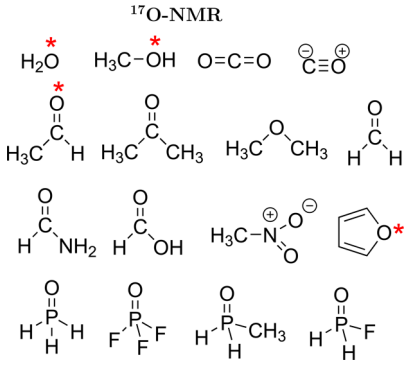


Figure 6. Molecular benchmark set for oxygen NMR shifts. The oxygen nuclei of the fitting set are labeled with a red star.

dx.doi.org/10.1021/ct5007295 | J. Chem. Theory Comput. 2015, 11, 37−44

Table 1. Comparison of SD, MSD, MAD, and MAXD for GIAO-SCS-MP2 vs Nonscaled MP2 (the Difference Is Denoted by Δ) Using the tz2p Basis[a]

	Fitting Set			Test Set			Total Set			Factors		
No. of ^{13}C shieldings	(20)			(20)			(40)					
	GIAO-SCS-MP2	MP2	Δ	GIAO-SCS-MP2	MP2	Δ	GIAO-SCS-MP2	MP2	Δ	c_{OS}	c_{SS}	MSD_{opt}
SD	2.2	3.1	−0.9	1.4	2.0	−0.6	1.9	2.6	−0.7	0.856	0.632	−1.457
MSD	0.6	−0.1	0.7	−0.4	−0.8	0.4	0.1	−0.5	0.6			
MAD	1.5	2.2	−0.7	1.0	1.6	−0.6	1.3	1.9	−0.6			
MaxD	6.5	5.8	0.7	4.5	4.8	−0.3	6.5	5.8	0.7			
No. of ^{13}C shieldings	(10)			(30)			(40)					
	GIAO-SCS-MP2	MP2	Δ	GIAO-SCS-MP2	MP2	Δ	GIAO-SCS-MP2	MP2	Δ	c_{OS}	c_{SS}	MSD_{opt}
SD	1.8	3.0	−1.2	2.1	2.5	−0.4	2.0	2.6	−0.6	0.724	0.927	−1.271
MSD	1.0	−0.6	1.6	1.1	−0.4	1.5	1.1	−0.5	1.6			
MAD	1.4	2.3	−0.9	1.4	1.8	−0.4	1.4	1.9	−0.5			
MaxD	4.1	4.9	−0.8	8.6	5.8	2.8	8.6	5.8	2.8			
No. of ^{13}C shieldings	(5)			(35)			(40)					
	GIAO-SCS-MP2	MP2	Δ	GIAO-SCS-MP2	MP2	Δ	GIAO-SCS-MP2	MP2	Δ	c_{OS}	c_{SS}	MSD_{opt}
SD	0.8	2.3	−1.5	2.0	2.6	−0.6	1.9	2.6	−0.7	0.886	0.649	−2.353
MSD	−0.2	−0.1	−0.1	0.0	−0.5	0.5	0.0	−0.5	0.5			
MAD	0.6	1.7	−1.1	1.4	1.9	−0.5	1.3	1.9	−0.6			
MaxD	1.2	3.4	−2.2	6.1	5.8	0.3	6.1	5.8	0.3			

[a]Data are shown for different fitting sets, test sets, and the total benchmark set (CCSD(T)/cc-pVQZ is employed as a reference). While the total benchmark set consists of 40 ^{13}C shieldings, the different fitting sets vary in the number of ^{13}C shieldings and the corresponding test sets comprise the remaining shieldings. In addition, the values for c_{OS}, c_{SS}, and MSD_{opt} are listed for the different fitting sets.

We follow the error criteria defined in previous work,[32] such as the mean signed deviation (MSD), the mean absolute deviation (MAD), the standard deviation (SD), and the maximum absolute deviation (MaxD):

$$\mathrm{MSD} = \frac{1}{N}\sum_{A}^{N} \overbrace{\delta_A - \delta_A^{\mathrm{CCSD(T)/cc\text{-}pVQZ}}}^{\Delta\delta_A}$$

$$\mathrm{MAD} = \frac{1}{N}\sum_{A}^{N} |\Delta\delta_A|$$

$$\mathrm{SD} = \sqrt{\frac{1}{N-1}\sum_{A}^{N} (\mathrm{MSD} - \Delta\delta_A)^2}$$

Here, N is the number of nuclei and δ_A the relative shielding of a nucleus A with respect to the reference nucleus such as for example TMS: $\delta_A = \sigma_{TMS} - \sigma_A$.

The opposite spin and the same spin terms of the MP2 correlation part of all nuclei were calculated with our AO-MP2 NMR method,[20] which is implemented in a development version of the program package Q-Chem.[37] These contributions were converged to a maximum deviation of 0.1 ppm with respect to MO-based MP2 shifts.

Furthermore, we employ the following basis sets: 6-31G** (6d),[38,39] 6-311G**,[40] pcS-0, pcS-1, pcS-2,[41] def2-SVP,[42] tz2p, qz2p,[33,34] cc-pVDZ, cc-pVTZ, and cc-pVQZ.[35] In all calculations, all electrons were correlated.

4. CARBON NMR SHIFTS

4.1. Accuracy Improvements of Carbon SCS-MP2 NMR Shifts for a Triple-ζ Basis Set. In the following, we discuss the influences of scaling the opposite and same spin part of carbon MP2 NMR shieldings using the triple-ζ basis set tz2p.[33,34] We optimized the scaling parameters for a fitting set with respect to CCSD(T)/cc-pVQZ results. The accuracy was studied for a test set with CCSD(T)/cc-pVQZ reference data. The CCSD(T)/cc-pVQZ data is the currently most reliable theoretical NMR data for the present benchmark set[32] and is independent of influences from structure and vibrational effects in comparison to experimentally determined reference data.

The total benchmark set consists of 40 different carbon shieldings[32] that we split into two sets each comprising 20 shieldings: one fitting set used for optimizing the scaling factors of GIAO-SCS-MP2/tz2p for approximating the CCSD(T)/cc-pVQZ values, while the other half (the test set) is used for benchmarking the obtained scaling factors. The carbon shieldings in the fitting set are labeled in Figure 1 and cover a broad range of the NMR scale for carbon nuclei (0−200 ppm) as shown in the Supporting Information.

The data shown in Table 1 for the fitting set, the test set, and the total set indicate, that the SD and the MAD are significantly reduced in comparison to nonscaled MP2. For example, the SD of the total set reduces from 2.6 to 1.9 ppm and the MAD from 1.9 to 1.3 ppm. The MSD and MaxD values are only slightly worse for some cases as, e.g., the MaxD of the total set increases from 5.8 to 6.5 ppm.

We have also investigated the influence of the number of ^{13}C shieldings in the fitting set. For this purpose, instead of 20, only 10 or 5 carbon shieldings were included in the fitting set. The remaining carbon shieldings are then employed as the test set (30 or 35 carbon shieldings, respectively). As shown in Table 1, the obtained scaling factors are quite similar. Also the SD and MAD values in comparison to nonscaled MP2 are reduced.

4.2. Scaling Factors for Various Basis Sets. While the triple-ζ basis sets yield quite accurate data, the size of the basis

dx.doi.org/10.1021/ct5007295 | J. Chem. Theory Comput. 2015, 11, 37−44

set plays a central role in the computation time and limits the size of the treatable systems. Therefore, the gain of accuracy by using our GIAO-SCS-MP2 and GIAO-SOS-MP2 methods for small and medium-sized basis sets is shown in Figure 2. Here, we employ the 20 carbon shieldings of the fitting set in optimizing the scaling factors to the CCSD(T)/cc-pVQZ data (the parameters c_{OS}, c_{SS}, and MSD_{opt} obtained by the fitting procedure are shown in Table 2).

Table 2. Parameters c_{OS}, c_{SS}, and MSD_{opt} for Carbon NMR Shifts As Obtained by Employing the 20-Carbon Fitting Set (Optimized Towards CCSD(T)/cc-pVQZ Results)

		GIAO-SCS-MP2			GIAO-SOS-MP2	
basis set	$\bar{N}_{bas}$ [a]	c_{OS}	c_{SS}	MSD_{opt}	c_{OS}	MSD_{opt}
6-31G**	9.8	−0.029	2.193	−12.765	0.452	−13.734
6-311G**	11.9	0.647	1.236	−3.175	0.927	−4.471
pcS-0	5.4	1.219	0.689	−8.938	1.400	−9.709
pcS-1	10.8	0.885	0.886	−1.120	1.093	−2.219
def2-SVP	9.4	0.207	1.700	−10.935	0.603	−11.965
tz2p	16.3	0.856	0.632	−1.457	0.998	−2.177
cc-pVDZ	9.4	0.196	1.812	−14.020	0.614	−15.056

[a]Average number of basis functions per atom (determined for the molecular benchmark set).

The SDs of GIAO-SCS-MP2 and GIAO-SOS-MP2 are for almost all basis sets smaller than those of the nonscaled MP2 method. Exceptions are GIAO-SOS-MP2/pcS-0 for which the SD stays the same and GIAO-SOS-MP2/pcS-1 for which the SD is slightly worse. The scaling of the MP2 values shows for the medium-sized basis sets 6-31G**, def2-SVP, and cc-pVDZ huge improvements in comparison to nonscaled MP2. The SD of GIAO-SCS-MP2 using the def2-SVP basis for example reduces from 6.3 down to 2.3 ppm and is much smaller than the corresponding 5.7 ppm of CCSD(T)/def2-SVP. While the poor performance of MP2 or CCSD(T) with the small def2-SVP basis is not a surprise, the GIAO-SCS-MP2 results illustrate that the scaling parameters can somewhat (pragmatically) compensate at least some basis set deficiencies. Finally, the SDs of GIAO-SCS-MP2 for larger basis sets such as pcS-1 and tz2p are close to the SDs of the corresponding results at the CCSD(T) level, showing the usefulness of the pragmatic approach (MSD, MAD, MaxD, and the carbon shifts for all basis sets are shown in the Supporting Information).

For the different basis sets, we also investigated the approach of not including the $\Delta\sigma_{HF}$ of eq 4 in the fitting procedure. Here, we employ the same 20 carbon shieldings in the fitting set as used above (the parameters c_{OS}, c_{SS}, and MSD_{opt} are given in Table 4). The results obtained are compared with nonscaled MP2 and those of the variant with $\Delta\sigma_{HF}$ included in the fitting procedure (SDs are shown in Table 3 and the other error criteria in the Supporting Information). Also for the correlation-only scaled approach we observe, for almost all basis sets, an improvement of the SDs compared with nonscaled MP2 values. Exceptions are, again, the pcS-0 and pcS-1 basis sets. However, by neglecting the $\Delta\sigma_{HF}$ term in the fitting procedure, we obtain SDs, across all basis sets, which are higher than in the corresponding fit procedure which includes the HF-difference term. We note that the difference in the quality of the fit is heavily dependent on the size of the chosen basis set, as shown in Table 3. This observation can be traced back to the large differences in the HF-shielding of the CCSD(T)/cc-pVQZ reference calculations. Overall, our data

Table 3. Standard Deviation (SD) of ^{13}C Chemical Shifts for the GIAO-SCS-MP2 and GIAO-SOS-MP2 Methods in Comparison to Nonscaled MP2 for Various Basis Sets[a]

basis set	$\bar{N}_{bas}$ [b]	MP2	GIAO-SCS-MP2 (without $\Delta\sigma_{HF}$)	GIAO-SCS-MP2 (with $\Delta\sigma_{HF}$)	GIAO-SOS-MP2 (without $\Delta\sigma_{HF}$)	GIAO-SOS-MP2 (with $\Delta\sigma_{HF}$)
6-31G**	9.8	7.5	5.6	2.4	6.4	4.4
6-311G**	11.9	3.9	3.1	2.6	3.7	3.7
pcS-0	5.4	6.7	12.4	6.5	12.5	6.7
pcS-1	10.8	2.5	2.5	2.2	2.9	2.9
def2-SVP	9.4	6.3	3.9	2.3	4.5	3.1
tz2p	16.3	2.6	2.0	1.9	2.5	2.5
cc-pVDZ	9.4	6.7	4.2	2.8	4.7	3.4

[a]For the scaled MP2 methods two different variants are shown, where the HF deviation $\Delta\sigma_{HF}$ of eq 4 is either omitted (without $\Delta\sigma_{HF}$) or included (with $\Delta\sigma_{HF}$) in the fitting procedure. As a reference, the CCSD(T)/cc-pVQZ data is employed. [b]Average number of basis functions per atom (determined for the molecular benchmark set).

Table 4. Parameters c_{OS}, c_{SS}, and MSD_{opt} for Carbon NMR Shifts as Obtained by Employing the 20-Carbon Fitting Set (Optimized Towards CCSD(T)/cc-pVQZ Results)[a]

		GIAO-SCS-MP2(without $\Delta\sigma_{HF}$)			GIAO-SOS-MP2(without $\Delta\sigma_{HF}$)	
basis set	$\bar{N}_{bas}$ [b]	c_{OS}	c_{SS}	MSD_{opt}	c_{OS}	MSD_{opt}
6-31G**	9.8	0.709	1.169	−1.367	0.966	−1.883
6-311G**	11.9	0.828	0.546	−0.167	0.952	−0.740
pcS-0	5.4	0.469	0.410	−1.688	0.577	−2.146
pcS-1	10.8	0.883	0.467	0.110	0.992	−0.468
def2-SVP	9.4	0.683	1.034	−0.892	0.924	−1.518
tz2p	16.3	0.949	0.395	0.143	1.038	−0.307
cc-pVDZ	9.4	0.656	1.063	−1.604	0.900	−2.211

[a]Here, the HF deviation $\Delta\sigma_{HF}$ of eq 4 is omitted (without $\Delta\sigma_{HF}$) in the fitting procedure. [b]Average number of basis functions per atom (determined for the molecular benchmark set).

strongly motivate the use of scaling factors, which are obtained from the fitting procedure, where the $\Delta\sigma_{HF}$ term is included.

4.3. Scaling Influences for Applying Different Basis Sets in the HF and in the Correlation Terms of Carbon NMR Shifts. In the previous sections, we employed the standard approach of using the same basis sets for the HF term σ_{HF} and the correlation parts of MP2 and CCSD(T) NMR shifts. Here, we propose a pragmatic alternative and less common approach to employ different basis sets for the two terms. Since HF-based NMR calculations are much faster than conventional correlation methods, larger basis sets are affordable for the HF term. As an example, an HF, MP2, or CCSD(T) calculation (with the program package CFOUR[36] on a single core of an Intel Xeon E5-2620 workstation using 128 GB RAM) for the molecule pyrimidine (from our benchmark set) and the pcS-1 basis set requires roughly 1, 3, or 116 min, respectively. With the larger basis set pcS-2, the calculation takes 18, 57, or 2208 min, respectively. These time differences become even larger when increasing the size of the molecule because of the different scaling behavior of these methods. As a consequence, it is typically affordable to perform an additional calculation of HF NMR shieldings with a larger basis set in contrast to the costly account of correlation effects with a larger basis set.

dx.doi.org/10.1021/ct5007295 | J. Chem. Theory Comput. 2015, 11, 37−44

Table 5. Standard Deviation (SD) of ^{13}C Shifts Based on CCSD(T), GIAO-SCS-MP2, GIAO-SOS-MP2, and Nonscaled MP2 with Respect to the CCSD(T)/cc-pVQZ Reference Data in Employing Different Basis Sets for the HF Part and the Correlation Part (corr. part)[a]

basis set					
corr. part ($\bar{N}_{bas}$[b])	HF part ($\bar{N}_{bas}$[b])	CCSD(T)	GIAO-SCS-MP2	GIAO-SOS-MP2	MP2
6-31G** (9.8)	6-31G** (9.8)	7.2	2.4	4.4	7.5
6-31G** (9.8)	qz2p (18.7)	2.2	1.7	2.6	2.6
6-31G** (9.8)	pcS-2 (23.2)	1.4	1.7	2.3	2.0
6-31G** (9.8)	cc-pVQZ (41.8)	2.2	1.5	2.5	2.7
pcS-1 (10.8)	pcS-1 (10.8)	2.1	2.2	2.9	2.5
pcS-1 (10.8)	qz2p (18.7)	1.2	1.7	2.1	2.4
pcS-1 (10.8)	pcS-2 (23.2)	0.9	2.1	2.1	2.5
pcS-1 (10.8)	cc-pVQZ (41.8)	1.4	1.6	2.0	2.6
tz2p (16.3)	tz2p (16.3)	1.7	1.9	2.5	2.6
tz2p (16.3)	qz2p (18.7)	0.8	1.6	1.9	2.1
tz2p (16.3)	pcS-2 (23.2)	1.1	1.9	2.0	2.5
tz2p (16.3)	cc-pVQZ (41.8)	0.8	1.5	1.8	2.2

[a]The new parameters for the GIAO-SCS-MP2 and GIAO-SOS-MP2 methods were optimized towards CCSD(T)/cc-pVQZ results of the 20-carbon fitting set. [b]Average number of basis functions per atom (determined for the molecular benchmark set).

Therefore, we tested the pragmatic approach of calculating the HF part with a larger basis and the correlation part with a smaller basis set. The corresponding results shown in Table 5 illustrate the gain of accuracy for carbon NMR shifts based on the CCSD(T), GIAO-SCS-MP2, GIAO-SOS-MP2, or nonscaled MP2 method vs the standard approach of employing the same basis for both parts. As before, we optimized the corresponding scaling factors for our GIAO-SCS-MP2 and GIAO-SOS-MP2 methods (with a fitting set consisting of 20 carbon NMR shieldings and with CCSD(T)/cc-pVQZ as a reference). The corresponding scaling factors are shown in Table 6.

Table 6. Parameters c_{OS}, c_{SS}, and MSD_{opt} for Using Different Basis Sets for the HF Part and the Correlation Part (corr. part) of MP2 Based Carbon Shifts Optimized for the 20-Carbon Fitting Set (towards CCSD(T)/cc-pVQZ Data)

basis set		GIAO-SCS-MP2			GIAO-SOS-MP2	
corr. part	HF part	c_{OS}	c_{SS}	MSD_{opt}	c_{OS}	MSD_{opt}
6-31G**	6-31G**	−0.029	2.193	−12.765	0.452	−13.734
6-31G**	qz2p	0.748	1.222	−1.041	1.016	−1.581
6-31G**	pcS-2	0.908	0.762	−0.228	1.075	−0.564
6-31G**	cc-pVQZ	0.709	1.169	−1.367	0.966	−1.883
pcS-1	pcS-1	0.885	0.886	−1.120	1.093	−2.219
pcS-1	qz2p	0.929	0.501	0.508	1.047	−0.114
pcS-1	pcS-2	1.081	0.058	1.181	1.094	1.109
pcS-1	cc-pVQZ	0.883	0.467	0.110	0.992	−0.468
tz2p	tz2p	0.856	0.632	−1.457	0.998	−2.177
tz2p	qz2p	1.001	0.418	0.532	1.095	0.056
tz2p	pcS-2	1.138	0.021	1.326	1.143	1.302
tz2p	cc-pVQZ	0.949	0.395	0.143	1.038	−0.307

By replacing for GIAO-SCS-MP2 and GIAO-SOS-MP2 the HF contribution of the MP2 part $\sigma_{HF(MP2)}$ in eq 4 with a different basis set, the described HF deviation $\Delta\sigma^{HF}$ in eq 4 changes. If one employs the cc-pVQZ basis for the HF contribution of the MP2 part, the deviation is zero ($\Delta\sigma^{HF} = 0$), since the cc-pVQZ basis set was used for the reference data (CCSD(T)/cc-pVQZ). This seems to be beneficial for the scaling procedure, as for most of the basis set combinations the lowest SD values can be obtained with this basis for the HF part (see Table 5).

Table 7. Parameters c_{OS}, c_{SS}, and MSD_{opt} for Fluorine, Nitrogen, Oxygen, and Phosphorus NMR Shifts As Obtained by Employing Four Shieldings in the Fitting Set (Optimized Towards CCSD(T)/cc-pVQZ Results)

^{19}F-NMR	GIAO-SCS-MP2			GIAO-SOS-MP2	
basis set	c_{OS}	c_{SS}	MSD_{opt}	c_{OS}	MSD_{opt}
6-31G**	0.026	1.896	−16.567	0.522	−28.920
pcS-1	0.817	1.193	−4.949	1.173	−13.256
cc-pVDZ	−0.486	1.882	−17.281	−0.170	−29.446
^{15}N NMR	GIAO-SCS-MP2			GIAO-SOS-MP2	
basis set	c_{OS}	c_{SS}	MSD_{opt}	c_{OS}	MSD_{opt}
6-31G**	−0.023	2.069	−8.413	0.686	−16.749
pcS-1	0.967	−0.261	−3.221	0.878	−2.015
cc-pVDZ	0.449	0.951	−12.712	0.780	−16.913
^{17}O NMR	GIAO-SCS-MP2			GIAO-SOS-MP2	
basis set	c_{OS}	c_{SS}	MSD_{opt}	c_{OS}	MSD_{opt}
6-31G**	0.009	0.946	−5.979	0.210	−8.926
pcS-1	1.078	0.188	−8.742	1.107	−9.216
cc-pVDZ	0.436	0.554	−12.386	0.543	−14.166
^{31}P NMR	GIAO-SCS-MP2			GIAO-SOS-MP2	
basis set	c_{OS}	c_{SS}	MSD_{opt}	c_{OS}	MSD_{opt}
6-31G**	−1.324	3.286	−19.962	2.412	−29.188
pcS-1	−2.492	2.533	15.884	0.591	9.891
cc-pVDZ	−3.947	4.807	−32.092	1.673	−52.584

While the cc-pVQZ basis is rather large, lower SDs can also be obtained with smaller basis sets for the HF contribution of the MP2 part. The carbon NMR shieldings based on GIAO-SCS-MP2 show a small SD of 1.7 ppm when using the qz2p basis for the HF part and 6-31G** for the correlation part (corr. part) as compared to 2.4 ppm in the standard GIAO-SCS-MP2/6-31G** approach—see Table 5. Also the combination of corr. part:tz2p and HF part:qz2p results in a low SD value of 1.6 ppm (vs 1.9 ppm). For the GIAO-SOS-MP2 approach, the combination of corr. part:pcS-1 and HF part:qz2p is a very cost-efficient way to calculate NMR shieldings and reduces the SD to 2.1 ppm (vs 2.9 ppm). An even lower SD of 1.9 ppm (vs 2.5 ppm) can be obtained by using the combination corr. part:tz2p and HF part:qz2p.

For nonscaled MP2, using the pcS-1 basis set for both parts remains a good compromise between accuracy and cost (SD of 2.5 ppm). Although a slightly lower SD can be obtained by using the combinations corr. part:6-31G**/HF part:pcS-2 (SD of 2.0 ppm) or corr. part:tz2p/HF part:qz2p (SD of 2.1 ppm), SD values under 2 ppm however can not be achieved by using a larger basis set for the HF-part.

Using different basis sets for the HF and correlation contributions is also beneficial for CCSD(T) NMR shieldings. For example, the SD reduces from 2.1 to 0.9 ppm for the combination corr. part:pcS-1/HF part:pcS-2 and from 1.7 to 0.8 ppm for corr. part:tz2p/HF part:qz2p. Furthermore, NMR shieldings based on CCSD(T) with a SD value under 1.5 ppm can be obtained even faster, if the combinations corr. part:6-31G**/HF part:pcS-2 (vs 7.2 ppm) or corr. part:pcS-1/HF part:qz2p (vs 2.1 ppm) are applied. Further data for this approach such as the MAD and MSD of the different methods are given in the Supporting Information.

dx.doi.org/10.1021/ct5007295 | J. Chem. Theory Comput. 2015, 11, 37−44

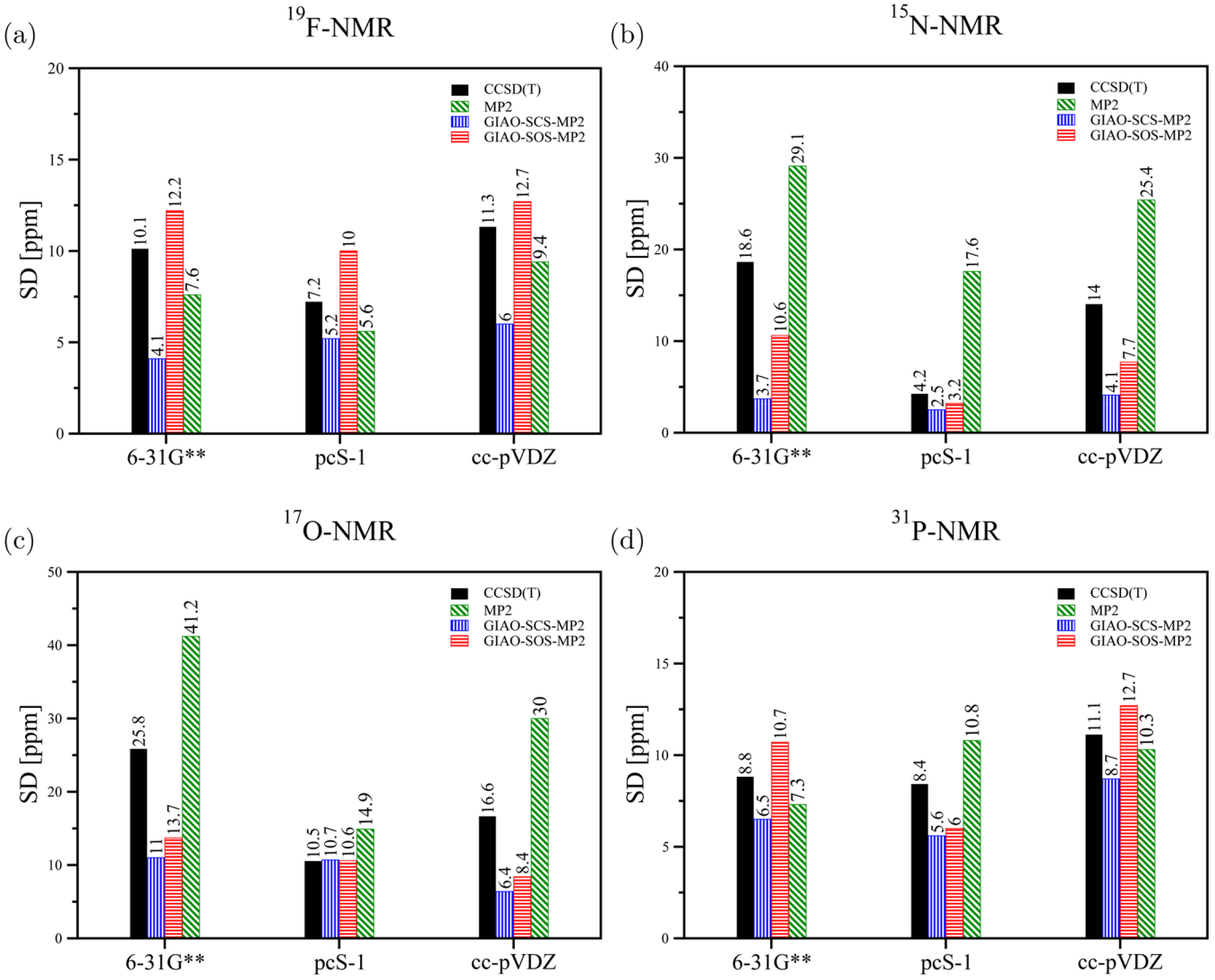


Figure 7. Standard deviation (SD) of fluorine, nitrogen, oxygen, and phosphorus chemical shifts for the GIAO-SCS-MP2 and GIAO-SOS-MP2 methods in comparison to nonscaled MP2 and CCSD(T) results are listed for different basis sets. As a reference the CCSD(T)/cc-pVQZ data is employed. Data for the total benchmark set are shown, while the fitting coefficients for GIAO-SCS-MP2 and GIAO-SOS-MP2 have been obtained with four shieldings in the fitting set.

5. PHOSPHORUS, NITROGEN, OXYGEN, AND FLUORINE NMR SHIFTS

The gain of accuracy by using our GIAO-SCS-MP2 and GIAO-SOS-MP2 methods were also investigated for phosphorus, nitrogen, oxygen, and fluorine NMR shifts. The benchmark sets of these nuclei consists of 10 phosphorus (relative to PH_3), 13 nitrogen (relative to CH_3NO_2), 18 oxygen (relative to H_2O) and 11 fluorine (relative to $CFCl_3$) NMR shifts. To optimize the scaling parameters with respect to CCSD(T)/cc-pVQZ results, we employ four shieldings in the fitting set in each benchmark set (the parameters c_{OS}, c_{SS}, and MSD_{opt} are shown in Table 7).

The SDs with respect to CCSD(T)/cc-pVQZ are presented for the different nuclei in Figure 7. While the SDs of GIAO-SOS-MP2 for fluorine and phosphorus nuclei are mostly larger than those of nonscaled MP2, the GIAO-SOS-MP2 shifts show huge accuracy improvements for nitrogen and oxygen. The SDs of GIAO-SCS-MP2 are for all shown basis sets smaller than those of the nonscaled MP2 method. Furthermore, for the same basis set (smaller than cc-pVQZ basis employed for computing the CCSD(T) reference values) our GIAO-SCS-MP2 method provides NMR shift SDs which are lower than those at CCSD(T) level using the same basis set if compared to the CCSD(T)/cc-pVQZ reference value. This indicates the advantages of the empirical fitting procedure with the drawback of lost ab initio character. The SD of oxygen GIAO-SCS-MP2 shifts using the cc-pVDZ basis for example reduces from 30 to 6.4 ppm and is much smaller than the corresponding 16.6 ppm of CCSD(T)/cc-pVDZ. Furthermore, the SD of nitrogen GIAO-SCS-MP2 shifts with the basis set pcS-1 reduces from 17.6 to 2.5 ppm and the SD of phosphorus GIAO-SCS-MP2 shifts using the basis set pcS-1 from 10.8 to 5.6 ppm.

Further data such as the MAD, MSD, MaxD, and the individual shifts of the different nuclei for all basis sets are given in the Supporting Information.

6. CONCLUSION

We have introduced GIAO-SCS-MP2 and GIAO-SOS-MP2 approaches and explored their usefulness for calculating NMR chemical shifts. The results improve significantly as compared

dx.doi.org/10.1021/ct5007295 | J. Chem. Theory Comput. 2015, 11, 37−44

to nonscaled MP2 and results are closer to the CCSD(T)/cc-pVQZ data used as a reference. While such scaled MP2 approaches are always empirical, they seem useful from a pragmatic point of view. Furthermore, we studied the effect of employing different basis sets for the Hartree−Fock part of MP2 and CC NMR chemical shielding calculations and show the usefulness as another pragmatic approach. The latter mixed basis approach is particularly useful if combined with scaled GIAO-MP2 approaches.

■ ASSOCIATED CONTENT

Supporting Information

Explicit values of the shifts determined with our GIAO-SCS-MP2 and GIAO-SOS-MP2 methods and all error criteria such as the SD, MSD, MAD, and MaxD. Furthermore, the CCSD(T)/cc-pVTZ optimized structures of the benchmark sets as well as the additional details of the approach of using different basis sets in the HF and in the correlation terms of carbon NMR shifts. This material is available free of charge via the Internet at http://pubs.acs.org/.

■ AUTHOR INFORMATION

Corresponding Author

*Electronic mail: christian.ochsenfeld@uni-muenchen.de.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors thank Dr. D. Flaig and Dr. S. A. Maurer (University of Munich, LMU) for useful discussions. C.O. acknowledges financial support by the "Deutsche Forschungsgemeinschaft" (DFG) funding proposal Oc35/4-1 and the DFG cluster of excellence EXC 114 "Center for Integrated Protein Science Munich" (CIPSM).

■ REFERENCES

(1) Ditchfield, R. *Mol. Phys.* **1974**, *27*, 789−807.
(2) Wolinski, K.; Hinton, J. F.; Pulay, P. *J. Am. Chem. Soc.* **1990**, *112*, 8251−8260.
(3) Häser, M.; Ahlrichs, R.; Baron, H.; Weis, P.; Horn, H. *Theor. Chim. Acta* **1992**, *83*, 455−470.
(4) Schreckenbach, G.; Ziegler, T. *J. Phys. Chem.* **1995**, *99*, 606−611.
(5) Rauhut, G.; Puyear, S.; Wolinski, K.; Pulay, P. *J. Phys. Chem.* **1996**, *100*, 6310−6316.
(6) Gauss, J. *Chem. Phys. Lett.* **1992**, *191*, 614−620.
(7) Gauss, J. *J. Chem. Phys.* **1993**, *99*, 3629−3643.
(8) Kollwitz, M.; Gauss, J. *Chem. Phys. Lett.* **1996**, *260*, 639−646.
(9) Kollwitz, M.; Häser, M.; Gauss, J. *J. Chem. Phys.* **1998**, *108*, 8295−8301.
(10) Gauss, J.; Stanton, J. F. *J. Chem. Phys.* **1995**, *103*, 3561−3577.
(11) Gauss, J.; Stanton, J. F. *J. Chem. Phys.* **1996**, *104*, 2574−2583.
(12) Kállay, M.; Gauss, J. *J. Chem. Phys.* **2004**, *120*, 6841−6848.
(13) Sumowski, C. V.; Hanni, M.; Schweizer, S.; Ochsenfeld, C. *J. Chem. Theory Comput.* **2014**, *10*, 122−133.
(14) Christensen, A. S.; Linnet, T. E.; Borg, M.; Boomsma, W.; Lindorff-Larsen, K.; Hamelryck, T.; Jensen, J. H. *PLoS One* **2013**, *8*, e84123.
(15) Ochsenfeld, C.; Kussmann, J.; Koziol, F. *Angew. Int. Ed.* **2004**, *43*, 4485−4489.
(16) Kussmann, J.; Ochsenfeld, C. *J. Chem. Phys.* **2007**, *127*, 054103.
(17) Beer, M.; Kussmann, J.; Ochsenfeld, C. *J. Chem. Phys.* **2011**, *134*, 074102.
(18) Gauss, J.; Werner, H.-J. *Phys. Chem. Chem. Phys.* **2000**, *2*, 2083.
(19) Loibl, S.; Schütz, M. *J. Chem. Phys.* **2012**, *137*, 084107.
(20) Maurer, M.; Ochsenfeld, C. *J. Chem. Phys.* **2013**, *138*, 174104.
(21) Maurer, S. A.; Lambrecht, D. S.; Kussmann, J.; Ochsenfeld, C. *J. Chem. Phys.* **2013**, *138*, 014101.
(22) Grimme, S. *J. Chem. Phys.* **2003**, *118*, 9095−9102.
(23) Grimme, S.; Goerigk, L.; Fink, R. F. *WIREs Comput. Mol. Sci.* **2012**, *2*, 886−906.
(24) Jung, Y.; Lochan, R. C.; Dutoi, A. D.; Head-Gordon, M. *J. Chem. Phys.* **2004**, *121*, 9793−9802.
(25) Fink, R. F. *J. Chem. Phys.* **2010**, *133*, 174113.
(26) Grimme, S.; Izgorodina, E. I. *Chem. Phys.* **2004**, *305*, 223−230.
(27) Rhee, Y. M.; Head-Gordon, M. *J. Phys. Chem. A* **2007**, *111*, 5314−5326.
(28) Hellweg, A.; Grün, S. A.; Hättig, C. *Phys. Chem. Chem. Phys.* **2008**, *10*, 4119−4127.
(29) London, F. *J. Phys. Radium* **1937**, *8*, 397.
(30) Hameka, H. F. *Mol. Phys.* **1958**, *1*, 203.
(31) Gauss, J. In *Modern Methods and Algorithms of Quantum Chemistry*, ed. Grotendorst, J., Vol. 3; John von Neumann Institute for Computing: Jülich, 2000; pp 541−592.
(32) Flaig, D.; Maurer, M.; Hanni, M.; Braunger, K.; Kick, L.; Thubauville, M.; Ochsenfeld, C. *J. Chem. Theory Comput.* **2014**, *10*, 572−578.
(33) Auer, A. A.; Gauss, J.; Stanton, J. F. *J. Chem. Phys.* **2003**, *118*, 10407−10417.
(34) Schäfer, A.; Horn, H.; Ahlrichs, R. *J. Chem. Phys.* **1992**, *97*, 2571−2577.
(35) Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007−1023.
(36) CFOUR, a quantum chemical program package written by Stanton, J. F.; Gauss, J.; Harding, M. E.; Szalay, P. G. with contributions from Auer, A. A.; Bartlett, R. J.; Benedikt, U.; Berger, C.; Bernholdt, D. E.; Bomble, Y. J.; Christiansen, O.; Heckert, M.; Heun, O.; Huber, C.; Jagau, T.-C.; Jonsson, D.; Jusélius, J.; , Klein, K.; Lauderdale, W. J.; Matthews, D. A.; Metzroth, T.; O'Neill, D. P.; Price, D. R.; Prochnow, E.; Ruud, K.; Schiffmann, F.; Stopkowicz, S.; Tajti, A.; Vázquez, J.; Wang, F.; Watts, J. D. and the integral packages MOLECULE (Almlöf, J.; Taylor, P. R.), PROPS (Taylor, P. R.), ABACUS (Helgaker, T.; Jensen, H. J. Aa.; Jørgensen, P.; Olsen, J.), and ECP routines by Mitin, A. V. and van Wüllen, C. For the current version, see http://www.cfour.de (accessed November 10, 2014).
(37) Development version of Q-Chem, www.q-chem.com (accessed November 10, 2014).
(38) Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257−2261.
(39) Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 213−222.
(40) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650−654.
(41) Jensen, F. *J. Chem. Theory Comput.* **2008**, *4*, 719−727.
(42) Weigend, F.; Ahlrichs, R. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297−3305.