Dissertation zur Erlangung des Doktorgrades
der Fakultät für Chemie und Pharmazie
der Ludwig-Maximilians-Universität München

Development of efficient and low-scaling methods to compute molecular properties at MP2 and double-hybrid DFT levels

Sigurd Vogler

aus

München

2018

Erklärung

Diese Dissertation wurde im Sinne von §7 der Promotionsordnung vom 28. November 2011 von Herrn Prof. Dr. Christian Ochsenfeld betreut.

Eidesstattliche Versicherung

Diese Dissertation wurde eigenständig und ohne unerlaubte Hilfe erarbeitet.

Sigurd Vogler
München, 30.10.2018

Dissertation eingereicht am	10.08.2018
1. Gutachter	Prof. Dr. Christian Ochsenfeld
2. Gutachter	Prof. Dr. Regina de Vivie-Riedle
Mündliche Prüfung am	29.10.2018

Danksagung

Bei meinem Doktorvater, Herrn Prof. Dr. Christian Ochsenfeld, möchte ich mich für die Themenstellung, für die Betreuung und für die Möglichkeit, mich während meiner Dissertation in seinem Arbeitskreis wissenschaftlich weiter zu bilden, bedanken. Mein Dank gilt auch Frau Prof. Dr. Regina de Vivie-Riedle für die freundliche Übernahme des Zweitgutachtens und den weiteren Mitgliedern der Promotionskommission. Die Finanzierung und ideelle Förderung der Studienstiftung des deutschen Volkes unterstützte mich maßgeblich bei meiner Promotion.

Einen wichtigen Anteil am Abschluss meiner Promotion haben meine ehemaligen und aktuellen Kollegen des Arbeitskreises Ochsenfeld. Besonders bedanken möchte ich mich bei Dr. Marina Maurer, die mich anfangs betreute und von der ich viel über Quantenchemie, Programmieren und vielem mehr lernen durfte. Die Zusammenarbeit mit Michael Glasbrenner, Martin Ludwig, Laurens Peters, Johannes Dietschreit und Gökcen Savasci waren immer lehrreich, und ich bin dankbar mit euch gemeinsam gearbeitet zu haben. Lebhafte Diskussionen mit meinen (ehemaligen) Bürokollegen Matthias Beuerle und Dr. Henry Schurkus halfen mit durch Zeiten, in denen die Forschung stockte, ebenso wie das gemeinsame Laufen mit dem AK-internen Laufteam "Runtime Error" und die Rennradtouren mit dem "Rennstall".

Ohne die Unterstützung meiner Familie, insbesondere meiner Eltern und meiner Geschwister, und auch meiner Freunde wäre diese Promotion nicht möglich gewesen. Besonders bedanken möchte ich bei den Mitgliedern meines nicht-unversitären Laufteams, Joel Sedlacko und Johanna Mandelartz, für fortwährenden Support, fürs Korrekturlesen und für alles andere, das hier noch nicht erwähnt wurde.

Summary

This thesis introduces new methods to compute molecular properties at the level of second-order Møller-Plesset perturbation theory (MP2) and double-hybrid density functional theory, building on a reformulation in atomic orbitals and exploiting the rank deficiency of the (pseudo-)density matrices, thus reducing the scaling behavior with respect to the size of the basis set. By furthermore employing the resolution-of-the-identity approximation, low-scaling and efficient MP2 energy gradients are presented, where significant two-electron integrals are screened using a distance-including integral estimation technique. With this, the forces and the hyperfine coupling constants of systems larger than previously computable at the MP2-level are obtained.

In the second part of this thesis, the locality of the spin density in many molecular systems is exploited in the computation of the hyperfine coupling constants, leading to further speed-ups and allowing for a thorough investigation of the effect of the protein environment on the hyperfine coupling within the core region of a pyruvate formate lyase. With this efficient method, studying the effect of nuclear motion on the accuracy of the computed hyperfine coupling constants is possible. The study presented in this thesis demonstrates that both electron correlation and vibrational motion are crucial for an accurate theoretical description.

When calculating magnetic properties, the dependence on the choice of gauge origins needs to be considered. This effect is studied systematically, and in detail, in a fourth project of this thesis for the computation of electronic g-tensors, for which it was previously assumed that the computation is largely independent of the choice of the gauge-origin. The study clearly contradicts this assumption and motivates the use of gauge including atomic orbitals in future work on electronic g-tensors.

In a last part, this work transfers the algorithmic developments on the computation of analytic gradients to the computation of nuclear magnetic resonance (NMR) shieldings at the MP2-level. Though a sublinear scaling ansatz to compute the NMR shielding tensor per nucleus is available, the lack of an efficient implementation and the large dependency on the size of the basis sets prohibits the accurate computation of the shielding tensor of medium- to large-sized molecules. Furthermore, while this ansatz in theory scales linearly when all nuclei in a system are computed, it is inefficient due to the dependence of the rate-determining steps on the nuclear magnetic moments. This thesis therefore presents a new all-nuclei ansatz and introduces the methodology for the efficient computation of the energy gradients developed in this thesis, highlighting significant computational savings.

Contents

List of Publications

This is a cumulative dissertation, comprising three articles in peer-reviewed journals (**I-III**) and a manuscript (**IV**), that is in preparation for publication. In the following, all articles and the manuscript are stated together with the author's contribution to each of them.

I
Sigurd Vogler, Martin Ludwig, Marina Maurer, Christian Ochsenfeld,
"Low-scaling first-order properties within second-order Møller-Plesset perturbation theory using Cholesky decomposed density matrices",
J. Chem. Phys., **147**, 024101 (2017)
Contribution by Sigurd Vogler: *All derivations, the implementation, all calculations, and writing the paper.*

II
Sigurd Vogler, Gökcen Savasci, Martin Ludwig, Christian Ochsenfeld,
"Selected-Nuclei Method for the Computation of Hyperfine Coupling Constants within Second-Order Møller-Plesset Perturbation Theory",
J. Chem. Theory Comput., **14**, 3014-3024 (2018)
Contribution by Sigurd Vogler: *All derivations, the implementation, all calculations, writing the paper except for the description of the PFL-system.*

III
Michael Glasbrenner, **Sigurd Vogler**, Christian Ochsenfeld,
"Gauge-origin dependence in electronic g-tensor calculations",
J. Chem. Phys., **148**, 214101 (2018)
Contribution by Sigurd Vogler: *Assisted with the derivations, the implementation, the design of the benchmark, and the writing of the paper.*

IV **Sigurd Vogler**, Johannes C. B. Dietschreit, Laurens D. M. Peters, Christian Ochsenfeld,
"Accurate Hyperfine Coupling Constants including Electron Correlation and Dynamic Effects", *in preparation*
Contribution by Sigurd Vogler: *Design of the study, all calculations, and the writing of most of the paper.*

1 Introduction

The extensive research on efficient and accurate quantum chemical methods has lead to their widespread application in aiding the understanding of reaction mechanisms and kinetics, as well as in determining molecular structures and their relationship with molecular properties. Two such examples are nuclear magnetic resonance (NMR) chemical shifts and hyperfine coupling constants from electron paramagnetic resonance (EPR) spectroscopy. While the theoretical foundation for an exact description of the electronic structure of molecular systems is well known, the computational complexity requires an introduction of different approximations. This necessity is due to their scaling behavior, which describes the exponent in the increase of required computational resources with respect to the size of the molecular system.

The basis of *ab initio* methods is the Hartree-Fock (HF) theory,[1–3] where the electron-electron interaction is described by a mean field-approach. However, the neglected electron correlation often proves vital for an accurate description. A variety of post-HF methods exist, which build either on a perturbation theory, describing electron correlation as a perturbation of the HF wavefunction, or on an expansion of the wavefunction in multiple determinants. The latter includes size-consistent coupled cluster (CC) approaches, such as the "gold standard" of quantum chemistry,[4] the CCSD(T) method, which includes singles and doubles excitations[5] and a perturbative description of triple excitations[6]. On the other hand, Møller-Plesset (MP) perturbation theory,[7] a special case of Rayleigh-Schrödinger perturbation theory,[8] provides a good compromise between accuracy and computational cost, while still being *ab initio* and wavefunction-based.

An alternative, popular route is density functional theory (DFT). Contrary to HF theory and post-HF methods, DFT is not an *a priori* wavefunction-based method. Based on the Hohenberg-Kohn theorem which states that the total energy is a unique functional of the electron density,[9] finding this functional is the quest of DFT. A common approach is the Kohn-Sham (KS) ansatz,[10] which reintroduces orbitals to accurately describe the kinetic energy. KS-DFT has proven to be reasonably accurate for a wide range of molecular

systems,[11] leading to its widespread use.[12] By adding exact exchange contributions in so-called hybrid functionals, this accuracy is typically increased.[13] Furthermore, including electron correlation in double-hybrid (DH-) DFT, e.g., by second-order perturbation theory (PT2) using Kohn-Sham orbitals, can lead to increased accuracy.[14] Recent work on improving these DH-DFT functionals, e.g., for open-shell systems, highlights their applicability.[15–17] The PT2 contribution to DH-DFT is from an algorithmic point of view identical to MP2, making developments for an efficient computation of MP2 energies, energy gradients, and molecular properties similarly applicable to DH-DFT.

Recent decades have seen significant advancements in the efficient computation of the MP2 equations. In order to reduce the scaling behavior, exploiting the local nature of the correlation contribution is crucial. An important approach is the use of localized orbitals based on the work by Pulay and Saebø[18–20] and for which local MP2 energies[21,22], energy gradients[23,24], and NMR shieldings[25,26] have been introduced. A second ansatz, which is especially efficient on parallelized architectures, is partitioning the orbital space[27–32] or fragmenting the molecule[33–37]. The third idea, which is the focus of this work, is based on a Laplace transformation[38–40], enabling a fully atomic orbital (AO-)based reformulation of the respective MP2 energy equation. The local nature of the AOs ultimately leads to linear scaling behavior.[41,42]

An accurate description of the electronic structure forms the basis of an *ab initio* computation of molecular properties, such as spectroscopic parameters. This constitutes an important link between experiment and theory and can help analyse and interpret experimental findings. As the Hellmann-Feynman theorem[43,44] only holds true for complete basis sets, molecular properties of a specific electronic state require analytic derivatives of the energy expression[45]. While numerical derivatives are often easy to implement, computational efficiency and accuracy mandate an analytic differentiation. The computation of these derivatives needs to be efficient with respect to computational resources and to the scaling behavior, allowing for a thorough analysis of the influence of the environment and solvation on the molecular property as well as of dynamic effects.

In this work, the computation of different molecular properties is investigated. Besides the nuclear gradients, which form the basis for geometry optimizations and *ab initio* Born-Oppenheimer molecular dynamics, these include different parameters from EPR and NMR spectroscopy. Both spectroscopic techniques are important analytical tools of chemistry due to their strong dependence on structural parameters and high sensitivity, allowing the detection of molecules, and their structure and dynamic effects. Their *ab initio* computa-

tion remains challenging and often electron correlation is crucial for accurate results[46,47].

Based on the previous AO-based reformulation of the MP2 energy gradient[48], Paper **I** of this work introduces an improved implementation by employing the resolution-of-the-identity (RI) approximation[49–55] and a Cholesky decomposition of the density matrices (CDD)[56,57], as was previously done in the realm of the energy computation[58,59]. Extending the distance-including QQR-type integral estimation by Maurer *et al.*[42,60] to the computation of MP2 energy gradients allows the exploitation of the locality of the atomic orbitals. This builds the foundation for a computation of hyperfine coupling constants (HFCCs) in the absence of spin-orbit coupling. Paper **II** improves on the efficient implementation of HFCCs in Paper **I** by taking into account the locality of the spin density in various realistic systems. By computing the HFCCs of selected nuclei, the computational cost is reduced further. When only the opposite spin contribution in the scaled opposite spin (SOS) approximation[61] is computed, the wall time required for the computation is reduced by migrating to graphics processing units (GPUs), as shown in computations on a realistic biomolecular system.

As outlined in Sec. 2.3.2, the choice of gauge including atomic orbitals (GIAOs) in the computation of magnetic properties, such as the electronic g-tensor and the NMR shielding, is vital. While the effect is known in the computation of NMR shieldings[45], it was assumed to be small or negligible in the case of the electronic g-tensor[62–67]. The work presented in Paper **III** contradicts this commonly adopted assumption by analyzing in detail the effect on larger molecular systems.

Furthermore, vibrational averaging significantly influences the accuracy of the computed molecular property. While this has been studied for small molecules at the DFT-level, Manuscript **IV** presents a thorough analysis at the DH-DFT and MP2-level of the influence of both molecular dynamics and electron correlation on the computed HFCCs using the efficient implementations presented in Papers **I** and **II**.

To allow for a description of dynamic effects in the computation of NMR shieldings, a low scaling efficient implementation is required. While Maurer and Ochsenfeld[68] presented a (sub-)linear scaling reformulation of the MP2 NMR shieldings in atomic orbitals, the pilot implementation entails a significant prefactor in the overall scaling behavior and a large dependency of the computational cost on the basis set size for fixed molecular system sizes. Therefore, extending the RI-CDD approach of the analytic gradients in Paper **I** to the computation of NMR shieldings might be beneficial. First results are discussed in Sec. 4.2.3, indicating significant computational savings.

This thesis will start by briefly reviewing HF and Møller-Plesset perturbation theory and subsequently summarizing the main ideas to reduce the computational cost: the reformulation in atomic orbitals and the screening of the integral products using distance-dependent estimates, the Cholesky decomposition, and the RI approximation. After this theoretical background, a brief summary of the findings in Paper **I-III** and in Manuscript **IV** is presented in Chapters 3 and 4.

2 Theoretical Background

Calculating the energy of molecular systems forms the basis for the computation of molecular properties. This chapter therefore starts by introducing the fundamental HF theory in Sec. 2.1, followed by an overview of the main equation and the atomic orbital-based approach to reduce the scaling behavior of Møller-Plesset perturbation theory in Sec. 2.2, and ending with a brief presentation of the theory to compute molecular properties based on the energy expression in Sec. 2.3.

2.1 Hartree-Fock theory

In order to obtain static molecular properties, the time-independent Schrödinger equation needs to be solved. Here, the non-relativistic molecular Hamiltonian $\hat{H}$ is composed of the kinetic energy of the nuclei and of the electrons, and of the potential energy contribution of the electron-electron, the electron-nuclei, and the nuclei-nuclei interaction. In atomic units, the Hamiltonian is written as:

$$\hat{H} = -\sum_A \frac{1}{2M_A}\nabla_A^2 - \sum_i \frac{1}{2}\nabla_i^2 + \sum_{j>i} \frac{1}{r_{ij}} - \sum_i \sum_A \frac{Z_A}{r_{iA}} + \sum_{B>A} \frac{Z_A Z_B}{R_{AB}}, \qquad (2.1)$$

where nuclei are represented by capital Latin letters, whereas lowercase Latin letters are used for electrons. M_A is the mass of nucleus A, Z_A is its charge, and R and r indicate the respective distances between the nuclei, the electrons, and the electrons to the nuclei.

The motion of the nuclei is considerably slower than that of the electrons. The common Born-Oppenheimer approximation thus views the motion of the electrons in a field of static nuclei.[69] The wavefunction is described as a product of a nuclear $\chi_n(\mathbf{R})$ and an electronic wavefunction $\psi_n(\mathbf{r}; \mathbf{R})$ of state n:

$$\Psi(\mathbf{r}, \mathbf{R}) = \sum_n \psi_n(\mathbf{r}; \mathbf{R})\chi_n(\mathbf{R}). \qquad (2.2)$$

Analogously, the Hamiltonian is separated into the operator of the kinetic energy of the

nuclei $\hat{T}_{\text{nuc}}$, the nuclear-nuclear repulsion $\hat{V}_{\text{NN}}$, and the remaining electronic Hamiltonian $\hat{H}_{\text{el}}$:

$$\hat{H} = \hat{T}_{\text{nuc}} + \hat{V}_{\text{NN}} + \hat{H}_{\text{el}}. \tag{2.3}$$

In the Born-Oppenheimer approximation one assumes the change of the electronic wavefunction with respect to the nuclear coordinates to be slow and thus negligible, leading to:

$$\left(\hat{V}_{\text{NN}}(\mathbf{R}) + \hat{T}_{\text{nuc}} + E_{\text{el}}(\mathbf{R})\right) \chi_n(\mathbf{R}) = E\chi_n(\mathbf{R}), \tag{2.4}$$

where $\hat{V}_{\text{NN}}(\mathbf{R})$ is the nuclear-nuclear interaction, and $E_{\text{el}}(\mathbf{R})$ is the eigenvalue of the corresponding electronic Schrödinger equation:

$$\hat{H}_{\text{el}}\psi_n(\mathbf{r};\mathbf{R}) = V_n(\mathbf{R})\psi_n(\mathbf{r};\mathbf{R}). \tag{2.5}$$

In HF theory[1–3], the wavefunction is approximated by a *single* Slater determinant $|\Psi\rangle$ built from an orthonormal set of molecular orbitals (MOs) $\{|\varphi\rangle\}$ to incorporate the antisymmetry of fermionic wavefunctions. These MOs are products of a spatial single-electron wavefunction and a spin function. In the closed-shell case, the common restricted HF ansatz dictates that each spatial orbital is occupied by two electrons of opposite spin. For open-shell systems, the unrestricted HF (UHF) ansatz[70] is a common choice, which can straightforwardly be applied in the computation of molecular properties but may introduce spin contamination as the wavefunction is not necessarily an eigenfunction of the $\hat{S}^2$ operator. Therefore, the UHF wavefunction may contain non-physical contributions from higher states.

Based on the Slater determinant, the MOs are optimized by minimizing the expectation value of the Slater determinant according to the variation principle, resulting in the canonical HF equations:

$$\hat{F}(i)|\varphi\rangle = \varepsilon_i|\varphi\rangle, \tag{2.6}$$

where ε_i is the orbital energy and $\hat{F}$ the Fock operator. This operator can be partitioned into a one-electron Hamiltonian $\hat{h}$, taking into account the kinetic energy of the electrons and the electron-nuclei interaction, and a two-electron contribution, where both operators for the classical Coulomb electron-electron interaction $\hat{J}_j$ and the non-classical exchange

energy $\hat{K}_j$ of the other electrons j are considered:

$$\hat{F}(i) = \hat{h}(i) + \sum_{j>i} \left(\hat{J}_j(i) - \hat{K}_j(i) \right). \tag{2.7}$$

The two-electron contribution thus behaves like a one-electron operator, where the other electrons are only considered in a mean field ansatz. This results in the missing so-called correlation energy, which is defined as the difference between the exact energy and the HF energy.

At this point it is worth mentioning that the Kohn-Sham (KS) theory, which forms the basis of modern DFT computations, is analogous to the canonical HF equation in Eq. 2.6, albeit with two exceptions: first, the two-electron contributions are evaluated as a function of the electron density ρ and, second, the exchange operator is replaced by the exchange-correlation (XC) potential, whose exact form is unknown but for which a variety of approximations exist. [12,71]

To solve Eq. 2.6, the molecular orbitals are often described by a linear combination of atomic orbitals (LCAO) $\{|\chi_\mu\rangle\}$:

$$|\varphi_i\rangle = \sum_\mu c_{\mu i} |\chi_\mu\rangle, \tag{2.8}$$

which leads to the Roothaan-Hall equations [72,73]:

$$\sum_\nu F_{\mu\nu} C_{\nu i} = \sum_\nu S_{\mu\nu} C_{\nu i} \varepsilon_i, \tag{2.9}$$

with the Fock matrix element $F_{\mu\nu} = \langle \mu | \hat{F} | \nu \rangle$ and the overlap $S_{\mu\nu} = \langle \mu | \nu \rangle$.[1] Accordingly, the HF energy can be given as:

$$E_{\text{HF}} = \sum_{\mu\nu} \left[h_{\mu\nu} P_{\mu\nu} + \frac{1}{2} P_{\mu\nu} G_{\mu\nu}(\mathbf{P}) \right] + V_{\text{nuc-nuc}}, \tag{2.10}$$

with the density matrix:

$$P_{\mu\nu} = \sum_{i \in \text{occ}} C_{\mu i}^* C_{\nu i}. \tag{2.11}$$

The term $G_{\mu\nu}(\mathbf{P})$ is the matrix representation of the two-electron contribution (or of the KS potential in KS-DFT), which itself depends on the density matrix. Due to this explicit dependence of the potential $\mathbf{G}$ on the density matrix, the HF energy, as well as the KS-

---

[1]In the following, the basis functions $\{|\chi_\mu\rangle\}$ will be abbreviated by $\{|\mu\rangle\}$.

energy, needs to be obtained iteratively in a self-consistent field (SCF) ansatz.

2.2 Møller-Plesset perturbation theory

One approach to describe the missing electron correlation is perturbation theory. Hereby, the Hamiltonian $\hat{H}$ is partitioned in two parts:

$$\hat{H} = \hat{H}_0 + \lambda \hat{V}, \tag{2.12}$$

where $\hat{H}_0$ is a Hamiltonian with known eigenfunctions and eigenvalues, and $\hat{V}$ is the perturbation operator.[8] The perturbation parameter λ is introduced for the following derivation. Inserting the partitioned Hamiltonian in the Schrödinger equation results in:

$$\left(\hat{H}_0 + \lambda \hat{V} \right) |\Psi_n\rangle = \mathcal{E}_n |\Psi_n\rangle. \tag{2.13}$$

Subsequently, the eigenfunctions and eigenvalues can be expanded in a Taylor series:

$$\begin{aligned} \mathcal{E}_n &= E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \cdots \\ |\Psi_n\rangle &= |\Psi_n^0\rangle + \lambda |\Psi_n^1\rangle + \lambda^2 |\Psi_n^2\rangle + \cdots \end{aligned} \tag{2.14}$$

Inserting Eq. 2.14 in Eq. 2.13, the following expressions for the i-th order energies can be derived:

$$\begin{aligned} E_n^{(0)} &= \langle \Psi_n^{(0)} | \hat{H}_0 | \Psi_n^{(0)} \rangle \\ E_n^{(1)} &= \langle \Psi_n^{(0)} | \hat{V} | \Psi_n^{(0)} \rangle \\ E_n^{(i)} &= \langle \Psi_n^{(0)} | \hat{V} | \Psi_n^{(i-1)} \rangle. \end{aligned} \tag{2.15}$$

In the quantum chemical context, this has been put forward by Møller and Plesset[7], where the perturbation operator $\hat{V}$ is the difference between the electronic Hamiltonian in the Born-Oppenheimer approximation (see Sec. 2.1) and the sum of all Fock-operators in Eq. 2.7. Therefore, the unperturbed energy is the sum of the orbital energies (see Eq. 2.6), while the sum of the unperturbed energy and first-order perturbation energy is the HF energy. A first description of correlation effects is obtained with second-order Møller-Plesset perturbation theory (MP2), which is the most employed order of perturbation and often viewed as a good compromise between accuracy and computational cost, leading to accurate results, e.g., for hydrogen bonding energies[74,75] and NMR shielding tensors[47,76,77]. Qualitatively, MP2 describes pair-wise electron correlation. It should be noted that the

MPn-series is not convergent, i.e., taking into account higher orders does not necessarily improve the accuracy.[78–82]

MPn methods are also used to describe open-shell systems, most straightforwardly when based on an UHF wavefunction.[83] Due to the spin contamination, UMPn results must be treated with care. Furthermore, MP2 has been formulated as a post-KS method in Grimme's DH-DFT,[14] which, when based on an unrestricted KS determinant, can resolve some of the spin contamination issues.[84]

2.2.1 Canonical representation

The canonical closed-shell MP2 energy reads as follows:

$$E_{\mathrm{MP2}} = -\sum_{ij} \sum_{ab} \frac{(ia|jb)[2(ia|jb) - (ib|ja)]}{\varepsilon_a + \varepsilon_b - \varepsilon_i - \varepsilon_j}, \tag{2.16}$$

where i, j represent occupied molecular orbitals and a, b virtual orbitals. To obtain the two-electron integrals in Eq. 2.16, the following transformations are necessary:

$$(ia|jb) = \sum_{\mu\nu\lambda\sigma} c_{\mu i} c_{\nu a} (\mu\nu \mid \lambda\sigma) c_{\lambda j} c_{\sigma b}. \tag{2.17}$$

When this transformation is performed consecutively, its formal scaling behavior is $\mathcal{O}(\mathrm{N}^5)$, where N is the number of basis functions. This is rate-determining and limits the applicability of canonical MP2 to small- to medium-sized molecules. This issue has been researched over the past decades, resulting in different ansätze to overcome the large scaling behavior. What they have in common is that they aim to exploit the locality of the correlation contribution, which is not taken into account in the canonical formulation in Eq. 2.17. This can be achieved by using local orbitals, as pioneered by Pulay and Saebø[18–20,85]. A different approach is partitioning the system either directly by fragmenting the molecule in the fragment molecular orbital (FMO) method[33–37], or by separating the orbital space in the divide-expand-consolidate (DEC) ansatz[27–32]. This work focuses on a third ansatz, which builds on a reformulation of the MP2 energy in atomic orbitals $\{\chi_\mu\}$, thus exploiting their locality.

2.2.2 Reformulation in atomic orbitals

For a reformulation of the MP2 energy expression in Eq. 2.16 in atomic orbitals the energy denominator needs to be avoided. This can be achieved by a Laplace-transformation, as introduced by Almlöf and Häser[38–40]:

$$\frac{1}{x} = \int_0^\infty e^{-xt}dt \approx \sum_{\zeta=1}^\tau \omega_\zeta e^{-xt_\zeta} \quad \rightarrow \quad \frac{1}{\varepsilon_a + \varepsilon_b - \varepsilon_i - \varepsilon_j} \approx \sum_{\zeta=1}^\tau \omega_\zeta e^{-\varepsilon_a t_\zeta} e^{-\varepsilon_b t_\zeta} e^{\varepsilon_i t_\zeta} e^{\varepsilon_j t_\zeta},$$

(2.18)

with the linear and exponential Laplace expansion coefficients ω_ζ and t_ζ. With this Laplace transformation, the MP2 equation can be rewritten as:

$$
\begin{aligned}
E_{\text{AO-MP2}} &= -\sum_{\zeta=1}^\tau \omega_\zeta \mathcal{E}_\zeta \\
&= -\sum_{\zeta=1}^\tau \omega_\zeta \sum_{\mu\nu\lambda\sigma} \left(\underline{\mu}\overline{\nu} \mid \underline{\lambda}\overline{\sigma} \right) \left[2 \left(\mu\nu \mid \lambda\sigma \right) - \left(\mu\sigma \mid \lambda\nu \right) \right],
\end{aligned}
$$

(2.19)

where $|\underline{\mu}\rangle$ and $|\overline{\nu}\rangle$ are transformed with occupied and virtual pseudodensities:

$$
\begin{aligned}
|\underline{\mu}\rangle &= \sum_{\mu'} \underline{P}_{\mu'\mu} |\mu'\rangle \\
|\overline{\nu}\rangle &= \sum_{\nu'} \overline{P}_{\nu'\nu} |\nu'\rangle.
\end{aligned}
$$

(2.20)

The pseudodensities arise from the Laplace transformation:

$$
\begin{aligned}
\underline{P}_{\mu'\mu} &= \sum_i c_{\mu'i} e^{\varepsilon_i t_\zeta} c_{\mu i} \\
\overline{P}_{\nu'\nu} &= \sum_a c_{\nu'i} e^{-\varepsilon_a t_\zeta} c_{\nu i}.
\end{aligned}
$$

(2.21)

To fully exploit the locality of the atomic orbitals and to achieve linear-scaling behavior, an efficient and accurate screening method for the required integral products is necessary. This will be outlined in the following section.

2.2.3 Distance-dependent screening of significant two-electron integrals

Preselecting significant two-electron integrals in the computation of the AO-MP2 contribution is necessary to achieve a reduction in the scaling behavior. A rigorous upper bound for the value of a specific two-electron integral is the classical Schwarz estimate:[86]

$$|(\mu\nu \mid \lambda\sigma)| \leq \underbrace{|(\mu\nu \mid \mu\nu)|^{\frac{1}{2}}}_{Q_{\mu\nu}}\underbrace{|(\lambda\sigma \mid \lambda\sigma)|^{\frac{1}{2}}}_{Q_{\lambda\sigma}}, \tag{2.22}$$

with the Schwarz-matrices $\mathbf{Q}$. When screening two-electron integrals using these Schwarz estimates, only a quadratic scaling behavior can be achieved since with the common Gaussian basis functions only a linear number of basis functions $\{\nu\}$ surrounding μ will have significant overlap. Thus, a linear number of basis function products is significant in both in the bra and the ket, i.e., a quadratic number. To accurately screen transformed two-electron integrals as in Eq. 2.19, the Schwarz estimates in Eq. 2.22 need to be transformed analogously.

To further reduce the scaling behavior, the dependence of the integral value on the distance between the bra and the ket charge distribution needs to be considered by a multipole expansion.[87,88] To this end, Maurer *et al.*[42,60] introduced a new screening ansatz named QQR, which incorporates the distance-dependence of the two electron integral over the charge distributions Ω_{bra} and Ω_{ket}:

$$(\Omega_{\text{bra}} \mid \Omega_{\text{ket}}) = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\sum_{i=-m}^{m}\sum_{j=-n}^{n}\frac{q_{mi}^{\text{bra}}T'_{mi,nj}q_{nj}^{\text{ket}}}{R_{\text{bra-ket}}^{m+n+1}}, \tag{2.23}$$

with the m-th order multipoles q_{mi}^{A} of the charge distribution A and the individual spherical multipoles i. $R_{\text{bra-ket}}$ indicates the distance between the two charge distributions and T' takes into account the relative orientation of the multipole moments.

The monopoles of transformed charge distributions are zero due to the orthogonality of the occupied and the virtual subspace:

$$q_{00}^{\underline{\mu}\bar{\nu}} = S_{\underline{\mu}\bar{\nu}} = \sum_{\mu'\nu'}\underline{P}_{\mu\mu'}S_{\mu'\nu'}\overline{P}_{\nu'\nu} = 0. \tag{2.24}$$

Therefore, when both bra and ket are fully transformed, as in Eq. 2.19, all terms with $m < 1$ or $n < 1$ will be zero, resulting in the slowest distance dependency being $R_{\text{bra-ket}}^3$.

As proposed by Maurer *et al.*[42], this can be exploited by screening the fully-transformed two-electron integral as follows:

$$\left(\underline{\mu}\overline{\nu} \mid \underline{\lambda}\overline{\sigma}\right) \approx \frac{Z_{\underline{\mu}\overline{\nu}} Z_{\underline{\lambda}\overline{\sigma}}}{\left(R - \text{ext}_{\underline{\mu}\overline{\nu}} - \text{ext}_{\underline{\lambda}\overline{\sigma}}\right)^3}. \tag{2.25}$$

Hereby, Häser's pseudo-Schwarz matrices $\mathbf{Z}$ are employed.[40] To correctly describe the distance between the charge distributions, their extents need to be taken into account. A thorough discussion of the extents can be found in Ref. 42. At this point it has to be noted that QQR estimates do not represent a rigorous upper bound although they do provide a fully controllable accuracy.[42] Thompson *et al.*[89] recently introduced distance-including rigorous upper bounds, which, at this stage, are not yet applicable to AO-MP2, as they only incorporate the distance-dependency of the operator. They do not include the additional distance dependency due to the orthogonality of the occupied and virtual subspace.

2.2.4 Cholesky decomposition

While linear scaling behavior can be achieved with the aforementioned QQR screening technique, widespread application is hampered by a large computational prefactor and dependence on the size of the basis set. This can, in part, be explained by the need to screen integral products, to access and process lists of significant integral products, and to evaluate the MP2 contribution per Laplace point. However, a further vital aspect needs to be considered: while for large molecule sizes the sum in Eq. 2.19 involves a linear number of atomic orbitals, when only the basis set is increased the result is an $\mathcal{O}(\mathrm{N}^5)$ behavior. The canonical formulation in Eq. 2.16 only involves the summation of $\mathrm{N}_{\text{occ}}^2 \cdot \mathrm{N}_{\text{virt}}^2$ orbitals, resulting in a $\mathrm{N} \cdot \mathrm{N}_{\text{occ}}^2 \cdot \mathrm{N}_{\text{virt}}^2$ scaling behavior in the transformation. For large basis sets, the number of occupied orbitals will be quasi constant, and the virtual orbital space alone increases with an enlarged basis set size, resulting in an asymptotically $\mathcal{O}(\mathrm{N}^3)$ scaling behavior.

Therefore, exploiting the rank deficiency of the density matrices is vital, reconnecting the summation in Eq. 2.19 to the number of electrons in the system. A plethora of methods reducing the rank deficiency exist. In the case of the AO-MP2 equation a method that preserves the locality of the (pseudo-)density matrices is necessary. This can be achieved by a Cholesky decomposition[56,57], which was previously applied to the computation of AO-MP2 energies[58,59].

Hereby, the pseudodensities are decomposed as:

$$\underline{P}_{\mu'\mu} = \sum_i \underline{L}_{\mu'i}\underline{L}_{\mu i}$$
$$\overline{P}_{\nu'\nu} = \sum_a \overline{L}_{\nu'a}\overline{L}_{\nu a}, \tag{2.26}$$

where the sums run over the ranks of the respective pseudodensity matrices. It has to be noted that such a Cholesky decomposition is only possible when the matrix is positive semi-definite. The pseudodensities satisfy this by definition.[59] As shown by Maurer *et al.*[59], the numerical rank of the pseudodensity is equal to or smaller than the rank of the density matrices, i.e., the size of the occupied or the virtual orbital subspace. When decomposing all pseudodensities in Eq. 2.19, one obtains the following contribution per Laplace point ζ of the AO-MP2 energy:

$$\mathcal{E}_\zeta = \sum_{ij}^{\mathrm{occ}} \sum_{ab}^{\mathrm{virt}} \left(\underline{i}\overline{a} \mid \underline{j}\overline{b} \right) \left[2 \left(\underline{i}\overline{a} \mid \underline{j}\overline{b} \right) - \left(\underline{i}\overline{b} \mid \underline{j}\overline{a} \right) \right]. \tag{2.27}$$

Therefore, when the basis set size is increased, a formal scaling behavior similar to the canonical formulation is obtained. Furthermore, the orthogonality of the occupied and virtual subspace is preserved, allowing for an efficient integral estimation using the QQR-technique (see Sec. 2.2.3). The preservation of the sparsity of the pseudodensities is shown in the work by Maurer *et al.*,[59] who introduced the term local pseudo-MOs (LPMOs) to describe the orbitals i, j and a, b. To clearly differentiate with respect to local MO methods, later work employed the term local Cholesky orbitals.[90,91]

2.2.5 Resolution-of-the-identity approximation

A further bottleneck of a widespread application of AO-based MP2 formulations is their large prefactor in the scaling behavior. An important ansatz reducing the computational prefactor as well as memory requirements is the RI approximation (also known as density fitting).[49–55] Hereby, the two-electron integral is described by:

$$(\mu\nu \mid \lambda\sigma) = \sum_{P,Q}^{\mathrm{N_{aux}}} (\mu\nu \mid P)(P \mid Q)^{-1}(Q \mid \lambda\sigma), \tag{2.28}$$

using the auxiliary basis set $\{P\}$. Thus, the two-electron integral can be described as a contraction of two three-center quantities:

$$
(\mu\nu \mid \lambda\sigma) = \sum_P B_{\mu\nu}^P B_{\lambda\sigma}^P
$$

$$
\text{with:} \quad B_{\mu\nu}^P = \sum_Q (\mu\nu \mid Q)\,(Q \mid P)^{-\frac{1}{2}}. \tag{2.29}
$$

This is exact for auxiliary basis sets that fully span the space of the integral products. For finite-sized basis sets, this approximation is typically accurate with respect to relative energies and molecular properties but, however, not for absolute energies. Therefore, a recent work by Schurkus *et al.*[92] reduces the error with respect to absolute energies by projecting out the unphysical parts of an oversized auxiliary basis set.

When applying the RI-approximation in the AO-based CDD-MP2 computation in Eq. 2.27, the fully transformed two-electron integrals can be computed by:

$$
\left(\underline{i}\overline{a} \mid \underline{j}\overline{b}\right) = \sum_P B_{\underline{i}\overline{a}}^P B_{\underline{j}\overline{b}}^P, \tag{2.30}
$$

where the transformed three-center $\mathbf{B}$ matrices are obtained by step-wise transformation with the Cholesky decomposed pseudodensities. The computation of the MP2 energy using this RI-CDD approach has been proven to be highly efficient, outperforming canonical implementations with decent basis sets for medium-sized molecules and providing manifold speed-ups for large molecular systems, such as a DNA strand consisting of up to eight base pairs.[59]

2.3 Molecular properties

2.3.1 Analytic derivative theory

The computation of molecular properties using quantum chemical methods is an important tool linking experiment and theory. As described by Gauss,[45] one can distinguish between three types of molecular properties: properties, (i) that are related to energy differences, such as reaction energies, dissociation energies; (ii) that are specific for a given electronic state, such as NMR chemical shifts, g-tensors, vibrational frequencies; and (iii) that are characterized by transitions between electronic states, such as excitation energies, transition strength, and electron affinities.

The focus of this work is on the second type of molecular properties, i.e., those that depend on a specific electronic case. These properties describe the response of the molecular system to an external perturbation, such as an external magnetic field. Assuming this perturbation to be weak, a Taylor expansion can be used to describe the energy in the presence of the perturbation ξ:

$$E(\xi) = E(\xi = 0) + \left.\frac{dE}{d\xi}\right|_{\xi=0} \xi + \frac{1}{2}\left.\frac{d^2E}{d\xi^2}\right|_{\xi=0} \xi^2 + \ldots \tag{2.31}$$

When considering the physical nature of the interaction with the perturbation, the different orders can be assigned to molecular properties. For example, when the perturbation is an external electric field, the first-order term is the dipole moment, whereas the second derivative describes the polarizability and the third the first hyperpolarizability, and so forth. As a note, the computation of the molecular property as an expectation value within the Hellmann-Feynman theorem[43,44] is not sufficient as it does not hold for approximate wave functions. Similarly, numerical differentiation is often insufficient.[45]

2.3.2 Magnetic properties

Both EPR and NMR spectroscopy involve the splitting of energy states in an external field: in the case of EPR, this involves the splitting of electronic spin states, whereas in NMR spectroscopy the loss of the degeneracy of nuclear spin states is studied. Therefore, the description of parameters of both spectroscopies involves an external magnetic field. In order to describe this perturbation, one has to extend the Hamiltonian. The interaction of the magnetic field with the magnetic moments, which are a result of the motion of the electrons, changes the kinetic energy operator. Hence, the momentum operator $\hat{p}$ is replaced by the generalized momentum operator $\hat{\pi}$ according to the principle of minimal electromagnetic coupling:[93,94]

$$\hat{\pi} = \hat{p} + \mathbf{A}(\mathbf{r}), \tag{2.32}$$

where the vector potential $\mathbf{A}$ is given by:

$$\mathbf{A}(\mathbf{r}) = \underbrace{\frac{1}{2}\mathbf{B} \times (\mathbf{r} - \mathbf{R}_0)}_{\mathbf{A}_0} + \sum_j \underbrace{\alpha^2 \frac{\mathbf{m}_j \times (\mathbf{r} - \mathbf{R}_j)}{|\mathbf{r} - \mathbf{R}_j|^3}}_{\mathbf{A}_j}, \tag{2.33}$$

taking into account both the external magnetic field $\mathbf{B}$ and the nuclear magnetic moments $\mathbf{m}_j$ of each nucleus j. α is the fine structure constant. Consequently, the magnetic field is given by:

$$\mathbf{B} = \nabla \times \mathbf{A} = \underbrace{\nabla \times \mathbf{A}_0}_{\mathbf{B}_0} + \sum_j \underbrace{\nabla \times \mathbf{A}_j}_{\mathbf{B}_j} \tag{2.34}$$

with the nuclear magnetic field $\mathbf{B}_j$.

If the interaction between the total electronic spin s_i and the magnetic field is to be considered, the Hamiltonian becomes:

$$\hat{H} \rightarrow \hat{\mathcal{H}} = \hat{H} - \sum_i \mathbf{B}_i \cdot \mathbf{s}_i. \tag{2.35}$$

The derivative of this additional term with respect to the nuclear magnetic moment $\mathbf{m}_j$ results in two terms: the anisotropic spin-dipole interaction and the isotropic Fermi-contact term.[95] In the absence of spin-orbit coupling, these two terms are the hyperfine coupling. As such, HFCCs can be obtained as a first-order property.

When treating magnetic properties, one additional issue needs to be addressed: the magnetic field is not uniquely defined by the vector potential $\mathbf{A}$. The choice of the gauge origin $\mathbf{R}_0$ in Eq. 2.33 does not change the results in a complete basis set. This, however, is not the case for truncated basis sets.[45,96] Here, the dependence on the choice of gauge-origin can be strong. Thus, gauge including atomic orbitals (GIAOs) are necessary:

$$\chi_\mu(\mathbf{B}) = \chi_\mu(0) exp\left(-\frac{i}{2}\mathbf{B} \times (\mathbf{R}_\mu - \mathbf{R}_0) \cdot \mathbf{r}\right) \tag{2.36}$$

where the field dependency of the basis function is included in the exponential gauge prefactor.[97–101]

3 First-order properties at the MP2-level

This chapter introduces the main theoretical developments and equations from the work on first-order properties shown in Paper **I** and **II** and briefly summarizes the results presented in Manuscript **IV**.

3.1 Analytical RI-CDD MP2 energy gradients

The computation of first-order properties at the MP2-level is challenging. While the first fully AO-based formulation of analytic energy gradients was put forward by Schweizer *et al.*,[48] the lack of an efficient implementation prohibits the testing and application on realistic systems. In order to achieve high computational efficiency with a reduced scaling behavior, both the RI-CDD ansatz and an extended QQR-screening is mandatory. This is shown and analyzed in detail in a first efficient implementation in Paper **I** included in this work. This section will briefly introduce analytical AO-MP2 gradients and then summarize the new RI-CDD ansatz in their computation. The reader is referred to the attached Paper **I** for a more detailed and thorough description and discussion.

Starting from Eq. 2.19, the analytic gradient with respect to a perturbation ξ involves two terms: the perturbation of the pseudodensities and the perturbation of the basis functions. The terms are separated as follows:

$$
\begin{aligned}
\frac{\partial E_{\text{AO-MP2}}}{\partial \xi} &= \sum_{\zeta=1}^{\tau} \omega_\zeta \frac{\partial \mathcal{E}_{\text{AO-MP2}}}{\partial \xi} \\
&= \sum_{\zeta=1}^{\tau} \left(2 \sum_{\mu'\mu} \overline{R}_{\mu'\mu}^{\zeta} \frac{\partial \underline{P}_{\mu'\mu}^{\zeta}}{\partial \xi} + 2 \sum_{\nu'\nu} \underline{R}_{\nu'\nu}^{\zeta} \frac{\partial \overline{P}_{\nu'\nu}^{\zeta}}{\partial \xi} + 2 \mathcal{I}_\zeta^\xi \right),
\end{aligned}
\tag{3.1}
$$

where the perturbation of the basis functions is included in $\mathcal{I}_\zeta^\xi$:

$$\mathcal{I}_\zeta^\xi = \sum_{\mu\nu\lambda\sigma} \frac{\partial (\mu\nu \mid \lambda\sigma)}{\partial \xi} \left[2 \left(\underline{\mu}\overline{\nu} \mid \underline{\lambda}\overline{\sigma} \right) - \left(\underline{\mu}\overline{\sigma} \mid \underline{\lambda}\overline{\nu} \right) \right], \tag{3.2}$$

and where the perturbed pseudodensities are contracted with the triple-transformed **R**-matrices:

$$\begin{aligned}
\underline{R}_{\nu'\nu}^\zeta &= \sum_{\mu\lambda\sigma} \left(\underline{\mu}\nu' \mid \underline{\lambda}\overline{\sigma} \right) \left[2 \left(\mu\nu \mid \lambda\sigma \right) - \left(\mu\sigma \mid \lambda\nu \right) \right], \\
\overline{R}_{\mu'\mu}^\zeta &= \sum_{\nu\lambda\sigma} \left(\mu'\overline{\nu} \mid \underline{\lambda}\overline{\sigma} \right) \left[2 \left(\mu\nu \mid \lambda\sigma \right) - \left(\mu\sigma \mid \lambda\nu \right) \right].
\end{aligned} \tag{3.3}$$

As shown in Ref. 48 in a pilot implementation, the number of significant integrals scales linearly. This can, however, only be exploited by an extended QQR-type integral estimation, which takes into account half-transformed charge distributions as presented in Paper **I**.

Issues challenging the computation of AO-MP2 energies similarly apply to their analytical gradients. Therefore, benefits can analogously be expected from the RI-CDD ansatz. By introducing a Cholesky decomposition of the pseudodensities, the $\mathcal{I}_\zeta^\xi$ term and the **R**-matrices read as:

$$\begin{aligned}
\mathcal{I}_\zeta^\xi &= \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} \left(\underline{i}\overline{a} \mid \underline{j}\overline{b} \right)^{(\xi)} \left[2 \left(\underline{i}\overline{a} \mid \underline{j}\overline{b} \right) - \left(\underline{i}\overline{b} \mid \underline{j}\overline{a} \right) \right], \\
\underline{R}_{\nu'\nu}^\zeta &= \sum_{ij}^{\text{occ}} \sum_{a}^{\text{virt}} \left(\underline{i}\nu' \mid \underline{j}\overline{b} \right) \left[2 \left(\underline{i}\nu \mid \underline{j}\overline{b} \right) - \left(\underline{i}\overline{b} \mid \underline{j}\nu \right) \right], \\
\overline{R}_{\mu'\mu}^\zeta &= \sum_{i}^{\text{occ}} \sum_{ab}^{\text{virt}} \left(\mu'\overline{a} \mid \underline{j}\overline{b} \right) \left[2 \left(\mu\overline{a} \mid \underline{j}\overline{b} \right) - \left(\mu\overline{b} \mid \underline{j}\overline{a} \right) \right],
\end{aligned} \tag{3.4}$$

where the perturbation in parentheses indicates a perturbation only with respect to the basis functions.

To reduce the computational prefactor, the transformed two-electron integrals are obtained using the RI-approximation. In contrast to the evaluation of the RI-CDD MP2 energy, a **B**-matrix transformed solely with the virtual Cholesky matrix is required. Furthermore, the perturbation of the two-electron integral itself needs to be accounted for. Its RI-approximated description reads as:

$$\frac{\partial (\mu\nu \mid \lambda\sigma)}{\partial \xi} = \sum_P \left[\frac{\partial B_{\mu\nu}^P}{\partial \xi} B_{\lambda\sigma}^P + B_{\mu\nu}^P \frac{\partial B_{\lambda\sigma}^P}{\partial \xi} \right] - \sum_{PQ} C_{\mu\nu}^P (P \mid Q)^\xi C_{\lambda\sigma}^Q$$

$$\text{with:} \quad \frac{\partial B_{\mu\nu}^P}{\partial \xi} = \sum_Q \frac{\partial (\mu\nu \mid Q)}{\partial \xi} (Q \mid P)^{-\frac{1}{2}} \tag{3.5}$$

$$\text{and:} \quad C_{\mu\nu}^P = \sum_Q (\mu\nu \mid Q)(Q \mid P)^{-1}.$$

This perturbed two-electron integral is fully transformed in the I_ζ^ξ term in Eq. 3.4. To circumvent the expensive transformation per perturbation, all perturbation-independent contributions are collected in terms Γ:

$$\mathcal{I}_\zeta^\xi = \sum_{\mu\nu} \sum_P \Gamma_{\mu\nu}^P \frac{\partial B_{\mu\nu}^P}{\partial \xi} + \sum_{\lambda\sigma} \sum_P \Gamma_{\lambda\sigma}^P \frac{\partial B_{\lambda\sigma}^P}{\partial \xi} - \Gamma^\xi(\zeta). \tag{3.6}$$

A definition of the perturbation-independent contributions Γ and of Γ^ξ can be found in Paper **I**.

By using the extended QQR-type integral estimation, a low-scaling, efficient algorithm is possible. The computation of the perturbed pseudodensity per perturbation can be circumvented by the Z-vector technique[102,103] in its AO-formulation[48] using the efficient density matrix-based Laplace-transform coupled perturbed SCF (DL-CPSCF)[104]. In the implementation presented in Paper **I**, the nuclear gradient of linear alkanes larger than $C_{60}H_{122}$ is obtained faster than with the canonical implementation. Due to the reduced scaling behavior, the speed-up increases for larger systems and is confirmed with glycine chains.

While nuclear gradients are important for, e.g., structure optimizations or accurate molecular dynamics, other first-order properties can similarly be obtained using the new RI-CDD ansatz. As outlined in Sec. 2.3.2, the HFCCs can be viewed in the absence of spin-orbit coupling as a first-order property, where the perturbation is the isotropic Fermi-Contact interaction and an anisotropic spin-dipole coupling. For this, open-shell RI-CDD gradients are necessary. A simple ansatz to describe open-shell properties is to base the computation on an unrestricted Hartree-Fock (UHF) wave function. It has to be noted that although this is mostly straightforward and can rely in large parts on respective developments in the closed-shell case, a drawback is the possible spin contamination[1]. Nonetheless, many systems exhibit only a small spin contamination. This can be improved

---

[1]The unrestricted slater determinant is not a satisfactory eigenfunction of the spin operator $\hat{S}^2$.

upon by double-hybrid density functional theory, which combines DFT and MP2.[14] The derivations for MP2 are equally transferable to DH-DFT with the same equations being solved using Kohn-Sham orbitals. With this, spin contamination issues can often be overcome[84] and reliable HFCCs can be obtained[46].

Based on an UHF-wavefunction, the RI-CDD-UMP2 energy gradient reads as:

$$
\frac{\partial E_{\text{RI-CDD-MP2}}}{\partial \xi} = -\frac{1}{2} \sum_{\zeta=1}^{\tau} \omega_\zeta \sum_{\eta}^{\alpha,\beta} \frac{\partial \mathcal{E}^\eta}{\partial \xi}
$$

$$
\text{with:} \quad \frac{\partial \mathcal{E}^\eta}{\partial \xi} = 2 \sum_{\mu'\mu} \overline{R}_{\mu'\mu}^{\zeta,\eta} \frac{\partial P_{\mu'\mu}^{\zeta,\eta}}{\partial \xi} + 2 \sum_{\nu'\nu} \underline{R}_{\nu'\nu}^{\zeta,\eta} \frac{\partial \overline{P}_{\nu'\nu}^{\zeta,\eta}}{\partial \xi} + 2\mathcal{I}_{\zeta,\eta}^{\xi},
$$

(3.7)

with the unrestricted **R**-matrices and $\mathcal{I}_{\zeta,\eta}^{\xi}$ term:

$$
\mathcal{I}_{\zeta,\eta}^{\xi} = \sum_{\eta'}^{\alpha,\beta} \sum_{ij}^{\text{occ}} \sum_{ab}^{\text{virt}} \left(\underline{i}\overline{a} \mid \underline{j}\overline{b} \right)_{\eta\eta'}^{(\xi)} \left[\left(\underline{i}\overline{a} \mid \underline{j}\overline{b} \right)_{\eta\eta'} - \delta_{\eta\eta'} \left(\underline{i}\overline{b} \mid \underline{j}\overline{a} \right)_{\eta\eta} \right],
$$

$$
\underline{R}_{\nu'\nu}^{\zeta,\eta} = \sum_{\eta'}^{\alpha,\beta} \sum_{ij}^{\text{occ}} \sum_{a}^{\text{virt}} \left(\underline{i}\nu' \mid \underline{j}\overline{b} \right)_{\eta\eta'} \left[\left(\underline{i}\nu \mid \underline{j}\overline{b} \right)_{\eta\eta'} - \delta_{\eta\eta'} \left(\underline{i}\overline{b} \mid \underline{j}\nu \right)_{\eta\eta} \right],
$$

(3.8)

$$
\overline{R}_{\mu'\mu}^{\zeta,\eta} = \sum_{\eta'}^{\alpha,\beta} \sum_{i}^{\text{occ}} \sum_{ab}^{\text{virt}} \left(\mu'\overline{a} \mid \underline{j}\overline{b} \right)_{\eta\eta'} \left[\left(\mu\overline{a} \mid \underline{j}\overline{b} \right)_{\eta\eta'} - \delta_{\eta\eta'} \left(\mu\overline{b} \mid \underline{j}\overline{a} \right)_{\eta\eta} \right],
$$

where the subscript η of the two-electron integrals indicates the respective spin of the bra and the ket. In the case of the HFCCs, the $\mathcal{I}_{\zeta,\eta}^{\xi}$ term equates to zero, as the basis functions are independent of the perturbation. With an unrestricted variant of the AO-based Z-vector technique presented in Paper **I**, this allows for an efficient computation of HFCCs for larger systems than previously computable. The largest system calculated in Paper **I**, $C_{100}H_{201}$, was computed five times faster than our conservative extrapolation of the respective canonical computation.

3.2 Selected-nuclei hyperfine coupling constants

HFCCs are directly related to the spin density. In most systems, the spin density is localized in close proximity to the radical center, which results in a small number of nuclei exhibiting HFCCs deviating significantly from zero. In Paper **II**, an approach to study the HFCCs of selected nuclei is therefore introduced, further exploiting their locality.

If only a specific HFCC is of interest, directly computing the perturbed pseudodensity is computationally favorable to the Z-vector technique. This results in the following equation to obtain the HFCC of nucleus k:

$$\frac{\partial E_{\text{RI-CDD-MP2}}}{\partial M_k} = -\frac{1}{2} \sum_\zeta^\tau \omega_\zeta \sum_\eta^{\alpha,\beta} \frac{\partial \mathcal{E}^\eta(\zeta)}{\partial M_k}$$

$$\text{with:} \quad \frac{\partial \mathcal{E}^\eta(\zeta)}{\partial M_k} = \sum_{\eta'}^{\alpha,\beta} \sum_\nu \sum_{ij}^{\text{occ}} \sum_b^{\text{virt}} \left(\underline{i} \overline{\nu}^{M_k} \mid \underline{j} \overline{b} \right)_{\eta\eta'} \left[\left(\underline{i}\nu \mid \underline{j}\overline{b} \right)_{\eta\eta'} - \delta_{\eta\eta'} \left(\underline{i}\overline{b} \mid \underline{j}\nu \right)_{\eta\eta} \right]$$

$$+ \sum_{\eta'}^{\alpha,\beta} \sum_\nu \sum_i^{\text{occ}} \sum_{ab}^{\text{virt}} \left(\underline{\mu}^{M_k}\overline{a} \mid \underline{j}\overline{b} \right)_{\eta\eta'} \left[\left(\underline{\mu}\overline{a} \mid \underline{j}\overline{b} \right)_{\eta\eta'} - \delta_{\eta\eta'} \left(\underline{\mu}\overline{a} \mid \underline{j}\overline{b} \right)_{\eta\eta} \right], \quad (3.9)$$

where M_k is the nuclear magnetic moment of nucleus k and the two-electron integrals are transformed directly with the perturbed pseudodensities:

$$\left(\underline{\mu}^{M_k}\overline{a} \mid \underline{j}\overline{b} \right)_{\eta\eta'} = \sum_{\mu'} \underline{P}_{\mu\mu'}^{\eta,M_k}(\zeta) \left(\mu'\overline{a} \mid \underline{j}\overline{b} \right)_{\eta\eta'}. \quad (3.10)$$

By including the perturbation in the screening, the scaling behavior of the number of significant two-electron integrals is sublinear due to the locality of the perturbation. This is shown in the computation of the HFCC of the terminal carbon atom of linear alkane chain radicals in Paper **II**. While alkane chains do not represent realistic molecular systems, they are the ideal test case to analyze asymptotic scaling behavior. When the RI-approximation is utilized, this sublinear scaling behavior cannot be fully exploited since the coupling with perturbation is lost in the ket distribution via the auxiliary basis set. Nonetheless, a significant speed-up can be achieved, allowing the computation of the HFCC of the terminal carbon atom in $C_{200}H_{401}$.

This can be further optimized when only the opposite spin terms, i.e., $\eta' \neq \eta$ in Eq. 3.9, are computed. The idea of neglecting same spin contributions has been pioneered in the evaluation of the MP2 energy. In work by Jung *et al.*[61], the opposite spin terms are scaled with c_{os} leading to scaled opposite spin (SOS)-MP2 resulting in a reliable approximation of the full MP2 energy. In a SOS-based MP2 formulation the use of the local attenuated Coulomb-metric[61,105,106] in the RI expansion couples the locality of the bra with the ket charge distribution.

With the erfc-attenuated Coulomb-metric, the SOS-CDD-MP2 HFCCs read as:

$$
\frac{\partial E_{\text{AO-MP2}}}{\partial M_k} = -c_{\text{os}} \sum_{\zeta}^{\tau} \sum_{\eta}^{\alpha,\beta} \left(\overline{\mathbf{Z}}_{\eta}^{\omega}(\zeta) \mathbf{C}^{\omega} \mathbf{Z}_{\eta'}^{\omega}(\zeta) \mathbf{C}^{\omega} \right.
$$
$$
\left. + \underline{\mathbf{Z}}_{\eta}^{\omega}(\zeta) \mathbf{C}^{\omega} \mathbf{Z}_{\eta'}^{\omega}(\zeta) \mathbf{C}^{\omega} \right) \tag{3.11}
$$

with:

$$
\left(C^{\omega} \right)_{PQ} = \sum_{P'Q'}^{N_{\text{aux}}} \left(P \mid P' \right)_{\omega}^{-1} \left(P' \mid Q' \right) \left(Q' \mid Q \right)_{\omega}^{-1}, \tag{3.12}
$$

where

$$
\left(P \mid Q \right)_{\omega} = \left(P \left| \frac{\text{erfc}(\omega r_{12})}{r_{12}} \right| Q \right). \tag{3.13}
$$

The Z-matrices are built from sparse three-center **B**-matrices as:

$$
\left(\underline{Z}_{\eta}^{\omega}(\zeta) \right)_{PP'} = \overline{\underline{\mathcal{B}}}_{\underline{M}_k}^{P}(\omega) \mathcal{B}^{P'}(\omega)
$$
$$
\left(\overline{Z}_{\eta}^{\omega}(\zeta) \right)_{PP'} = \overline{\underline{\mathcal{B}}}_{\overline{M}_k}^{P}(\omega) \mathcal{B}^{P'}(\omega) \tag{3.14}
$$
$$
\left(Z_{\eta}^{\omega}(\zeta) \right)_{PP'} = \overline{\underline{\mathcal{B}}}^{P}(\omega) \mathcal{B}^{P'}(\omega).
$$

A definition of $\mathcal{B}$ and $\overline{\underline{\mathcal{B}}}$ can be found in Paper **II**. The attenuation of the Coulomb-metric is controlled by ω, for which a value of 0.1 in the work by Luenser *et al.*[107] demonstrated no loss in accuracy with respect to the full Coulomb-metric while still providing the computational efficiency of the local overlap metric.

A further speed-up can be obtained by migrating to massively-parallelized architectures such as graphics processing units (GPUs). Analogous to work on the computation of SOS-MP2 energies using the RI-CDD ansatz on GPUs,[108] the previously established J-engine can be used to efficiently compute the HFCC on GPUs.

In Paper **II** the accuracy of the SOS-approach is thoroughly studied and the entailed speed-ups are shown for linear alkane chains. With the new SOS-MP2 method employing the attenuated erfc-Coulomb metric, the HFCC of the terminal carbon atom of $C_{200}H_{402}$ was computed in 33 hours on one GPU node. The broad applicability of this new method was shown in an exemplary study on the influence of the environment on the HFCC of a glycyl radical enyzme, namely pyruvate formate lyase (PFL), where, despite the high globularity of the system, the protein environment could be studied extensively, even with taking into account over 400 atoms.

3.3 Influence of correlation and nuclear dynamics

Besides electron correlation, a second important influence on the accuracy of the obtained molecular property is, especially for structurally sensitive parameters, the contribution of nuclear dynamics. Most measurements are performed in solution, resulting in vibrational averaging of the measured results. Recent research has highlighted the importance of taking this into consideration when computing HFCCs.[109–112] In Manuscript **IV**, we analyzed in detail the dynamic and the correlation contribution on the HFCC.

Assuming a perfect harmonic distribution of all vibrational degrees of freedom, as well as a strictly point-symmetric relationship between the structural parameters and the HFCC, dynamic effects would be zero. These premises are typically not fulfilled, as can be seen in Manuscript **IV**. Various approaches exist to consider dynamic contributions. The most straightforward one is to perform a simulation of the nuclear motion and subsequently compute a set of frames from the simulation, averaging the obtained property accordingly. For this, the computational cost per frame needs to be sufficiently low. Using the new methods presented in Papers **I** and **II** of this work, such an analysis taking into account both the correlation effects and dynamic contributions is possible.

The results in Manuscript **IV** show that dynamic contributions are non-negligible and their neglect can in severe cases lead to erratic results. Furthermore, electron correlation proves to be equally important and considering both effects leads to accurate and reliable HFCCs.

4 Second-order properties

This section will deal with two important second-order properties, namely the electronic g-tensor and the NMR shielding. Both are of great importance in the determination and evaluation of molecular structures and their dynamical behavior. As the interpretation of the experimental spectra, especially for large molecular systems, remains challenging, *ab initio* calculations are vital to not only aid the interpretation, support the experimental evidence, but also link experiment to theory. The computation of NMR shieldings is well established, especially using HF- and DFT-ansätze.[96,113–118] However, when correlation effects need to be adressed,[45,76] MP2 was shown to be the method of choice[47,119–121]. Similar studies with respect to the electronic g-tensor are still hampered by the lack of an efficient implementation, although results using coupled cluster approaches indicate the influence of the correlation contribution[122]. Recent work by Grimme *et al.*[123] highlights the necessity to include dynamic effects in the computation of NMR shieldings and to correctly sample the conformational space, for which highly efficient codes to compute the molecular properties are required.

Sec. 4.1 summarizes the main findings in our systematic study of the gauge-dependence of the electronic g-tensor especially for large molecular systems, shown in detail in Paper **III**. The subsequent section introduces the main theory and working equations of the previous work on nuclei-selected AO-MP2 NMR shieldings[68] and discusses how the extension of the RI-CDD ansatz in the computation of analytic gradients in Paper **I** and **II** can be transferred in a new all-nuclei ansatz for NMR shieldings at the MP2-level.

4.1 Gauge-dependence of the electronic g-tensor

As outlined in Sec. 2.3.2, the computation of magnetic properties is challenged by the dependency of the result on the choice of the gauge origin when using incomplete basis sets. This is circumvented by the use of GIAOs (see Eq. 2.36). While the use of GIAOs is widespread in the computation of NMR parameters,[45,96] this is not the case with the

computation of g-tensors. The g-tensor is given as a mixed derivative of the energy with respect to the magnetic field and the electronic spin, and contains different contributions: the relativistic mass correction and diamagnetic and paramagnetic one- and two-electron terms. The relativistic mass correction is due to the kinetic energy correction to the electron spin Zeeman effect and is readily computed from kinetic energy integrals and the spin density. The one-electron contributions stem from the electron-nuclear spin-orbit coupling, whereas the two-electron terms involve both the electron-electron spin-orbit operator and the spin-other-orbit operator.[124]

Despite the g-tensor being a magnetic property, most previous studies found this property to be relatively insensitive to the choice of gauge origin,[62–67] whereas others point to a significant contribution[66,125–128]. The effect of an error due to the choice of gauge origin is expected to be greater for larger molecular systems, whose computation remains challenging due to the lack of an efficient low-scaling implementation.

We therefore conducted a study systematically analyzing the effect of the choice of the gauge origin on the accuracy of the computation of the electronic g-tensor, which is presented in the work in Paper **III**. Our results highlight that for larger molecular systems a suitable choice of the gauge origin is inevitable. If a single-choice origin is used, this must, however, be situated in the center of the spin density. Therefore, the common electronic charge centroid approach, which places the gauge origin in the center of mass of the molecular system, fails dramatically in cases where the maximum of the spin density is situated away from the center of the molecule. Paper **III** introduces a more physical and accurate choice of gauge origin, based on the spin density. For a general description of systems with delocalized spin densities or multiple spin centers, the use of GIAOs is inevitable, and thus the GIAO ansatz is the method of choice.

Furthermore, the sparsity of the perturbed spin densities in Paper **III** lays the foundation for future work on exploiting the locality of the perturbation which forms the basis for a sublinear computation of electronic g-tensors. This requires sophisticated work, especially with regards to the sublinear solution of the required CPSCF equations analogous to the work on the sublinear scaling computation of NMR shieldings at the HF/DFT-level.[129]

4.2 RI-CDD ansatz for NMR shieldings at the MP2-level

4.2.1 AO-based MP2-NMR shieldings for selected nuclei

The magnetic shieldings tensors σ_j of nucleus j can be obtained as a second derivative of the energy expression with respect to the magnetic field $\mathbf{B}$ and the nuclear magnetic spin moment $\mathbf{m}_j$:[130,131]

$$\sigma_j = \left(\frac{\partial^2 E}{\partial \mathbf{m}_j \partial \mathbf{B}} \right)_{\mathbf{B}, \mathbf{m}_j = 0}. \tag{4.1}$$

For NMR shieldings at the MP2 level, the respective MP2 energy expression is chosen. Similar to the computation of g-tensors in Sec. 4.1, gauge dependency is an important issue. Therefore, gauge-including orbitals are used,[97,113] which cancel any dependencies on the choice of the gauge origin.[45,96]

Maurer and Ochsenfeld[68] proposed a complete AO-reformulation of the computation of NMR shieldings at the MP2-level using GIAOs. Their ansatz allows for the sublinear computation of the shielding tensor of a specific nucleus. The sublinear computation of NMR shieldings has previously been demonstrated at the HF- and DFT-levels by Beer *et al.*[129] The starting point of AO-MP2 NMR shieldings is a formulation of the MP2 energy gradient with respect to the magnetic field $\mathbf{B}$, following the approach by Schweizer *et al.*[48] Subsequently, the second derivative with respect to the nuclear magnetic spin moment $\mathbf{m}_j$ is needed, taking into account the independence of the basis functions of the perturbation of $\mathbf{m}_j$.

Overall, the following equation for the computation of the shielding tensor of nucleus j can be given:[68]

$$\sigma_j^{AO-MP2} = -\sum_{\alpha=1}^{\tau} \omega_\alpha \Bigg\{ 4\underline{\mathcal{I}}^{\mathbf{B},\mathbf{m}_j} + 4\overline{\mathcal{I}}^{\mathbf{B},\mathbf{m}_j}$$

$$+ 2\operatorname{Tr}\Big[\big(\overline{\mathbf{Y}}_{\mathbf{1}}^{\mathbf{m_j}} - \underline{\mathbf{Y}}_{\mathbf{1}}^{\mathbf{m_j}} + \mathbf{G}[\overline{\mathbf{Y}}_{\mathbf{2}}^{\mathbf{m_j}} + \underline{\mathbf{Y}}_{\mathbf{2}}^{\mathbf{m_j}}] + \overline{\mathbf{R}}^{\mathbf{m_j}} e^{t_\alpha \mathbf{P}_{\mathrm{occ}} \mathbf{F}}$$

$$+ \overline{\mathbf{R}}\left(e^{t_\alpha \mathbf{P}_{\mathrm{occ}} \mathbf{F}}\right)^{\mathbf{m}_j} - \underline{\mathbf{R}}^{\mathbf{m_j}} e^{-t_\alpha \mathbf{P}_{\mathrm{virt}} \mathbf{F}} - \underline{\mathbf{R}}\left(e^{-t_\alpha \mathbf{P}_{\mathrm{virt}} \mathbf{F}}\right)^{\mathbf{m}_j} \big)\mathbf{P}_{\mathbf{occ}}^{\mathbf{B}}\Big]$$

$$+ 2\operatorname{Tr}\Big[\mathcal{P}\mathbf{P}_{\mathbf{occ}}^{\mathbf{Bm_j}}\Big]$$

$$+ 2\operatorname{Tr}\Big[\big(\overline{\mathbf{Y}}_{\mathbf{2}}^{\mathbf{m_j}} + \underline{\mathbf{Y}}_{\mathbf{2}}^{\mathbf{m_j}}\big)\mathbf{F}^{(\mathbf{B})} + \big(\overline{\mathbf{Y}}_{\mathbf{2}} + \underline{\mathbf{Y}}_{\mathbf{2}}\big)\mathbf{F}^{(\mathbf{Bm_j})}\Big]$$

$$+ 2\operatorname{Tr}\Big[\big(-(\underline{\mathbf{Y}}_{\mathbf{1}}^{\mathbf{m_j}} + \underline{\mathbf{R}}^{\mathbf{m_j}} e^{-t_\alpha \mathbf{P}_{\mathrm{virt}} \mathbf{F}} + \underline{\mathbf{R}}\left(e^{-t_\alpha \mathbf{P}_{\mathrm{virt}} \mathbf{F}}\right)^{\mathbf{m}_j}\big)\mathbf{S}^{-1}\mathbf{S}^{\mathbf{B}}\mathbf{S}^{-1}\Big]\Bigg\}. \tag{4.2}$$

The definition of the $\mathbf{Y}$-matrices, which depend on $\underline{\mathbf{R}}$ and $\overline{\mathbf{R}}$ can be found in Ref. 48. A detailed description of the recursive calculation of the derivatives of the $\mathbf{Y}$-matrices with respect to the nuclear spin moment can be found in the Appendix B of Ref. 68. The derivatives of the $\mathbf{R}$-matrices (see Eq. 3.3 for details) are defined as:[68]

$$\underline{R}_{\nu'\nu}^{\mathbf{m}_j} = \sum_{\mu\lambda\sigma}\left(\underline{\mu}^{\mathbf{m}_j}\nu' \mid \underline{\lambda}\overline{\sigma}\right)(\mu\nu \mid\mid \lambda\sigma) + \sum_{\mu\lambda\sigma}\left(\underline{\mu}\nu' \mid \underline{\lambda}^{\mathbf{m}_j}\overline{\sigma}\right)(\mu\nu \mid\mid \lambda\sigma)$$

$$+ \sum_{\mu\lambda\sigma}\left(\underline{\mu}\nu' \mid \underline{\lambda}\overline{\sigma}^{\mathbf{m}_j}\right)(\mu\nu \mid\mid \lambda\sigma), \tag{4.3}$$

and

$$\overline{R}_{\mu'\mu}^{\mathbf{m}_j} = \sum_{\nu\lambda\sigma}(\mu'\overline{\nu}^{\mathbf{m}_j} \mid \underline{\lambda}\overline{\sigma})(\mu\nu \mid\mid \lambda\sigma) + \sum_{\nu\lambda\sigma}(\mu'\overline{\nu} \mid \underline{\lambda}^{\mathbf{m}_j}\overline{\sigma})(\mu\nu \mid\mid \lambda\sigma)$$

$$+ \sum_{\nu\lambda\sigma}(\mu'\overline{\nu} \mid \underline{\lambda}\overline{\sigma}^{\mathbf{m}_j})(\mu\nu \mid\mid \lambda\sigma). \tag{4.4}$$

The perturbed two-electron matrices $\underline{\mathcal{I}}^{\mathbf{B},\mathbf{m}_j}$ and $\overline{\mathcal{I}}^{\mathbf{B},\mathbf{m}_j}$ are given by:[68]

$$\underline{\mathcal{I}}^{\mathbf{B},\mathbf{m}_j} = \sum_{\mu\nu\lambda\sigma}\left(\underline{\mu}^{\mathbf{m}_j}\overline{\nu} \mid \underline{\lambda}\overline{\sigma}\right)(\mu\nu \mid\mid \lambda\sigma)^{\mathbf{B}},$$

$$\overline{\mathcal{I}}^{\mathbf{B},\mathbf{m}_j} = \sum_{\mu\nu\lambda\sigma}\left(\underline{\mu}\overline{\nu}^{\mathbf{m}_j} \mid \underline{\lambda}\overline{\sigma}\right)(\mu\nu \mid\mid \lambda\sigma)^{\mathbf{B}}. \tag{4.5}$$

As the computation of $\underline{\mathbf{R}}^{\mathbf{m_j}}$ and $\overline{\mathbf{R}}^{\mathbf{m_j}}$ involves perturbed pseudo-densities, these need to be calculated by:

$$
\begin{aligned}
\underline{P}_{\mu'\mu}^{\mathbf{m}_j} &= \left(e^{t_\alpha \mathbf{P}_{\mathrm{occ}}\mathbf{F}}\right)^{\mathbf{m}_j} + e^{t_\alpha \mathbf{P}_{\mathrm{occ}}\mathbf{F}}\mathbf{P}_{\mathbf{occ}}^{\mathbf{m}_j} \\
\overline{P}_{\nu'\nu}^{\mathbf{m}_j} &= \left(e^{-t_\alpha \mathbf{P}_{\mathrm{virt}}\mathbf{F}}\right)^{\mathbf{m}_j} + e^{-t_\alpha \mathbf{P}_{\mathrm{virt}}\mathbf{F}}\mathbf{P}_{\mathbf{virt}}^{\mathbf{m}_j}.
\end{aligned}
\tag{4.6}
$$

The derivatives of the matrix exponentials can be obtained in recursions and the method of squaring and scaling[48,132,133] can be employed, avoiding numerical instabilities. The number of perturbations are three per atom, but involve highly local perturbed pseudo-densities.[68]

The computation of the nuclear magnetic shielding tensor is a fully AO-based reformulation, thus allowing for a (sub-)linear-scaling computation. To this end, the second derivative of the density matrix $\mathbf{P}_{\mathbf{occ}}^{\mathbf{Bm_j}}$ needs to be circumvented. This can be achieved by the Z-vector method.[48,102,103] In order to obtain the derivative of the density matrix, the following CPSCF equations need to be solved for any perturbation ξ_1:

$$
\underline{\underline{\mathbf{A}}}\mathbf{P}^{\xi_1} = \mathbf{b}^{\xi_1}.
\tag{4.7}
$$

The second derivative with respect to a perturbation ξ_2 leads to:

$$
\underline{\underline{\mathbf{A}}}\mathbf{P}^{\xi_1\xi_2} + \underline{\underline{\mathbf{A}}}^{\xi_2}\mathbf{P}^{\xi_1} = \mathbf{b}^{\xi_1\xi_2}.
\tag{4.8}
$$

As $\underline{\underline{\mathbf{A}}}$ is a symmetric, positive-definite Hessian, it can be inverted. Multiplying Eq. 4.8 from the left with both the inverse of $\underline{\underline{\mathbf{A}}}$ and $\mathcal{P}$ leads to:

$$
\begin{aligned}
\mathcal{P}\mathbf{P}^{\xi_1\xi_2} &= \mathcal{P}\underline{\underline{\mathbf{A}}}^{-1}\mathbf{b}^{\xi_1\xi_2} - \mathcal{P}\underline{\underline{\mathbf{A}}}^{-1}\underline{\underline{\mathbf{A}}}^{\xi_2}\mathbf{P}^{\xi_1} \\
&= \mathcal{P}\underline{\underline{\mathbf{A}}}^{-1}\left(\mathbf{b}^{\xi_1\xi_2} - \underline{\underline{\mathbf{A}}}^{\xi_2}\mathbf{P}^{\xi_1}\right).
\end{aligned}
\tag{4.9}
$$

Introducing the unperturbed Z-vector $\mathbf{Z}^T = \mathcal{P}\underline{\underline{\mathbf{A}}}^{-1}$ enables the critical trace $\mathrm{Tr}\left[\mathcal{P}\mathbf{P}_{\mathbf{occ}}^{\mathbf{Bm_j}}\right]$ to be expressed by:

$$
\mathrm{Tr}\left[\mathcal{P}\mathbf{P}_{\mathbf{occ}}^{\mathbf{Bm_j}}\right] = \mathrm{Tr}\left[\mathbf{Z}^T\left(\mathbf{b}^{\mathbf{Bm_j}} - \underline{\underline{\mathbf{A}}}^{\mathbf{m}_j}\mathbf{P}^{\mathbf{B}}\right)\right].
\tag{4.10}
$$

To achieve sublinearity, all terms that depend on $\mathbf{P}^{\mathbf{B}}$ need to be collected and a perturbed Z-vector needs to be introduced, which is defined by:

$$
\mathcal{O}^{\mathbf{m}_j}\mathbf{P}^{\mathbf{B}} = \left(\mathbf{Z}^T\right)^{\mathbf{m}_j}\mathbf{b}^{\mathbf{B}},
\tag{4.11}
$$

with $\left(\mathbf{Z}^T\right)^{\mathbf{m}_j} = \mathcal{O}^{\mathbf{m}_j} \underline{\underline{\mathbf{A}}}^{-1}$ (for details see Ref. 68).

Based on the pilot implementation by Maurer and Ochsenfeld[68], QQR-type integral estimates were introduced and applied, showcasing the sublinear scaling behavior for small basis sets[134]. Due to the large basis set dependency and prefactor in the scaling behavior, this could not be applied using larger and more diffuse basis sets.

4.2.2 AO-based MP2-NMR shieldings for all nuclei

While the computation of NMR shieldings of selected nuclei is often sufficient, a method to compute the shieldings of all nuclei of the system is also of interest. In principle, the approach by Maurer and Ochsenfeld allows for a linear-scaling computation of all nuclei,[68] but this entails a large prefactor. This is due to the fact that all computationally expensive steps involve matrices depending on the perturbation of the nuclear magnetic spin moment. When specific nuclei are computed, the dependence on the perturbation of the nuclear magnetic spin moment is favorable as this exploits the locality of this perturbation. This is undesirable when the NMR shieldings of all, or a large number of, nuclei are calculated, as the number of nuclear magnetic spin moment perturbations scales with the number of nuclei. This work therefore introduces a reversed order of differentiation of the AO-MP2 energy expression in order to obtain a dependence of these matrices on the perturbation of the magnetic field, which involves three non-local perturbations irrespective of the size of the molecule.

Based on the analytical AO-MP2 energy gradient by Schweizer *et al.*[48], the perturbation of the MP2 energy with respect to the nuclear magnetic spin moment $\mathbf{m}_j$ of nucleus j is given by:

$$
\begin{aligned}
E_{AO-MP2}^{\mathbf{m}_j} = -\sum_{\alpha=1}^{\tau} \omega_\alpha &\Big\{ 2\,\mathrm{Tr}\left[\left(\overline{\mathbf{Y}}_2 + \underline{\mathbf{Y}}_2\right)\mathbf{h}^{\mathbf{m}_j}\right] \\
&+ 2\,\mathrm{Tr}\left[\left(\overline{\mathbf{Y}}_1 - \underline{\mathbf{Y}}_1 + \mathbf{G}\left[\overline{\mathbf{Y}}_2 + \underline{\mathbf{Y}}_2\right] + \overline{\mathbf{R}}e^{t_\alpha \mathbf{P}_{\mathrm{occ}}\mathbf{F}} - \underline{\mathbf{R}}e^{-t_\alpha \mathbf{P}_{\mathrm{virt}}\mathbf{F}}\right)\mathbf{P}_{\mathrm{occ}}^{\mathbf{m}_j}\right] \Big\} \quad (4.12) \\
= -\sum_{\alpha=1}^{\tau} \omega_\alpha &\Big\{ 2\,\mathrm{Tr}\left[\mathcal{F}\mathbf{h}^{\mathbf{m}_j}\right] + 2\,\mathrm{Tr}\left[\mathcal{P}\mathbf{P}_{\mathrm{occ}}^{\mathbf{m}_j}\right] \Big\}.
\end{aligned}
$$

In contrast to the differentiation with respect to the magnetic field, $\mathcal{I}^{\mathbf{m}_j}$ and $\mathbf{S}_j^{\mathbf{m}}$ are zero due to the independence of the basis functions on the perturbation $\mathbf{m}_j$. The chemical

shielding tensor for a nucleus j can now be obtained by a second differentiation of Eq. 4.12 with respect to the magnetic field $\mathbf{B}$. This leads to the following expression for the shielding tensor σ_j:

$$
\begin{aligned}
\sigma_j^{AO-MP2} = -\sum_{\alpha=1}^{\tau} \omega_\alpha \Big\{ &2\,\mathrm{Tr}\,\big(\overline{\mathbf{Y}}_{\mathbf{1}}^{\mathbf{B}} - \underline{\mathbf{Y}}_{\mathbf{1}}^{\mathbf{B}} + \mathbf{G}[\overline{\mathbf{Y}}_{\mathbf{2}}^{\mathbf{B}} + \underline{\mathbf{Y}}_{\mathbf{2}}^{\mathbf{B}}] + \mathbf{G}^{\mathbf{B}}[\overline{\mathbf{Y}}_{\mathbf{2}} + \underline{\mathbf{Y}}_{\mathbf{2}}] \\
&+ \overline{\mathbf{R}}^{\mathbf{B}} e^{t_\alpha \mathbf{P}_{\mathrm{occ}} \mathbf{F}} + \overline{\mathbf{R}}\,\big(e^{t_\alpha \mathbf{P}_{\mathrm{occ}} \mathbf{F}}\big)^{\mathbf{B}} - \underline{\mathbf{R}}^{\mathbf{B}} e^{-t_\alpha \mathbf{P}_{\mathrm{virt}} \mathbf{F}} - \underline{\mathbf{R}}\,\big(e^{-t_\alpha \mathbf{P}_{\mathrm{virt}} \mathbf{F}}\big)^{\mathbf{B}}\,\big)\mathbf{P}_{\mathrm{occ}}^{m_j}\Big] \\
&+ 2\,\mathrm{Tr}\,\Big[\big(\overline{\mathbf{Y}}_{\mathbf{2}}^{\mathbf{B}} + \underline{\mathbf{Y}}_{\mathbf{2}}^{\mathbf{B}}\big)\,\mathbf{h}^{\mathbf{m_j}} + \big(\overline{\mathbf{Y}}_{\mathbf{2}} + \underline{\mathbf{Y}}_{\mathbf{2}}\big)\,\mathbf{h}^{\mathbf{m_j B}}\Big] \\
&+ 2\,\mathrm{Tr}\,\Big[\mathcal{P}\mathbf{P}_{\mathrm{occ}}^{\mathbf{m_j B}}\Big]\Big\},
\end{aligned}
\tag{4.13}
$$

where the perturbed $\mathbf{Y}^{\mathbf{B}}$-matrices depend on the respective perturbed $\mathbf{R}^{\mathbf{B}}$-matrices:

$$
\begin{aligned}
\underline{R}_{\nu'\nu}^{\mathbf{B}} = &\sum_{\mu\lambda\sigma}\big(\underline{\mu}^{\mathbf{B}}\nu' \mid \underline{\lambda}\overline{\sigma}\big)\,(\mu\nu \mid\mid \lambda\sigma) + \sum_{\mu\lambda\sigma}\big(\underline{\mu}\nu' \mid \underline{\lambda}^{\mathbf{B}}\overline{\sigma}\big)\,(\mu\nu \mid\mid \lambda\sigma) \\
&+ \sum_{\mu\lambda\sigma}\big(\underline{\mu}\nu' \mid \underline{\lambda}\overline{\sigma}^{\mathbf{B}}\big)\,(\mu\nu \mid\mid \lambda\sigma) + \sum_{\mu\lambda\sigma}\big(\underline{\mu}\nu' \mid \underline{\lambda}\overline{\sigma}\big)\,(\mu\nu \mid\mid \lambda\sigma)^{\mathbf{B}},
\end{aligned}
\tag{4.14}
$$

and

$$
\begin{aligned}
\overline{R}_{\mu'\mu}^{\mathbf{B}} = &\sum_{\nu\lambda\sigma}\big(\mu'\overline{\nu}^{\mathbf{B}} \mid \underline{\lambda}\overline{\sigma}\big)\,(\mu\nu \mid\mid \lambda\sigma) + \sum_{\nu\lambda\sigma}\big(\mu'\overline{\nu} \mid \underline{\lambda}^{\mathbf{B}}\overline{\sigma}\big)\,(\mu\nu \mid\mid \lambda\sigma) \\
&+ \sum_{\nu\lambda\sigma}\big(\mu'\overline{\nu} \mid \underline{\lambda}\overline{\sigma}^{\mathbf{B}}\big)\,(\mu\nu \mid\mid \lambda\sigma) + \sum_{\nu\lambda\sigma}\big(\mu'\overline{\nu} \mid \underline{\lambda}\overline{\sigma}\big)\,(\mu\nu \mid\mid \lambda\sigma)^{\mathbf{B}}.
\end{aligned}
\tag{4.15}
$$

The equation for the new computation of the NMR shielding tensor σ_j in Eq. 4.13 can be abbreviated by:

$$
\sigma_j^{AO-MP2} = \mathrm{Tr}[\mathcal{O}^{\mathbf{B}}\mathbf{P}_{\mathrm{occ}}^{\mathbf{m_j}}] + \mathrm{Tr}[\mathcal{P}\mathbf{P}_{\mathrm{occ}}^{\mathbf{m_j B}}] + \mathrm{Tr}[\mathcal{X}],
\tag{4.16}
$$

where quantities independent of a perturbation of the density matrix are combined in $\mathcal{X}$.

The CPSCF equation for $\mathbf{P}_{\mathrm{occ}}^{\mathbf{m_j}}$ is given by:

$$
\underline{\underline{\mathbf{A}}}\mathbf{P}_{\mathrm{occ}}^{\mathbf{m_j}} = \mathbf{b}^{\mathbf{m}_j}.
\tag{4.17}
$$

Multiplying Eq. 4.17 from the left with $\mathcal{O}^{\mathbf{B}}$ leads to:

$$
\mathcal{O}^{\mathbf{B}}\mathbf{P}_{\mathrm{occ}}^{\mathbf{m_j}} = \mathcal{O}^{\mathbf{B}}\underline{\underline{\mathbf{A}}}^{-1}\mathbf{b}^{\mathbf{m_j}}.
\tag{4.18}
$$

The perturbed Z-vector can now be defined as:

$$(\mathbf{Z}^T)^{\mathbf{B}} = \mathcal{O}^{\mathbf{B}} \underline{\underline{\mathbf{A}}}^{-1}.$$ (4.19)

Therefore, $\text{Tr}[\mathcal{O}^{\mathbf{B}} \mathbf{P}_{\mathbf{occ}}^{\mathbf{m_j}}]$ is given by:

$$\text{Tr}[\mathcal{O}^{\mathbf{B}} \mathbf{P}_{\mathbf{occ}}^{\mathbf{m_j}}] = \text{Tr}\left[\left(\mathbf{Z}^T\right)^{\mathbf{B}} \mathbf{b}^{\mathbf{m_j}}\right].$$ (4.20)

Similarly to Eq. 4.10, the trace $\text{Tr}[\mathcal{P}\mathbf{P}_{\mathbf{occ}}^{\mathbf{m_j B}}]$ can be computed as:

$$\text{Tr}\left[\mathcal{P}\mathbf{P}_{\mathbf{occ}}^{\mathbf{m_j B}}\right] = \text{Tr}\left[\mathbf{Z}^T \left(\mathbf{b}^{\mathbf{m_j B}} - \underline{\underline{\mathbf{A}}}^{\mathbf{B}} \mathbf{P}^{\mathbf{m_j}}\right)\right],$$ (4.21)

where the Z-vector $\mathbf{Z}^T$ is defined as $\mathcal{P}\underline{\underline{\mathbf{A}}}^{-1}$.

This new ansatz leads to a reduced number of perturbations in the Z-vector technique. In the approach by Maurer and Ochsenfeld, the Z-vector in Eq. 4.11 needs to be solved for each $\mathbf{m}_j$ perturbation, which is desired when specific nuclei are calculated, but which will provide a high prefactor when all nuclei are computed. In the reverse order of differentiation, the Z-vector in Eq. 4.20 needs to be solved per magnetic field perturbation, which is independent of the size of the molecule. It has to be noted, though, that the computation of the Z-vectors is not rate-determining.

All computationally expensive terms have been reformulated to asymptotically allow for a linear-scaling calculation of the three magnetic field perturbations in contrast to N_{atoms} sublinear-scaling nuclear magnetic spin moment perturbations.

4.2.3 RI-CDD ansatz in the computation of MP2-NMR shieldings

Similar to the computation of MP2 energies[59] and of analytical MP2 energy gradients (see Sec. 3.1 and Paper **I**), using the RI-CDD ansatz is beneficial to reduce the prefactor in the scaling behavior and the basis set dependency. This work therefore extends the ansatz to the computation of the perturbed **R**-matrices in the nuclei-selected approach (see Eqs. 4.3 and 4.4), resulting in:

$$\underline{R}^{\mathbf{m}_j}_{\nu'\nu} = \sum_\mu \sum_j^{\text{occ}} \sum_b^{\text{virt}} \left(\underline{\mu}^{\mathbf{m}_j}\nu' \mid \underline{j}\bar{b}\right) \left(\mu\nu \mid\mid \underline{j}\bar{b}\right) + \sum_\lambda \sum_i^{\text{occ}} \sum_b^{\text{virt}} \left(\underline{i}\nu' \mid \underline{\lambda}^{\mathbf{m}_j}\bar{b}\right) \left(\underline{i}\nu \mid\mid \lambda\bar{b}\right)$$
$$+ \sum_\sigma \sum_{i,j}^{\text{occ}} \left(\underline{i}\nu' \mid \underline{j}\bar{\sigma}^{\mathbf{m}_j}\right) \left(\underline{i}\nu \mid\mid \underline{j}\sigma\right) \tag{4.22}$$

and

$$\overline{R}^{\mathbf{m}_j}_{\mu'\mu} = \sum_\nu \sum_j^{\text{occ}} \sum_b^{\text{virt}} \left(\mu'\underline{\nu}^{\mathbf{m}_j} \mid \underline{j}\bar{b}\right) \left(\mu\nu \mid\mid \underline{j}\bar{b}\right) + \sum_\lambda \sum_{a,b}^{\text{virt}} \left(\mu'\bar{a} \mid \underline{\lambda}^{\mathbf{m}_j}\bar{b}\right) \left(\mu\bar{a} \mid\mid \lambda\bar{b}\right)$$
$$+ \sum_\sigma \sum_a^{\text{virt}} \sum_j^{\text{occ}} \left(\mu'\bar{a} \mid \underline{j}\bar{\sigma}^{\mathbf{m}_j}\right) \left(\mu\bar{a} \mid\mid \underline{j}\sigma\right). \tag{4.23}$$

Furthermore, the $\mathcal{I}^{\mathbf{B},\mathbf{m}_j}$ matrices in Eq. 4.5 can be obtained as:

$$\underline{\mathcal{I}}^{\mathbf{B},\mathbf{m}_j} = \sum_\mu \sum_j \sum_{a,b} \left(\underline{\mu}^{\mathbf{m}_j}\bar{a} \mid \underline{j}\bar{b}\right) \left(\mu\bar{a} \mid\mid \underline{j}\bar{b}\right)^{(\mathbf{B})},$$
$$\overline{\mathcal{I}}^{\mathbf{B},\mathbf{m}_j} = \sum_\nu \sum_{i,j} \sum_b \left(\underline{i}\bar{\nu}^{\mathbf{m}_j} \mid \underline{j}\bar{b}\right) \left(\underline{i}\nu \mid\mid \underline{j}\bar{b}\right)^{(\mathbf{B})}. \tag{4.24}$$

If the NMR shieldings of all nuclei following this work's new ansatz are to be computed, the RI-CDD approach in the computation of the perturbed **R**-matrices leads to:

$$\underline{R}^{\mathbf{B}}_{\nu'\nu} = \sum_\mu \sum_j^{\text{occ}} \sum_b^{\text{virt}} \left(\underline{\mu}^{\mathbf{B}}\nu' \mid \underline{j}\bar{b}\right) \left(\mu\nu \mid\mid \underline{j}\bar{b}\right) + \sum_\lambda \sum_i^{\text{occ}} \sum_b^{\text{virt}} \left(\underline{i}\nu' \mid \underline{\lambda}^{\mathbf{B}}\bar{b}\right) \left(\underline{i}\nu \mid\mid \lambda\bar{b}\right)$$
$$+ \sum_\sigma \sum_{i,j}^{\text{occ}} \left(\underline{i}\nu' \mid \underline{j}\bar{\sigma}^{\mathbf{B}}\right) \left(\underline{i}\nu \mid\mid \underline{j}\sigma\right) + \sum_{i,j}^{\text{occ}} \sum_b^{\text{virt}} \left(\underline{i}\nu' \mid \underline{j}\bar{b}\right) \left(\underline{i}\nu \mid\mid \underline{j}\bar{b}\right)^{(\mathbf{B})}, \tag{4.25}$$

and

$$\overline{R}^{\mathbf{B}}_{\mu'\mu} = \sum_\nu \sum_j^{\text{occ}} \sum_b^{\text{virt}} \left(\mu'\bar{\nu}^{\mathbf{B}} \mid \underline{j}\bar{b}\right) \left(\mu\nu \mid\mid \underline{j}\bar{b}\right) + \sum_\lambda \sum_{a,b}^{\text{virt}} \left(\mu'\bar{a} \mid \underline{\lambda}^{\mathbf{B}}\bar{b}\right) \left(\mu\bar{a} \mid\mid \lambda\bar{b}\right)$$
$$+ \sum_\sigma \sum_a^{\text{virt}} \sum_j^{\text{occ}} \left(\mu'\bar{a} \mid \underline{j}\bar{\sigma}^{\mathbf{B}}\right) \left(\mu\bar{a} \mid\mid \underline{j}\sigma\right) + \sum_{a,b}^{\text{virt}} \sum_j^{\text{occ}} \left(\mu'\bar{a} \mid \underline{j}\bar{b}\right) \left(\mu\bar{a} \mid\mid \underline{j}\bar{b}\right)^{(\mathbf{B})}. \tag{4.26}$$

The superscripted perturbation in brackets of the two-electron integrals with respect to the magnetic field indicates a perturbation of the basis functions only. When the RI-approximation is employed, as outlined in Sec. 2.2.5, the perturbation of the auxiliary space has to be disregarded to maintain gauge invariance[26,135], resulting in:

$$
\left(\mu \overline{a} \mid\mid \underline{j}\overline{b}\right)^{(\mathbf{B})} = \sum_{P}^{\mathrm{N_{aux}}} \left[B^{(\mathbf{B})}_{\mu\overline{a},P} B_{\underline{j}\overline{b},P} + B_{\mu\overline{a},P} B^{(\mathbf{B})}_{\underline{j}\overline{b},P} \right]
$$

$$
\text{with:} \quad B^{(\mathbf{B})}_{\mu\overline{a},P} = \sum_{\nu} \overline{L}_{\overline{a}\nu} B^{(\mathbf{B})}_{\mu\nu,P},
$$

$$
\text{and:} \quad B^{(\mathbf{B})}_{\mu\nu,P} = \sum_{Q}^{\mathrm{N_{aux}}} \frac{\partial \left(\mu\nu \mid Q\right)}{\partial \mathbf{B}} \left(Q \mid P\right)^{-\frac{1}{2}}.
$$

(4.27)

The perturbed pseudodensity matrices are not positive semi-definite per definition. Therefore, a Cholesky decomposition is not possible. Both $\overline{R}^{\mathbf{B}}_{\mu'\mu}$ and $\overline{R}^{\mathbf{m}_j}_{\mu'\mu}$ thus exhibit a less pronounced reduction of the scaling behavior with respect to the size of the basis set than in RI-CDD MP2 energies and energy gradients, namely from N^5 to $N^3 \cdot N^2_{\mathrm{virt}}$. Other decomposition techniques proved so far to be incapable of retaining the sparsity and locality of the perturbed pseudodensities.

With the formulation in Eqs. 4.22-4.26, the efficient code to compute $\mathbf{R}$-matrices using the RI-CDD approach, presented in Paper **I**, can be employed. If a selected nucleus is computed, accordingly, ten $\underline{\mathbf{R}}$-, ten $\overline{\mathbf{R}}$-type matrices, and the $\mathcal{I}^{\mathbf{B},\mathbf{m}_j}$ matrices need to be computed. These $\mathbf{R}$-type matrices involve the transformation with the local perturbed pseudodensities either instead of with the occupied or with the virtual decomposed pseudodensity. In the computation of all nuclei, similarly ten $\underline{\mathbf{R}}$- and $\overline{\mathbf{R}}$-type matrices are required per x-, y-, and z-direction of the magnetic field, though transformed with the less favorable non-local perturbed pseudodensity. The pseudodensity perturbed with respect to the external magnetic field is sparse, as can be seen in Fig. 4.1, where $\underline{\mathbf{P}}^{\mathbf{B}}$ and $\overline{\mathbf{P}}^{\mathbf{B}}$ are shown for for $C_{60}H_{122}$ using the basis set pcS-1.[1] In addition, the perturbed two-electron integrals need to be evaluated explicitly in Eqs. 4.25 and 4.26.

Fig. 4.2 shows the computational cost of the computation of the rate-determining perturbed $\mathbf{R}$-matrices in the new all-nuclei ansatz compared to the computation of its unperturbed variant. This clearly shows that while the computational effort is larger for a second-order molecular property, the favorable characteristics, i.e., the scaling behavior

---

[1]pcS-n are basis sets that were shown to provide accurate NMR shieldings.[47,136] pcS-1 corresponds to a double-ζ basis set in size.
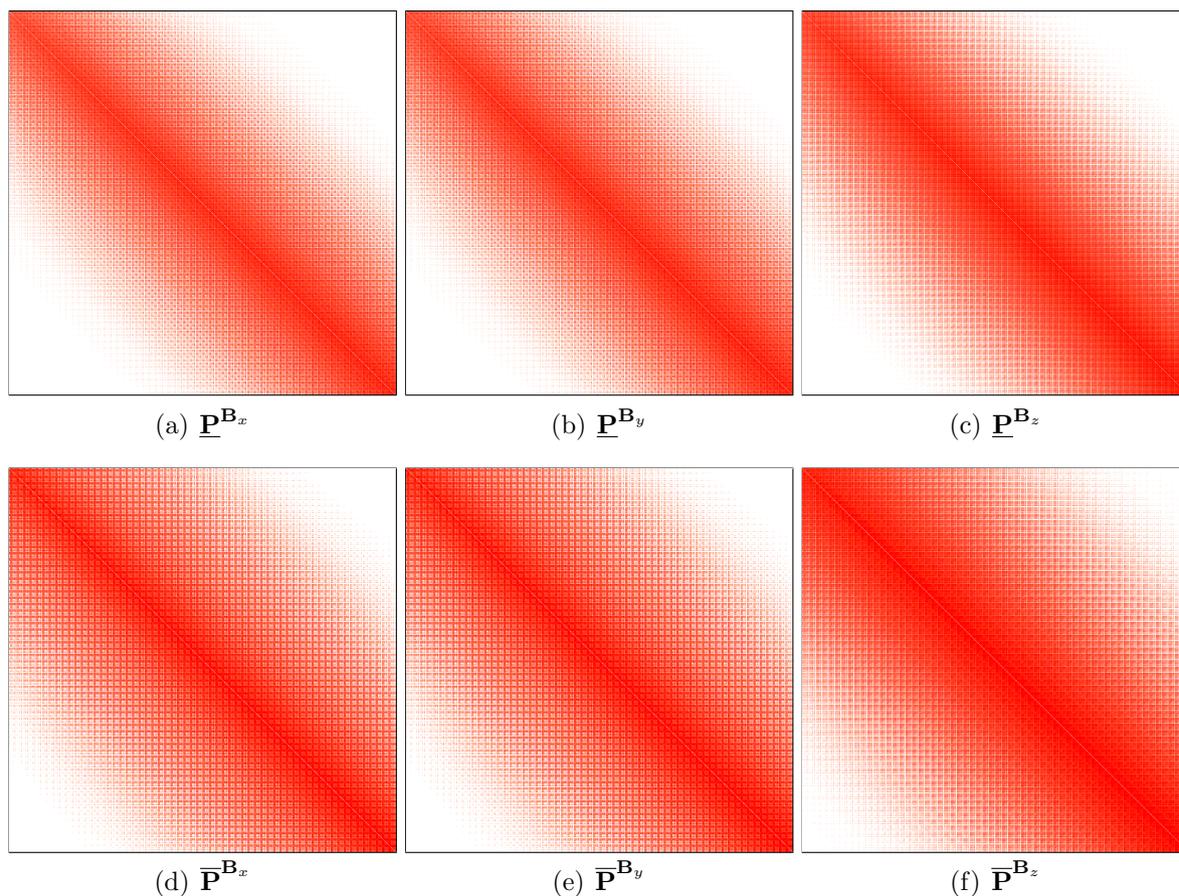
Figure 4.1: Sparsity patterns for the perturbed occupied and virtual pseudodensities with respect to the external magnetic field $\mathbf{B}$ obtained from DL-CPSCF for $C_{60}H_{122}$ using the basis set pcS-1. The sparsity patterns are taken from the first Laplace expansion point and elements smaller than 10^{-5} are discarded (white).
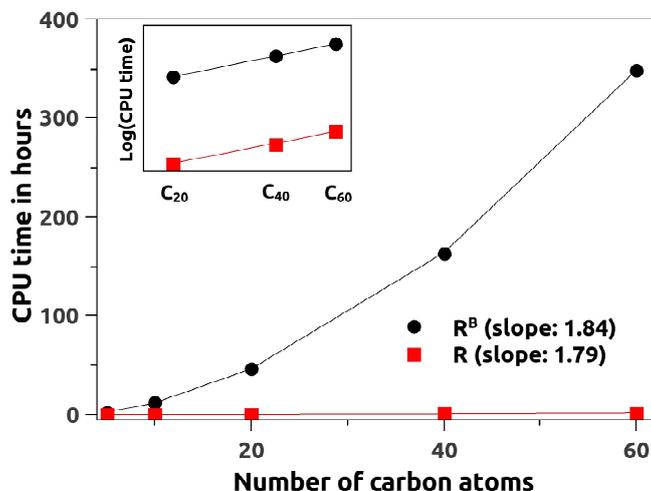
Figure 4.2: CPU time in the computation of the $\mathbf{R}$ and the $\mathbf{R^B}$ matrices of linear alkanes. The inset is a double-logarithmic plot of the CPU time of the three largest systems, on which the slope indicated in the legend is based. Basis set: pcS-1/def2-SVP-RI. Values taken from the first Laplace expansion point.

shown in Paper $\mathbf{I}$, can be exploited in the computation of the NMR shieldings. Nonetheless, the necessity to store a multitude of transformed $\mathbf{B}$-matrices, including those transformed with the perturbed pseudodensities, constitutes a memory bottleneck.

The sacrifice of the RI-CDD ansatz is the linear scaling behavior, which is lost due to the auxiliary space. While SOS-methods, in combination with local metrics such as the attenuated Coulomb-metric, can asymptotically regain this favorable scaling behavior, as, e.g., shown in work on MP2 energies[61] and on a linear-scaling AO-based reformulation of the random phase approximation[107], this is not possible for computations including both the opposite and the same spin contribution. Although work by Maurer and Ochsenfeld introduces new scaling parameters to make SOS-MP2 NMR shieldings accurate,[137] approaches including both contributions can be considered to be more stable and reliable. The new approaches introduced in this chapter already constitute a significant improvement in the computation of perturbed $\mathbf{R}$-matrices, though memory requirements and the scaling behavior remain issues for future work.

5 Conclusion and Outlook

In this work, new efficient methods to accurately compute analytic gradients, hyperfine couplings, and NMR shieldings were introduced. Based on previous studies on linear scaling algorithms to compute these molecular properties at the MP2-level, the methodology introduced in the realm of MP2 energies to reduce the basis set dependency and the prefactor in the scaling behavior was applied and extended. The reduction of the basis set dependency is achieved by decomposing the pseudodensities, thus exploiting their rank deficiency. By combining with the RI approximation, the memory requirements and the prefactor in the scaling behavior is reduced. Furthermore, to efficiently exploit the locality of the atomic orbitals, or Cholesky orbitals in the case of Cholesky decomposed pseudodensities, an extension of the previously introduced distance-including QQR-type integral estimation is required.

With these new algorithmic developments, the computations presented in Paper **I** showcase the reduction of the computational cost in the computation of the analytic gradients and an early cross-over with respect to the canonical implementation. By extending the methodology to unrestricted MP2, taking into account both spin cases in the evaluation of the MP2 contribution, HFCCs can be obtained in the absence of spin-orbit coupling. The computation of the HFCCs is highly efficient and, similarly to the analytic gradients, exhibits a cross-over to the canonical implementation for medium-sized molecules. Using the new ansatz, larger molecular systems become accessible at the MP2-level than before the present work.

The size of computable systems can be extended further by computing the HFCCs of selected nuclei. Hereby, the locality of the perturbation is exploited by a perturbation-including extended QQR-type integral estimation. In the work presented in Paper **II**, this leads to a sublinear number of significant integrals. While this significantly reduces the computational cost, the RI approximation hampers the efficient exploitation of this reduced scaling behavior. Within the SOS-approximation, this can be circumvented by the use of an attenuated Coulomb-metric, as shown in Paper **II**. Furthermore, by adapting

the previously introduced J-engine-type scheme, the computation can be transferred to massively parallelized architectures such as GPUs, leading to a further significant speed-up. Thus, studies on the influence of the environment on the HFCC of the radical center are possible, as showcased using the pyruvate formate lyase, a glycyl radical enzyme.

This efficient implementation to compute HFCCs lays the foundation for widespread applications. As shown in previous work, reliable HFCCs can only be obtained when taking into account vibrational averaging. Since the HFCCs are strongly dependent on the structure, anharmonicity in the distribution of bond angles, bond lengths, or dihedral angles, as well as deviations from the point-symmetric relationship between the structure and the HFCCs, results in considerable dynamic contributions. The study in Manuscript **IV** clearly shows that both electron correlation and incorporating dynamic effects can significantly change the computed HFCCs and must therefore be included in future studies.

When computing magnetic properties, the choice of the gauge origin can significantly alter the results. While in the computation of NMR shieldings, GIAOs, which overcome the gauge origin problem, are used, this is not commonly the case in computations of electronic g-tensors. In the benchmark presented in Paper **III**, the influence of the choice of gauge origin on the accuracy of the results is analyzed, highlighting a strong dependence for larger molecular systems. By introducing GIAOs, this error can be removed completely - though at a higher computational cost. Paper **III** therefore additionally presents a new pragmatic approach to determine a more reliable approximation of the gauge origin based on the spin density, which largely improves the results for systems with localized spin densities. For systems with delocalized spin densities or more than one radical center, GIAOs will need to be employed and constitute the basis set of choice.

This analysis of the necessity to employ GIAOs in the computation of electronic g-tensors forms the basis for future developments towards their efficient calculation, exploiting the locality of the (perturbed) spin density. In the case of NMR shieldings, the strong dependence on the choice of gauge origin is well known. Therefore, all recent developments towards (sub-)linear scaling AO-based MP2 NMR shieldings resort to GIAOs. Their widespread application is, however, hampered by the lack of an efficient implementation. Analogous to AO-MP2 energies and energy gradients, AO-MP2 NMR shieldings suffer from a strong basis set dependency with respect to the computational cost. In this work, a first attempt towards a reduced prefactor in the computational scaling behavior as well as a reduced basis set dependency of the computation of the NMR shieldings of all nuclei is presented by extending the RI-CDD ansatz. A Cholesky decomposition of the perturbed pseudo-

densities in order to exploit rank deficiency is not possible as they are skew-symmetric and indefinite. The results presented in this work show that a significant reduction in the computational cost for medium-sized molecules can be achieved. While the significant computational cost and large memory requirements remain and warrant future work on their reduction, this work already provides fundamental research results forming the basis for following studies.

The developments presented in this work form the foundation for future large scale studies of HFCCs taking into account both electron correlation and vibrational averaging. Combined with a future low scaling implementation to compute the electronic g-tensor, EPR spectra of challenging molecular systems can be obtained reliably *in silico*. Furthermore, the work on the analytic gradients can be applied to methods that involve MP2-like equations. For example, in the case of the random phase approximation, SOS-MP2-type equations are rate-determining, allowing for the results presented in Paper **I** and **II** to be transferable. Furthermore, since a recent study at the HF/DFT-level shows that a computationally efficient description of medium-sized systems is sufficient when the environment is described by molecular mechanics within the QM/MM framework,[138] the new advances regarding the computation of NMR shielding tensors of this work provide a significant step towards their reliable computation.

6 Bibliography

[1] D. R. Hartree, The Wave Mechanics of an Atom with a non-Coulomb Central Field. Part I. Theory and Methods, *Math. Proc. Cambridge Philos. Soc.* **1928**, *24*, 89.

[2] D. R. Hartree, The Wave Mechanics of an Atom with a non-Coulomb Central Field. Part II. Some Results and Discussion, *Math. Proc. Cambridge Philos. Soc.* **1928**, *24*, 111.

[3] V. Fock, Näherungsmethode zur Lösung des quantenmechanischen Mehrkörperproblems, *Z. Phys.* **1930**, *61*, 126.

[4] R. O. Ramabhadran, K. Raghavachari, Extrapolation to the Gold-Standard in Quantum Chemistry: Computationally Efficient and Accurate CCSD(T) Energies for Large Molecules Using an Automated Thermochemical Hierarchy, *J. Chem. Theory Comput.* **2013**, *9*, 3986.

[5] G. D. Purvis III, R. J. Bartlett, A full coupled-cluster singles and doubles model: The inclusion of disconnected triples, *J. Chem. Phys.* **1982**, *76*, 1910.

[6] K. Raghavachari, G. W. Trucks, J. A. Pople, M. Head-Gordon, A fifth-order perturbation comparison of electron correlation theories, *Chem. Phys. Lett.* **1989**, *157*, 479.

[7] C. Møller, M. S. Plesset, Note on an Approximation Treatment for Many-Electron Systems, *Phys. Rev.* **1934**, *46*, 618.

[8] E. Schrödinger, Quantisierung als Eigenwertproblem, *Ann. Phys.* **1926**, *385*, 437.

[9] P. Hohenberg, W. Kohn, Inhomogeneous Electron Gas, *Phys. Rev.* **1964**, *136*, B864.

[10] W. Kohn, L. J. Sham, Self-Consistent Equations Including Exchange and Correlation Effects, *Phys. Rev.* **1965**, *140*, A1133.

[11] W. Koch, M. C. Holthausen, *A Chemist's Guide to Density Functional Theory*, Wiley-VCH, New York, **2001**.

[12] R. G. Parr, W. Yang, *Density-Functional Theory of Atoms and Molecules*, Oxford University Press, New York, **1989**.

[13] N. Mardirossian, M. Head-Gordon, Thirty year of density functional theory in computational chemistry: an overview and extensive assessment of 200 density functionals, *Mol. Phys.* **2017**, *115*, 2315.

[14] S. Grimme, Semiempirical hybrid density functional with perturbative second-order correlation, *J. Chem. Phys.* **2006**, *124*, 034108.

[15] L. Goerigk, S. Grimme, Efficient and Accurate Double-Hybrid-Meta-GGA Density Functionals - Evaluation with the Extended GMTKN30 Database for General Main Group Thermochemistry, Kinetics, and Noncovalent Interactions, *J. Chem. Theory Comput.* **2011**, *7*, 291.

[16] S. Kozuch, J. M. L. Martin, Spin-component-scaled double hybrids: An extensive search for the best fifth-rung functionals blending DFT and perturbation theory, *J. Comput. Chem.* **2013**, *34*, 2327.

[17] C. Kalai, J. Toulouse, A general range-separated double-hybrid density-functional theory, *J. Chem. Phys.* **2018**, *148*, 164105.

[18] P. Pulay, Localizability of dynamic electron correlation, *Chem. Phys. Lett.* **1983**, *100*, 151.

[19] S. Saebø, P. Pulay, Local configuration interaction: An efficient approach for larger molecules, *Chem. Phys. Lett.* **1985**, *113*, 13.

[20] P. Pulay, S. Saebø, Orbital-invariant formulation and second-order gradient evaluation in Møller-Plesset perturbation theory, *Theor. Chim. Acta* **1986**, *69*, 357.

[21] M. Schütz, G. Hetzer, H.-J. Werner, Low-order scaling local electron correlation methods. I. Linear scaling local MP2, *J. Chem. Phys.* **1999**, *111*, 5691.

[22] G. Hetzer, M. Schütz, H. Stoll, H.-J. Werner, Low-order scaling local correlation methods. II. Splitting the Coulomb operator in linear scaling local second-order Møller-Plesset perturbation theory, *J. Chem. Phys.* **2000**, *113*, 9443.

[23] A. E. Azhary, G. Rauhut, P. Pulay, H.-J. Werner, Analytical energy gradients for local second-order Møller-Plesset perturbation theory, *J. Chem. Phys.* **1998**, *108*, 5185.

[24] M. Schütz, H.-J. Werner, R. Lindh, F. R. Manby, Analytical energy gradients for local second-order Møller-Plesset perturbation theory using density fitting approximations, *J. Chem. Phys.* **2004**, *121*, 737.

[25] J. Gauss, H.-J. Werner, NMR chemical shift calculations within local correlation methods: the GIAO-LMP2 approach, *Phys. Chem. Chem. Phys.* **2000**, *2*, 2083.

[26] S. Loibl, M. Schütz, NMR shielding tensors for density fitted local second-order Møller-Plesset perturbation theory using gauge including atomic orbitals, *J. Chem. Phys.* **2012**, *137*, 084107.

[27] I.-M. Høyvik, K. Kristensen, B. Jansik, P. Jørgensen, The divide-expand-consolidate family of coupled cluster methods: Numerical illustrations using second order Møller-Plesset perturbation theory, *J. Chem. Phys.* **2012**, *136*, 014105.

[28] K. Kristensen, M. Ziółkowski, B. Jansik, T. Kjærgaard, P. Jørgensen, A Locality Analysis of the Divide-Expand-Consolidate Coupled Cluster Amplitude Equations, *J. Chem. Theory Comput.* **2011**, *7*, 1677.

[29] P. Baudin, P. Ettenhuber, S. Reine, K. Kristensen, T. Kjærgaard, Efficient linear-scaling second-order Møller-Plesset perturbation theory: The divide-expand-consolidate RI-MP2 model, *J. Chem. Phys.* **2016**, *144*, 054102.

[30] K. Kristensen, T. Kjærgaard, I.-M. Høyvik, P. Ettenhuber, P. Jørgensen, B. Jansik, S. Reine, J. Jakowski, The divide-expand-consolidate MP2 scheme goes massively parallel, *Mol. Phys.* **2013**, *111*, 1196.

[31] K. Kristensen, P. Jørgensen, B. Jansik, T. Kjærgaard, S. Reine, Molecular gradient for second-order Møller-Plesset perturbation theory using the divide-expand-consolidate (DEC) scheme, *J. Chem. Phys.* **2012**, *111*, 114102.

[32] D. Bykov, K. Kristensen, T. Kjærgaard, The molecular gradient using the divide-expand-consolidate resolution of the identity second-order Møller-Plesset perturbation theory: The DEC-RI-MP2 gradients, *J. Chem. Phys.* **2016**, *145*, 024106.

[33] K. Kitaura, T. Sawai, T. Asada, T. Nakano, M. Uebayasi, Pair interaction molecular orbital method: an approximate computational method for molecular interactions, *Chem. Phys. Lett.* **1999**, *312*, 319.

[34] K. Kitaura, E. Ikeo, T. Asada, T. Nakano, M. Uebayasi, Fragment molecular orbital method: an approximate computational method for large scale molecules, *Chem. Phys. Lett.* **1999**, *313*, 701.

[35] Y. Mochizuki, K. Yamashita, T. Murase, T. Nakano, K. Fukuzawa, K. Takematsu, H. Watanabe, S. Tanaka, Large scale FMO-MP2 calculations on a massively parallel-vector computer, *Chem. Phys. Lett.* **2008**, *457*, 396.

[36] D. G. Fedorov, T. Ishida, M. Uebayasi, K. Kitaura, The Fragment Molecular Orbital Method for Geometry Optimizations of Polypeptides and Proteins, *J. Phys. Chem. A* **2007**, *111*, 2722.

[37] T. Ishikawa, N. Yamamoto, K. Kuwata, Partial energy gradient based on the fragment molecular orbital method: Application to geometry optimization, *Chem. Phys. Lett.* **2010**, *500*, 149.

[38] J. Almlöf, Elimination of energy denominators in Møller-Plesset perturbation theory by a Laplace transform approach, *Chem. Phys. Lett.* **1991**, *181*, 319.

[39] M. Häser, J. Almlöf, Laplace transform techniques in Møller-Plesset perturbation theory, *J. Chem. Phys.* **1992**, *96*, 489.

[40] M. Häser, Møller-Plesset (MP2) perturbation theory for large molecules, *Theor. Chim. Acta* **1993**, *87*, 147.

[41] P. Y. Ayala, G. E. Scuseria, Linear scaling second-order Moller-Plesset theory in the atomic orbital basis for large molecular systems, *J. Chem. Phys.* **1999**, *110*, 3660.

[42] S. A. Maurer, D. S. Lambrecht, J. Kussmann, C. Ochsenfeld, Efficient distance-including integral screening in linear-scaling Møller-Plesset perturbation theory, *J. Chem. Phys.* **2013**, *138*, 014101.

[43] H. Hellmann, *Einführung in die Quantenchemie*, Franz Deutlicke, Leipzig und Wien, **1937**.

[44] R. P. Feynman, Forces in Molecules, *Phys. Rev.* **1939**, *56*, 340.

[45] J. Gauss, *Molecular Properties*, John von Neumann Institute for Computing, Jülich, **2000**, S. 509.

[46] S. Kossmann, B. Kirchner, F. Neese, Performance of modern density functional theory for the prediction of hyperfine structure: meta-GGA and double hybrid functionals, *Mol. Phys.* **2007**, *105*, 2049.

[47] D. Flaig, M. Maurer, M. Hanni, K. Braunger, L. Kick, M. Thubauville, C. Ochsenfeld, Benchmarking Hydrogen and Carbon NMR Chemical Shifts at HF, DFT, and MP2 Levels, *J. Chem. Theory Comput.* **2014**, *10*, 572.

[48] S. Schweizer, B. Doser, C. Ochsenfeld, An atomic orbital-based reformulation of energy gradients in second-order MøllerPlesset perturbation theory, *J. Chem. Phys.* **2008**, *128*, 154101.

[49] J. L. Whitten, Coulombic potential energy integrals and approximations, *J. Chem. Phys.* **1973**, *58*, 4496.

[50] B. I. Dunlap, J. W. D. Connolly, J. R. Sabin, On some approximations in applications of Xα theory, *J. Chem. Phys.* **1979**, *71*, 3396.

[51] M. Feyereisen, G. Fitzgerald, A. Komornicki, Use of approximate integrals in ab initio theory: An application in MP2 energy calculations, *Chem. Phys. Lett.* **1993**, *208*, 359.

[52] O. Vahtras, J. Almlöf, M. Feyereisen, Integral approximations for LCAO-SCF calculations, *Chem. Phys. Lett.* **1993**, *213*, 514.

[53] D. E. Bernholdt, R. J. Harrison, Large-scale correlated electronic structure calculations: the RI-MP2 method on parallel computers, *Chem. Phys. Lett.* **1996**, *250*, 477.

[54] F. Weigend, M. Häser, RI-MP2: first derivatives and global consistency, *Theor. Chem. Acc.* **1997**, *97*, 331.

[55] F. Weigend, M. Häser, H. Patzelt, R. Ahlrichs, RI-MP2: optimized auxiliary basis sets and demonstration of efficiency, *Chem. Phys. Lett.* **1998**, *294*, 143.

[56] H. Koch, A. Sánchez de Meras, T. B. Pedersen, Reduced scaling in electronic structure calculations using Cholesky decompositions, *J. Chem. Phys.* **2003**, *118*, 9481.

[57] J. Boström, M. Pitoňák, F. Aquilante, P. Neogrády, T. B. Pedersen, R. Lindh, Coupled Cluster and Møller-Plesset Perturbation Theory Calculations of Noncovalent Intermolecular Interactions using Density Fitting with Auxiliary Basis Sets from Cholesky Decompositions, *J. Chem. Theory Comput.* **2012**, *8*, 1921.

[58] J. Zienau, L. Clin, B. Doser, C. Ochsenfeld, Cholesky-decomposed densities in Laplace-based second-order Møller-Plesset perturbation theory, *J. Chem. Phys.* **2009**, *130*, 204112.

[59] S. A. Maurer, L. Clin, C. Ochsenfeld, Cholesky-decomposed density MP2 with density fitting: Accurate MP2 and double-hybrid DFT energies for large systems, *J. Chem. Phys.* **2014**, *140*, 224112.

[60] S. A. Maurer, D. S. Lambrecht, D. Flaig, C. Ochsenfeld, Distance-dependent Schwarz-based integral estimates for two-electron integrals: Reliable tightness vs. rigorous upper bounds, *J. Chem. Phys.* **2012**, *136*, 144107.

[61] Y. Jung, R. C. Lochan, A. D. Dutoi, M. Head-Gordon, Scaled opposite-spin second order Møller-Plesset correlation energy: An economical electronic structure method, *J. Chem. Phys.* **2004**, *121*, 9793.

[62] O. L. Malkina, J. Vaara, B. Schimmelpfennig, M. Munzarova, V. G. Malkin, M. Kaupp, Density Functional Calculations of Electronic g-Tensors Using Spin-Orbit Pseudopotentials and Mean-Field All-Electron Spin-Orbit Operators, *J. Am. Chem. Soc.* **2000**, *122*, 9206.

[63] G. H. Lushington, F. Grein, Complete to second-order ab initio level calculations of electronic g-tensors, *Theor. Chim. Acta* **1996**, *93*, 259.

[64] D. Jayatilaka, Electron spin resonance g tensors from general Hartree-Fock calculations, *J. Chem. Phys.* **1998**, *108*, 7587.

[65] E. van Lenthe, P. E. S. Wormer, A. van der Avoird, Density functional calculations of molecular g-tensors in the zero-order regular approximation for relativistic effects, *J. Chem. Phys.* **1997**, *107*, 2488.

[66] G. H. Lushington, P. Bündgen, F. Grein, Ab initio study of molecular g-tensors, *Int. J. Quantum Chem.* **1995**, *55*, 377.

[67] F. Neese, Analytic derivative calculation of electronic g-tensors based on multireference configuration interaction wavefunctions, *Mol. Phys.* **2007**, *105*, 2507.

[68] M. Maurer, C. Ochsenfeld, A linear- and sublinear-scaling method for calculating NMR shieldings in atomic orbital-based second-order Møller-Plesset perturbation theory, *J. Chem. Phys.* **2013**, *138*, 174104.

[69] M. Born, R. Oppenheimer, Zur Quantentheorie der Molekeln, *Ann. Phys.* **1927**, *389*, 457.

[70] J. A. Pople, R. K. Nesbet, Self-Consistent Orbitals for Radicals, *J. Chem. Phys.* **1954**, *22*, 571.

[71] E. Engel, R. M. Dreizler, *Density Functional Theory*, Springer-Verlag Berlin Heidelberg, **2011**.

[72] C. C. J. Roothaan, New Developments in Molecular Orbital Theory, *Rev. Mod. Phys.* **1951**, *23*, 69.

[73] G. G. Hall, The Molecular Orbital Theory of Chemical Valency VIII. A Method of Calculating Ionization Potentials, *Proc. R. Soc. A Math. Phys. Eng. Sci.* **1951**, *205*, 541.

[74] P. Jurečka, J. Šponer, J. Černý, P. Hobza, Benchmark database of accurate (MP2 and CCSD(T) complete basis set limit) interaction energies of small model complexes, DNA base pairs, and amino acid pairs, *Phys. Chem. Chem. Phys.* **2006**, *8*, 1985.

[75] A. D. Boese, Density Functional Theory and Hydrogen Bonds: Are We There Yet?, *Chem. Phys. Chem.* **2015**, *16*, 978.

[76] J. Gauss, Effects of electron correlation in the calculation of nuclear magnetic resonance chemical shifts, *J. Chem. Phys.* **1993**, *99*, 3629.

[77] M. Bühl, T. van Mourik, NMR spectroscopy: quantum-chemical calculations, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1*, 634.

[78] O. Christiansen, J. Olsen, P. Jørgensen, H. Koch, P.-Å. Malmqvist, On the inherent divergence in the Møller-Plesset series. The neon atom - a test case, *Chem. Phys. Lett.* **1996**, *261*, 369.

[79] J. Olsen, O. Christiansen, H. Koch, P. Jørgensen, Surprising cases of divergent behavior in Møller-Plesset perturbation theory, *J. Chem. Phys.* **1996**, *105*, 5082.

[80] J. Olsen, P. Jørgensen, T. Helgaker, O. Christiansen, Divergence in Møller-Plesset theory: A simple explanation based on a two-state model, *J. Chem. Phys.* **2000**, *112*, 9736.

[81] J. Olsen, M. P. Fülscher, On the convergence of multi-reference perturbation theory, *Chem. Phys. Lett.* **2000**, *326*, 225.

[82] M. L. Leininger, W. D. Allen, H. F. Schaefer, C. D. Sherrill, Is Møller-Plesset perturbation theory a convergent ab initio method?, *J. Chem. Phys.* **2000**, *112*, 9213.

[83] N. C. Handy, P. J. Knowles, K. Somasundram, On the convergence of the Møller-Plesset perturbation series, *Theor. Chim. Acta* **1985**, *68*, 87.

[84] A. S. Menon, L. Radom, Consequences of Spin Contamination in Unrestricted Calculations on Open-Shell Species: Effect of Hartree-Fock and Møller-Plesset Contributions in Hybrid and Double-Hybrid Density Functional Theory Approaches, *J. Phys. Chem. A* **2008**, *112*, 13225.

[85] S. Saebø, P. Pulay, Fourth-order Møller-Plesset perturbation theory in the local correlation treatment. I. Method., *J. Chem. Phys.* **1987**, *86*, 914.

[86] M. Häser, R. Ahlrichs, Improvements on the direct SCF Method, *J. Comput. Chem.* **1989**, *10*, 104.

[87] D. S. Lambrecht, C. Ochsenfeld, Multipole-based integral estimates for the rigorous description of distance dependence in two-electron integrals, *J. Chem. Phys.* **2005**, *123*, 184101.

[88] D. S. Lambrecht, B. Doser, C. Ochsenfeld, Rigorous integral screening for electron correlation methods, *J. Chem. Phys.* **2005**, *123*, 184102.

[89] T. H. Thompson, C. Ochsenfeld, Distance-including rigorous upper bounds and tight estimates for two-electron integrals over long- and short-range operators, *J. Chem. Phys.* **2017**, *147*, 144101.

[90] D. Graf, M. Beuerle, H. F. Schurkus, A. Luenser, G. Savasci, C. Ochsenfeld, Accurate and Efficient Parallel Implementation of an Effective Linear-Scaling Direct Random Phase Approximation Method, *J. Chem. Theory Comput.* **2018**, *14*, 2505.

[91] M. Beuerle, D. Graf, H. F. Schurkus, C. Ochsenfeld, Efficient calculation of beyond RPA correlation energies in the dielectric matrix formalism, *J. Chem. Phys.* **2018**, *148*, 204104.

[92] H. F. Schurkus, A. Luenser, C. Ochsenfeld, Communication: Almost error-free resolution-of-the-identity correlation methods by null space removal of the particle-hole interactions, *J. Chem. Phys.* **2017**, *146*, 211106.

[93] D. P. Craig, T. Thirunamachandran, *Molecular Quantum Electrodynamics*, Dover Publications, New York, **1984**.

[94] W. Kutzelnigg, U. Fleischer, M. Schindler, *NMR-Basic Principles and Progress: Deuterium and Shift Calculation*, Springer-Verlag Berlin Heidelberg, **1991**.

[95] T. Helgaker, S. Coriani, P. Jørgensen, K. Kristensen, J. Olsen, K. Ruud, Recent Advances in Wave Function-Based Methods of Molecular-Property Calculations, *Chem. Rev.* **2012**, *112*, 543.

[96] K. Wolinski, J. F. Hinton, P. Pulay, Efficient implementation of the gauge-independent atomic orbital method for NMR chemical shift calculations, *J. Am. Chem. Soc.* **1990**, *112*, 8251.

[97] F. London, Théorie quantique des courants interatomiques dans les combinaisons aromatiques, *J. Phys. Radium* **1937**, *8*, 397.

[98] H. F. Hameka, On the nuclear magnetic shielding in the hydrogen molecule, *Mol. Phys.* **1958**, *1*, 203.

[99] H. F. Hameka, Berechnung der magnetischen Eigenschaften des Wasserstoffmoleküls, *Z. Naturforsch. A* **1959**, *14*, 599.

[100] H. F. Hameka, Theory of Magnetic Properties of Molecules with Particular Emphasis on the Hydrogen Molecule, *Rev. Mod. Phys.* **1962**, *34*, 87.

[101] R. Ditchfield, Molecular Orbital Theory of Magnetic Shielding and Magnetic Susceptibility, *J. Chem. Phys.* **1972**, *56*, 5688.

[102] N. C. Handy, H. F. Schaefer III, On the evaluation of analytic energy derivatives for correlated wave functions, *J. Chem. Phys.* **1984**, *81*, 5031.

[103] Y. Yamaguchi, J. D. Goddard, Y. Osamura, H. F. Schaefer III, *New Dimension to Quantum Chemistry: Analytic Derivative Methods in Ab Initio Molecular Electronic Structure Theory*, Oxford University Press Inc., Oxford, **1994**.

[104] M. Beer, C. Ochsenfeld, Efficient linear-scaling calculation of response properties: Density matrix-based Laplace-transformed coupled-perturbed self-consistent field theory, *J. Chem. Phys.* **2008**, *128*, 221102.

[105] Y. Jung, A. Sodt, P. M. W. Gill, M. Head-Gordon, Auxiliary basis expansions for large-scale electronic structure calculations, *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 6692.

[106] S. Reine, E. Tellgren, A. Krapp, T. Kjærgaard, T. Helgaker, B. Jansik, S. Høst, P. Salek, Variational and robust density fitting of four-center two-electron integrals in local metrics, *J. Chem. Phys.* **2008**, *129*, 104101.

[107] A. Luenser, H. F. Schurkus, C. Ochsenfeld, Vanishing-Overhead Linear-Scaling Random Phase Approximation by Cholesky Decomposition and an Attenuated Coulomb-Metric, *J. Chem. Theory Comput.* **2017**, *13*, 1647.

[108] S. A. Maurer, J. Kussmann, C. Ochsenfeld, Communication: A reduced scaling J-engine based reformulation of SOS-MP2 using graphics processing units, *J. Chem. Phys.* **2014**, *141*, 051106.

[109] H. Tachikawa, M. Igarashi, T. Ishibashi, Ab initio molecular dynamics (MD) calculations of hyperfine coupling constants of methyl radical, *Chem. Phys. Lett.* **2002**, *352*, 113.

[110] J. R. Asher, M. Kaupp, Hyperfine Coupling Tensors of the Benzosemiquinone Radical Anion from Car-Parrinello Molecular Dynamics, *Chem. Phys. Chem.* **2007**, *8*, 69.

[111] V. Barone, P. Cimino, A. Pedone, An integrated computational protocol for the accurate prediction of EPR and PNMR parameters of aminoxyl radicals in solution, *Magn. Reson. Chem.* **2010**, *48*, S11.

[112] A. Massolle, T. Dresselhaus, S. Eusterwiemann, C. Doerenkamp, H. Eckert, A. Studer, J. Neugebauer, Towards reliable references for electron paramagnetic resonance parameters based on quantum chemistry: the case of verdazyl radicals, *Phys. Chem. Chem. Phys.* **2018**, *20*, 7661.

[113] R. Ditchfield, Self-consistent perturbation theory of diamagnetism I. A gauge-invariant LCAO method for N.M.R. chemical shifts, *Mol. Phys.* **1974**, *27*, 789.

[114] M. Häser, R. Ahlrichs, H. Baron, P. Weis, H. Horn, Direct computation of second-order SCF properties of large molecules on workstation computers with an application to large carbon clusters, *Theor. Chim. Acta* **1992**, *83*, 455.

[115] R. Stevens, R. Pitzer, W. Lipscomb, Perturbed Hartree-Fock Calculations. I. Magnetic Susceptibility and Shielding in the LiH Molecule, *J. Chem. Phys.* **1963**, *38*, 550.

[116] G. Schreckenbach, T. Ziegler, Calculation of NMR Shielding Tensors Using Gauge-Including Atomic Orbitals and Modern Density Functional Theory, *J. Phys. Chem.* **1995**, *99*, 606.

[117] J. R. Cheeseman, G. W. Trucks, T. A. Keith, M. J. Frisch, A comparison of models for calculating nuclear magnetic resonance shielding tensors, *J. Chem. Phys.* **1996**, *104*, 5497.

[118] G. Rauhut, S. Puyear, K. Wolinski, P. Pulay, Comparison of NMR Shieldings Calculated from Hartree-Fock and Density Functional Wave Functions Using Gauge-Including Atomic Orbitals, *J. Phys. Chem.* **1996**, *100*, 6310.

[119] A. A. Auer, J. Gauss, J. F. Stanton, Quantitative prediction of gas-phase ^{13}C nuclear magnetic shielding constants, *J. Chem. Phys.* **2003**, *118*, 10407.

[120] L. Olsson, D. Cremer, Sum-over states density functional perturbation theory: Prediction of reliable ^{13}C, ^{15}N, and ^{17}O nuclear magnetic resonance chemical shifts, *J. Chem. Phys.* **1996**, *105*, 8995.

[121] S. Moon, D. Case, A Comparison of Quantum Chemical Methods for Calculating NMR Shielding Parameters in Peptides: Mixed Basis Set and ONIOM Methods Combined with a Complete Basis Set Extrapolation, *J. Comput. Chem.* **2006**, *27*, 825.

[122] A. Perera, J. Gauss, P. Verma, J. A. Morales, Benchmark coupled-cluster g-tensor calculations with full inclusion of the two-particle spin-orbit contributions, *J. Chem. Phys.* **2017**, *146*, 164104.

[123] S. Grimme, C. Bannwarth, S. Dohm, A. Hansen, J. Pisarek, P. Pracht, J. Seibert, F. Neese, Fully Automated Quantum-Chemistry-Based Computation of Spin-Spin-Coupled Nuclear Magnetic Resonance Spectra, *Angew. Chem. Int. Ed.* **2017**, *56*, 14763.

[124] G. Schreckenbach, T. Ziegler, Calculation of the G-Tensor of Electron Paramagnetic Resonance Spectroscopy Using Gauge-Including Atomic Orbitals and Density Functional Theory, *J. Phys. Chem. A* **1997**, *101*, 3388.

[125] A. J. Sadlej, Medium-size polarized basis sets for high-level correlated calculations of molecular electric properties, *Collect. Czech. Chem. Commun.* **1988**, *53*, 1995.

[126] M. Kaupp, T. Gress, R. Reviakine, O. L. Malkina, V. G. Malkin, g Tensor and Spin Density of the Modified Tyrosyl Radical in Galactose Oxidase: A Density Functional Study, *J. Phys. Chem. B* **2003**, *107*, 331.

[127] S. Patchkovskii, R. T. Strong, C. J. Pickard, S. Un, Gauge invariance of the spin-other-orbit contribution to the g-tensors of electron paramagnetic resonance, *J. Chem. Phys.* **2005**, *122*, 214101.

[128] G. H. Lushington, F. Grein, The electronic g-tensor of MgF: A comparison of ROHF and MRD-CI level results, *Int. J. Quantum Chem.* **1996**, *60*, 1679.

[129] M. Beer, J. Kussmann, C. Ochsenfeld, Nuclei-selected NMR shielding calculations: A sublinear-scaling quantum-chemical method, *J. Chem. Phys.* **2011**, *134*, 074102.

[130] J. Gauss, Calculation of NMR chemical shifts at second-order many-body perturbation theory using gauge-including atomic orbitals, *Chem. Phys. Lett.* **1992**, *191*, 614.

[131] J. Gauss, J. F. Stanton, *Electron-Correlated Approaches for the Calculation of NMR Chemical Shifts*, Wiley-Blackwell, Kap. 6, **2003**, S. 355.

[132] C. Moler, C. Van Loan, Nineteen Dubious Ways to Compute the Exponential of a Matrix, *SIAM Rev.* **1978**, *20*, 801.

[133] T. Fung, Computation of the matrix exponential and its derivatives by scaling and squaring, *Int. J. Numer. Methods Eng.* **2004**, *59*, 1273.

[134] M. Maurer, Dissertation, Ochsenfeld group, University of Munich (LMU), **2014**.

[135] S. Loibl, F. Manby, M. Schütz, Density fitted, local Hartree-Fock treatment of NMR chemical shifts using London atomic orbitals, *Mol. Phys.* **2010**, *108*, 477.

[136] F. Jensen, Basis Set Convergence for Nuclear Magnetic Shielding Constants Calculated by Density Functional Methods, *J. Chem. Theory Comput.* **2008**, *4*, 719.

[137] M. Maurer, C. Ochsenfeld, Spin Component-Scaled second-Order Møller-Plesset Perturbation Theory for Calculating NMR Shieldings, *J. Chem. Theory Comput.* **2015**, *11*, 37.

[138] D. Flaig, M. Beer, C. Ochsenfeld, Convergence of Electronic Structure with the Size of the QM region: Example of QM/MM NMR Shieldings, *J. Chem. Theory Comput.* **2012**, *8*, 2260.

7 Publications

This chapter contains three paper published in peer-reviewed journals and a manuscript as part of this thesis. The supporting information to Paper **II** can be found attached to Paper **II**.

7.1 Paper I: Analytical RI-CDD MP2 energy gradients

> "Low-scaling first-order properties within second-order Møller-Plesset perturbation theory using Cholesky decomposed density matrices",
> **S. Vogler**, M. Ludwig, M. Maurer, C. Ochsenfeld,
> *J. Chem. Phys*, **147**, 024101 (2017)

The following article is reproduced in agreement with its publisher (AIP Publishing LLC) and can be found online at:

http://aip.scitation.org/doi/pdf/10.1063/1.4990413

Low-scaling first-order properties within second-order Møller-Plesset perturbation theory using Cholesky decomposed density matrices

Sigurd Vogler, Martin Ludwig, Marina Maurer, and Christian Ochsenfeld[a)]
Chair of Theoretical Chemistry and Center for Integrated Protein Science Munich (CIPSM),
Department of Chemistry, University of Munich (LMU), Butenandtstr. 7, 81377 Munich, Germany

An efficient implementation of energy gradients and of hyperfine coupling constants in second-order Møller-Plesset perturbation theory (MP2) is presented based on our fully atomic orbital (AO)-based approach. For the latter, an unrestricted AO-based MP2 formulation is introduced. A reduction in the dependency of the computational efficiency on the size of the basis set is achieved by a Cholesky decomposition and the prefactor is reduced by the resolution-of-the-identity approximation. Significant integral contributions are selected based on distance-including integral estimates (denoted as QQR-screening) and its reliability as a fully controlled screening procedure is demonstrated. The rate-determining steps are shown via model computations to scale cubically in the computation of energy gradients and quadratically in the case of hyperfine coupling constants. Furthermore, a significant speed-up of the computational time with respect to the canonical formulation is demonstrated. *Published by AIP Publishing.* [http://dx.doi.org/10.1063/1.4990413]

I. INTRODUCTION

Analytic energy gradients, as first introduced in a quantum-chemical context by Pulay,[1] play a crucial role in theoretical chemistry. Besides providing equilibrium and transition structures of a molecule, they are necessary to perform on-the-fly molecular dynamics and form the basis for calculating molecular properties, such as hyperfine coupling constants (HFCCs) and nuclear magnetic resonance (NMR) shifts.[2] For accurate gradients, electron correlation needs to be accounted for. This can be achieved by various means, e.g., by coupled-cluster (CC)[3–6] or second-order Møller-Plesset perturbation theory (MP2).[7,8] Such electron correlation methods entail a large computational effort and a high-order scaling behavior of the cost. Even the relatively cheap MP2 method scales in its canonical formulation with the fifth power of the size of the molecule. The computation of large molecular systems at low computational cost therefore needs a reduction of the scaling behavior. Several approaches exist to exploit the locality of dynamic correlation. A group of methods is based on the work by Pulay and Saebø[9–12] using localized orbitals. Various local correlation methods were derived,[13,14] including local MP2 energies[15,16] and analytical gradients.[17,18] Another approach is to fragment the whole molecule, such as in the fragment molecular orbital (FMO) method by Kitaura *et al.*[19,20] This enables large scale energy computations at the MP2-level on massively parallel-vector computers[21] and has been extended for geometry optimizations.[22–24] An alternative approach is the divide-expand-consolidate (DEC) ansatz in partitioning the orbital space,[25–27] thus allowing for massive parallelization[28] and for the computation of large-scale systems. Recently, DEC-MP2 gradients were presented, providing a fast

computation using large computer clusters.[29,30] Besides these groups of methods, an alternative ansatz exploits the locality of the electron correlation by a formulation of the MP2 equations in atomic orbitals (AOs) via a Laplace transformation.[31–33] AO methods do not rely on a spatial restriction of the correlation space. Instead significant integral contributions are determined directly.[34–36]

The present work is based on the Laplace transform. A linear scaling behavior can hereby be achieved by our efficient QQR-type integral estimates that include the distance dependency of the charge distributions in the two-electron integrals.[35,36] Due to redundancies in the AO basis set, this scaling behavior cannot be expected when the basis set is enlarged at a constant size of the molecular system. Using a Cholesky decomposition,[37–39] Maurer *et al.*[40] introduced an approach to reduce the redundancy of the basis set in the computation of AO-MP2 energies whilst maintaining its locality. Furthermore, the prefactor in the scaling behavior can be reduced by the resolution-of-the-identity (RI) approximation.[41–47] Combining the Cholesky decomposition of the pseudodensity matrices (CDD) with the RI-approximation to form the RI-CDD ansatz and using the QQR-estimation technique, a low-scaling, efficient computation of the MP2 energy is achieved.[40]

We introduced the theory for the computation of MP2 energy gradients[48] avoiding any transformations between the AO and the molecular orbital (MO) basis. While the working equations and first results of a preliminary implementation have been presented earlier,[48] an efficient implementation has not yet been put forward. This is achieved with the present work, where we apply the RI-CDD approach to the computation of AO-MP2 energy gradients and present an efficient implementation with extended QQR-type integral estimates.

Furthermore, based on an open-shell formulation of the RI-CDD MP2 energy gradients, the efficient computation of

a)Electronic mail: christian.ochsenfeld@uni-muenchen.de

HFCCs in the absence of spin-orbit coupling becomes possible. The equations to compute HFCCs at the MP2-level using the RI-CDD approach are presented, accounting for both spin cases by an extended Laplace expansion. We want to point out at this point two approaches for improving the reliability and accuracy of the computed HFCCs. The first one is to use spin-component scaling, as introduced by Grimme,[49] which, in combination with orbital optimization, provides reliable HFCCs.[50] The second one is the double-hybrid density functional theory (DH-DFT).[51] DH-DFT combines density functional theory and perturbation theory thus overcoming the pitfalls and short-comings of the MP2 approach. It was shown for bond dissociation energies that including MP2 contributions improves the results while not being as sensitive with respect to spin contamination as a pure MP2 computation.[52] With the use of analytical gradients of the DH-DFT energy,[53] reliable HFCCs can be computed.[54] Both approaches can also be used within our RI-CDD ansatz by adapting the respective MP2 terms.

In this work, we first briefly summarize the main equations for the calculation of AO-MP2 gradients and present the algorithm to compute unrestricted MP2 energy gradients by the RI-CDD ansatz. We then introduce the extension of the QQR-type integral estimates. Subsequently, we derive the equations to compute closed-shell nuclear gradients and HFCCs based on the RI-CDD approach. Finally, we analyze both the accuracy and the efficiency of our approach.

II. THEORY

A. AO-MP2 energy gradients

Following the Laplace-transform approach by Almlöf and Häser,[31–33] we introduce the unrestricted AO-MP2 energy as

$$
\begin{aligned}
E_{\text{AO-MP2}} = &-\frac{1}{2} \sum_{\zeta_{\alpha\alpha}}^{\tau_{\alpha\alpha}} w_{\zeta_{\alpha\alpha}}^{\alpha\alpha} \sum_{\mu\nu\lambda\sigma}^{N} \left(\underline{\mu}\overline{\nu} \mid \underline{\lambda}\overline{\sigma} \right)_{\alpha\alpha} (\mu\nu \parallel \lambda\sigma) \\
&- \frac{1}{2} \sum_{\zeta_{\beta\beta}}^{\tau_{\beta\beta}} w_{\zeta_{\beta\beta}}^{\beta\beta} \sum_{\mu\nu\lambda\sigma}^{N} \left(\underline{\mu}\overline{\nu} \mid \underline{\lambda}\overline{\sigma} \right)_{\beta\beta} (\mu\nu \parallel \lambda\sigma) \\
&- \sum_{\zeta_{\alpha\beta}}^{\tau_{\alpha\beta}} w_{\zeta_{\alpha\beta}}^{\alpha\beta} \sum_{\mu\nu\lambda\sigma}^{N} \left(\underline{\mu}\overline{\nu} \mid \underline{\lambda}\overline{\sigma} \right)_{\alpha\beta} (\mu\nu \mid \lambda\sigma), \quad (1)
\end{aligned}
$$

where integration over the spin coordinate is applied. The two-electron integrals are transformed with the local pseudo-densities (PDs),

$$
\left(\underline{\mu}\overline{\nu} \mid \underline{\lambda}\overline{\sigma} \right)_{\eta\eta'} = \sum_{\mu'\nu'\lambda'\sigma'}^{N} \underline{P}_{\mu'\mu}^{\eta} \overline{P}_{\nu'\nu}^{\eta} \underline{P}_{\lambda'\lambda}^{\eta'} \overline{P}_{\sigma'\sigma}^{\eta'} (\mu'\nu' \mid \lambda'\sigma'). \quad (2)
$$

Different Laplace expansions are chosen for the three spin cases, indicated by the linear Laplace expansion coefficients $w^{\eta\eta'}$, which depend on the Laplace point $\zeta_{\eta\eta'}$. The number of Laplace points is given by $\tau_{\eta\eta'}$. The Laplace expansions of the different spin cases depend on a generalized range, which are given by

$$
R_{\eta\eta'} = \frac{(\epsilon_{\max}^{\eta} + \epsilon_{\max}^{\eta'}) - (\epsilon_{\min}^{\eta} + \epsilon_{\min}^{\eta'})}{(\epsilon_{\text{LUMO}}^{\eta} + \epsilon_{\text{LUMO}}^{\eta'}) - (\epsilon_{\text{HOMO}}^{\eta} + \epsilon_{\text{HOMO}}^{\eta'})}. \quad (3)
$$

In the closed shell case, only one range is necessary ($\eta = \eta'$). In a similar manner, we propose a simplification for the open-shell case by taking the maximum range of the three different spin cases leading to the Laplace points ζ. In the unrestricted formulation, this leads to the following equation for the AO-MP2 energy:

$$
\begin{aligned}
E_{\text{AO-MP2}} = &-\frac{1}{2} \sum_{\zeta}^{\tau} w_{\zeta} \sum_{\mu\nu\lambda\sigma}^{N} \Big[\left(\underline{\mu}\overline{\nu} \mid \underline{\lambda}\overline{\sigma} \right)_{\alpha\alpha} (\mu\nu \parallel \lambda\sigma) \\
&+ \left(\underline{\mu}\overline{\nu} \mid \underline{\lambda}\overline{\sigma} \right)_{\beta\beta} (\mu\nu \parallel \lambda\sigma) \\
&+ 2\left(\underline{\mu}\overline{\nu} \mid \underline{\lambda}\overline{\sigma} \right)_{\alpha\beta} (\mu\nu \mid \lambda\sigma) \Big], \quad (4)
\end{aligned}
$$

which can be abbreviated as

$$
E_{\text{AO-MP2}} = -\frac{1}{2} \sum_{\zeta}^{\tau} w_{\zeta} \sum_{\eta}^{\alpha,\beta} \mathcal{E}^{\eta}(\zeta), \quad (5)
$$

with

$$
\begin{aligned}
\mathcal{E}^{\eta}(\zeta) = &\sum_{\mu\nu\lambda\sigma}^{N} \left(\underline{\mu}\overline{\nu} \mid \underline{\lambda}\overline{\sigma} \right)_{\eta\eta} (\mu\nu \parallel \lambda\sigma) \\
&+ \left(\underline{\mu}\overline{\nu} \mid \underline{\lambda}\overline{\sigma} \right)_{\eta\eta'} (\mu\nu \mid \lambda\sigma), \quad (6)
\end{aligned}
$$

which depends on the Laplace expansion via the PDs and where η' represents the opposite electron spin state of η. To obtain the energy gradients, this energy expression is differentiated with respect to the perturbation ξ,

$$
\begin{aligned}
\frac{\partial E_{\text{AO-MP2}}}{\partial \xi} = &-\frac{1}{2} \sum_{\zeta}^{\tau} w_{\zeta} \sum_{\eta}^{\alpha,\beta} \frac{\partial \mathcal{E}^{\eta}(\zeta)}{\partial \xi} \\
= &-\frac{1}{2} \sum_{\zeta}^{\tau} w_{\zeta} \sum_{\eta}^{\alpha,\beta} \Big[2 \sum_{\mu'\mu}^{N} \overline{R}_{\mu'\mu}^{\eta}(\zeta) \frac{\partial \underline{P}_{\mu'\mu}^{\eta}}{\partial \xi} \\
&+ 2 \sum_{\nu'\nu}^{N} \underline{R}_{\nu'\nu}^{\eta}(\zeta) \frac{\partial \overline{P}_{\nu'\nu}^{\eta}}{\partial \xi} + 2 I^{\eta,\xi}(\zeta) \Big]. \quad (7)
\end{aligned}
$$

Hereby, the terms depending on the derivative of the two-electron integral are sorted in opposite-spin (OS) and same-spin (SS) $I^{\eta,\xi}$ contributions, given by

$$
\begin{aligned}
I^{\eta,\xi}(\zeta) = &I_{\text{SS}}^{\eta,\xi}(\zeta) + I_{\text{OS}}^{\eta,\xi}(\zeta) \\
= &\sum_{\mu\nu\lambda\sigma}^{N} \left(\underline{\mu}\overline{\nu} \mid \underline{\lambda}\overline{\sigma} \right)_{\eta\eta} \frac{\partial (\mu\nu \parallel \lambda\sigma)}{\partial \xi} \\
&+ \sum_{\mu\nu\lambda\sigma}^{N} \left(\underline{\mu}\overline{\nu} \mid \underline{\lambda}\overline{\sigma} \right)_{\eta\eta'} \frac{\partial (\mu\nu \mid \lambda\sigma)}{\partial \xi}, \quad (8)
\end{aligned}
$$

and the terms contracted with the derivative of the PDs are included in so-called **R**-matrices,

$$
\begin{aligned}
\overline{R}_{\mu'\mu}^{\eta}(\zeta) = &\overline{R}_{\mu'\mu}^{\eta,SS}(\zeta) + \overline{R}_{\mu'\mu}^{\eta,OS}(\zeta) \\
= &\sum_{\nu\lambda\sigma}^{N} \left(\mu'\overline{\nu} \mid \underline{\lambda}\overline{\sigma} \right)_{\eta\eta} (\mu\nu \parallel \lambda\sigma) \\
&+ \sum_{\nu\lambda\sigma}^{N} \left(\mu'\overline{\nu} \mid \underline{\lambda}\overline{\sigma} \right)_{\eta\eta'} (\mu\nu \mid \lambda\sigma) \quad (9)
\end{aligned}
$$

and

$$\underline{R}^{\eta}_{\nu'\nu}(\zeta) = \underline{R}^{\eta,SS}_{\nu'\nu}(\zeta) + \underline{R}^{\eta,OS}_{\nu'\nu}(\zeta)$$

$$= \sum_{\mu\lambda\sigma}^{N} \left(\underline{\mu}\nu' \mid \underline{\lambda}\overline{\sigma} \right)_{\eta\eta} (\mu\nu \parallel \lambda\sigma)$$

$$+ \sum_{\mu\lambda\sigma}^{N} \left(\underline{\mu}\nu' \mid \underline{\lambda}\overline{\sigma} \right)_{\eta\eta'} (\mu\nu \mid \lambda\sigma). \quad (10)$$

Following the density matrix-based formulation of the PDs,[34,55] Schweizer *et al.*[48] introduced an efficient formulation of their derivative based on an equivalent formulation of the energy-weighted density matrix,[56,57]

$$\frac{\partial \underline{P}^{\eta}}{\partial \xi} = \frac{\partial e^{t\zeta \mathbf{P}^{\eta}_{occ} \mathbf{F}^{\eta}}}{\partial \xi} \mathbf{P}^{\eta}_{occ} + e^{t\zeta \mathbf{P}^{\eta}_{occ} \mathbf{F}^{\eta}} \frac{\partial \mathbf{P}^{\eta}_{occ}}{\partial \xi},$$

$$\frac{\partial \overline{P}^{\eta}}{\partial \xi} = \frac{\partial e^{-t\zeta \mathbf{P}^{\eta}_{virt} \mathbf{F}^{\eta}}}{\partial \xi} \mathbf{P}^{\eta}_{virt} + e^{-t\zeta \mathbf{P}^{\eta}_{virt} \mathbf{F}^{\eta}} \frac{\partial \mathbf{P}^{\eta}_{virt}}{\partial \xi}. \quad (11)$$

The perturbed matrix exponentials can be obtained by expansion in power series[58,59] and numerical difficulties can be avoided by using the method of squaring and scaling.[59] As shown by Schweizer *et al.*[48] the corresponding equation for the computation of the MP2 gradient using Eq. (11) can be re-ordered, leading to the following asymptotically linear-scaling equation for the computation of AO-MP2 energy gradients:

$$\frac{\partial \mathcal{E}^{\eta}(\zeta)}{\partial \xi} = 2I^{\eta,\xi} + 2\text{Tr}\left[\left(\overline{\mathbf{Y}}^{\eta}_2 + \underline{\mathbf{Y}}^{\eta}_2 \right) \mathbf{F}^{\eta,(\xi)} \right]$$

$$+ 2\text{Tr}\left[\left(\overline{\mathbf{Y}}^{\eta}_1 - \underline{\mathbf{Y}}^{\eta}_1 + \mathbf{G}[\overline{\mathbf{Y}}^{\eta}_2 + \underline{\mathbf{Y}}^{\eta}_2] + \mathbf{J}[\overline{\mathbf{Y}}^{\eta'}_2 + \underline{\mathbf{Y}}^{\eta'}_2] \right. \right.$$

$$\left. + \overline{\mathbf{R}}^{\eta} e^{t\zeta \mathbf{P}^{\eta}_{occ} \mathbf{F}^{\eta}} - \underline{\mathbf{R}}^{\eta} e^{-t\zeta \mathbf{P}^{\eta}_{virt} \mathbf{F}^{\eta}} \right) \frac{\partial \mathbf{P}^{\eta}_{occ}}{\partial \xi} \right]$$

$$+ 2\text{Tr}\left[-\left(\underline{\mathbf{Y}}^{\eta}_1 + \underline{\mathbf{R}}^{\eta} e^{-t\zeta \mathbf{P}^{\eta}_{virt} \mathbf{F}^{\eta}} \right) \right.$$

$$\left. \times \left(\mathbf{P}^{\eta}_{occ} + \mathbf{P}^{\eta}_{virt} \right) \frac{\partial \mathbf{S}}{\partial \xi} \mathbf{S}^{-1} \right]$$

$$= 2I^{\eta,\xi} + 2\text{Tr}\left[\mathcal{F}^{\eta} \mathbf{F}^{\eta,(\xi)} \right] + 2\text{Tr}\left[\mathcal{P}^{\eta} \frac{\partial \mathbf{P}^{\eta}_{occ}}{\partial \xi} \right]$$

$$+ 2\text{Tr}\left[\mathcal{S}^{\eta} \frac{\partial \mathbf{S}}{\partial \xi} \mathbf{S}^{-1} \right], \quad (12)$$

where $\mathbf{F}^{\eta,(\xi)}$ denotes the derivative of the Fock-matrix with respect to the basis functions only, and where $\mathbf{G} = \mathbf{J} - \mathbf{K}$ denotes the usual two-electron contributions by the Coulomb term $\mathbf{J}$ and the exchange term $\mathbf{K}$.

The $\mathbf{Y}$-matrices are built in recursions from the corresponding $\mathbf{R}$-matrices; a detailed description can be found in Ref. 48. The expensive determination of the explicit first derivative of the density matrix can be avoided using our AO-based reformulation of the Z-vector method[48] that had been originally formulated by Handy and Schaefer in the MO-basis.[60,61] Hereby, Eq. (12) can be abbreviated as

$$\frac{\partial \mathcal{E}^{\eta}(\zeta)}{\partial \xi} = 2\text{Tr}\left[\mathcal{P}^{\eta} \frac{\partial \mathbf{P}^{\eta}_{occ}}{\partial \xi} \right] + 2\text{Tr}\left[\mathcal{X}^{\eta,\xi} \right], \quad (13)$$

where the contributions independent of the perturbed density matrix are collected in $\mathcal{X}$. Subsequently, the Z-vector technique can be applied,

$$\mathcal{P}^{\eta} \frac{\partial \mathbf{P}^{\eta}_{occ}}{\partial \xi} = \underbrace{\mathcal{P}^{\eta} (\mathbf{A}^{\eta})^{-1}}_{(\mathbf{Z}^{\eta})^T} \mathbf{B}^{\eta,\xi} \quad (14)$$

with $\mathbf{B}^{\eta,\xi}$ as the right hand side of the Coupled-Perturbed Self-Consistent Field Theory (CPSCF) equations. This leads to the following reformulation of the trace summation in Eq. (13):

$$\text{Tr}\left[\mathcal{P}^{\eta} \frac{\partial \mathbf{P}^{\eta}_{occ}}{\partial \xi} \right] = \text{Tr}\left[(\mathbf{Z}^{\eta})^T \mathbf{B}^{\eta,\xi} \right]. \quad (15)$$

Using the efficient density matrix-based Laplace-transformed CPSCF (DL-CPSCF) by Beer and Ochsenfeld,[62] one obtains the following unrestricted Z-vector equation:

$$\mathbf{Z}^{\eta-vo} = \sum_{\zeta}^{\tau} w_{\zeta} \overline{\mathbf{P}}^{\eta} \left(\mathcal{P}^{\eta-vo} - \mathbf{G}^{\eta-vo} \left[\mathbf{Z}^{\eta-vo} + \mathbf{Z}^{\eta-ov} \right] \right.$$

$$\left. - \mathbf{J}^{\eta-vo} \left[\mathbf{Z}^{\eta'-vo} + \mathbf{Z}^{\eta'-ov} \right] \right) \underline{\mathbf{P}}^{\eta}, \quad (16)$$

where the Coulomb-type two-electron integral $\mathbf{J}^{\eta-vo}$ depends on the Z-vector of the opposite spin η'. It has to be noted that for real perturbations in the unrestricted open-shell case, the two Z-vectors need to be determined simultaneously. Here, the η–vo superscript describes the virtual/occupied-projection into the η-space.[63]

B. RI-CDD approach

Based on this linear-scaling AO-MP2 energy gradient theory, an efficient algorithm with a non-redundant local basis set can be obtained via the RI-CDD ansatz,[40] combining the RI approximation (also known as density fitting)[41–47] to reduce the prefactor in the scaling behavior with a Cholesky decomposition of the PDs to remove redundancies in the AO basis set.[37–39] The use of density fitting is advisable in the case of AO-MP2 gradients, as this avoids the tedious and computational time- and memory-intensive book-keeping necessary to obtain a linear scaling computation of the transformed two-electron integrals.

In the RI-CDD approach, the PDs in the AO-formalism are decomposed using a completely pivoted Cholesky decomposition[64,65] to account for their semi-definiteness,

$$\underline{P}^{\eta}_{\mu'\mu} = \sum_{i}^{\text{rank}(\mathbf{P}_{occ})=occ} \underline{L}^{\eta}_{\mu'\underline{i}} \underline{L}^{\eta}_{\mu\underline{i}},$$

$$\overline{P}^{\eta}_{\nu'\nu} = \sum_{a}^{\text{rank}(\mathbf{P}_{virt})=virt} \overline{L}^{\eta}_{\nu'\overline{a}} \overline{L}^{\eta}_{\nu\overline{a}}. \quad (17)$$

$\underline{\mathbf{L}}$ and $\overline{\mathbf{L}}$ can be considered as coefficient matrices of localized occupied and virtual pseudo-MOs (LPMOs).[40] The number of occupied and virtual LPMOs is equal to or smaller than the rank of the PDs, which is linked to the size of the respective MO space.

Decomposing the PDs in the computation of the $\mathbf{R}$-matrices and of $I^{\eta,\xi}$ accordingly leads to

$$\overline{R}^{\eta}_{\mu'\mu}(\zeta) = \sum_{j}^{\text{occ}} \sum_{a,b}^{\text{virt}} \left(\mu'\overline{a} \mid \underline{j}\overline{b}\right)_{\eta\eta} \left(\mu\overline{a} \parallel \underline{j}\overline{b}\right)_{\eta\eta}$$
$$+ \sum_{j}^{\text{occ}} \sum_{a,b}^{\text{virt}} \left(\mu'\overline{a} \mid \underline{j}\overline{b}\right)_{\eta\eta'} \left(\mu\overline{a} \mid \underline{j}\overline{b}\right)_{\eta\eta'}, \quad (18)$$

$$\underline{R}^{\eta}_{\nu'\nu}(\zeta) = \sum_{i,j}^{\text{occ}} \sum_{b}^{\text{virt}} \left(i\nu' \mid \underline{j}\overline{b}\right)_{\eta\eta} \left(i\nu \parallel \underline{j}\overline{b}\right)_{\eta\eta}$$
$$+ \sum_{i,j}^{\text{occ}} \sum_{b}^{\text{virt}} \left(i\nu' \mid \underline{j}\overline{b}\right)_{\eta\eta'} \left(i\nu \mid \underline{j}\overline{b}\right)_{\eta\eta'}, \quad (19)$$

and

$$I^{\eta,\xi}(\zeta) = \sum_{i,j}^{\text{occ}} \sum_{a,b}^{\text{virt}} \left(i\overline{a} \mid \underline{j}\overline{b}\right)_{\eta\eta} \left(i\overline{a} \parallel \underline{j}\overline{b}\right)^{(\xi)}_{\eta\eta}$$
$$+ \sum_{i,j}^{\text{occ}} \sum_{a,b}^{\text{virt}} \left(i\overline{a} \mid \underline{j}\overline{b}\right)_{\eta\eta'} \left(i\overline{a} \mid \underline{j}\overline{b}\right)^{(\xi)}_{\eta\eta'}, \quad (20)$$

where (ξ) indicates the perturbation with respect to the basis functions only, and $|\underline{i}\rangle_\eta = \sum_\mu \underline{L}^\eta_{\mu\underline{i}} |\mu\rangle$ and $|\overline{a}\rangle_\eta = \sum_\nu \overline{L}^\eta_{\nu\overline{a}} |\nu\rangle$.

Following the approach by Maurer *et al.*,[40] we furthermore introduce the RI approximation. The RI-approximated calculation of the fully transformed two-electron integrals using the Coulombic metric[44] reads as

$$\left(i\overline{a} \mid \underline{j}\overline{b}\right)_{\eta\eta'} = \sum_{PQ}^{N_{\text{aux}}} \left(i\overline{a} \mid P\right)_\eta (P \mid Q)^{-1} \left(Q \mid \underline{j}\overline{b}\right)_{\eta'}, \quad (21)$$

where P and Q denote the auxiliary basis. The inverse Coulomb metric $(P \mid Q)^{-1}$ is separated to describe the RI-approximated two-electron integral as a contraction of three-center matrices **B**,

$$\left(i\overline{a} \mid \underline{j}\overline{b}\right)_{\eta\eta'} = \sum_{P}^{N_{\text{aux}}} B^\eta_{i\overline{a},P} B^{\eta'}_{\underline{j}\overline{b},P},$$
$$\text{with } B^\eta_{i\overline{a},P} = \sum_{Q}^{N_{\text{aux}}} \left(i\overline{a} \mid Q\right)_\eta (Q \mid P)^{-\frac{1}{2}}. \quad (22)$$

The necessary transformed matrices $B^\eta_{i\nu,P}$, $B^\eta_{i\overline{a},P}$ and $B^\eta_{\mu\overline{a},P}$ are obtained from the untransformed three-center integrals in the AO-basis as follows:

$$B^\eta_{\mu\overline{a},P} = \sum_{\nu}^{N} \overline{L}^\eta_{\nu\overline{a}} B_{\mu\nu,P},$$
$$B^\eta_{i\nu,P} = \sum_{\mu}^{N} \underline{L}^\eta_{\mu\underline{i}} B_{\mu\nu,P}, \quad (23)$$
$$B^\eta_{i\overline{a},P} = \sum_{\nu}^{N} \overline{L}^\eta_{\nu\overline{a}} B^\eta_{i\nu,P}.$$

The transformation with the coefficient matrix of the occupied LPMOs is an asymptotically quadratic scaling step with respect to the size of the basis set at constant size of the molecular system, as is the consecutive transformation with its virtual variant. The transformation of $B_{\mu\nu,P}$ to obtain $B^\eta_{\mu\overline{a},P}$, however, scales in a cubical manner.

The description of the perturbed two-electron integral using the RI-approximation reads as

$$\frac{\partial (\mu\nu \mid \lambda\sigma)}{\partial \xi} = \sum_{P}^{N_{\text{aux}}} \left[\frac{\partial B_{\mu\nu,P}}{\partial \xi} B_{\lambda\sigma,P} + B_{\mu\nu,P} \frac{\partial B_{\lambda\sigma,P}}{\partial \xi} \right]$$
$$- \sum_{P,Q}^{N_{\text{aux}}} C_{\mu\nu,P}(P \mid Q)^\xi C_{\lambda\sigma,Q},$$
$$\text{with } \frac{\partial B_{\mu\nu,P}}{\partial \xi} = \sum_{Q}^{N_{\text{aux}}} \frac{\partial (\mu\nu \mid Q)}{\partial \xi}(Q \mid P)^{-\frac{1}{2}}$$
$$\text{and } C_{\mu\nu,P} = \sum_{Q}^{N_{\text{aux}}} (\mu\nu \mid Q)(Q \mid P)^{-1}. \quad (24)$$

For the computation of RI-CDD MP2 energy gradients, perturbed two-electron integrals need to be fully transformed with the Cholesky decomposed PDs [see Eq. (20)]. This computationally expensive transformation would thus need to be performed per perturbation. In order to circumvent this, we aim to pre-compute all perturbation-independent quantities.

First, Ξ-matrices are computed for the SS- and the OS-case,

$$\Xi^{\eta,\text{OS}}_{i\overline{a},P} = \sum_{j}^{\text{occ}} \sum_{b}^{\text{virt}} B^{\eta'}_{\underline{j}\overline{b},P} \left(i\overline{a} \mid \underline{j}\overline{b}\right)_{\eta\eta'},$$
$$\Xi^{\eta,\text{SS}}_{i\overline{a},P} = \sum_{j}^{\text{occ}} \sum_{b}^{\text{virt}} B^{\eta}_{\underline{j}\overline{b},P} \left(i\overline{a} \parallel \underline{j}\overline{b}\right)_{\eta\eta},$$
$$\Xi^{\eta,\text{OS}}_{\underline{j}\overline{b},P} = \sum_{i}^{\text{occ}} \sum_{a}^{\text{virt}} B^{\eta}_{i\overline{a},P} \left(i\overline{a} \mid \underline{j}\overline{b}\right)_{\eta\eta'},$$
$$\Xi^{\eta,\text{SS}}_{\underline{j}\overline{b},P} = \sum_{i}^{\text{occ}} \sum_{a}^{\text{virt}} B^{\eta}_{i\overline{a},P} \left(i\overline{a} \parallel \underline{j}\overline{b}\right)_{\eta\eta}, \quad (25)$$

which form the basis of Γ-matrices. These correspond to the corrections to the two-particle density matrix in canonical RI-MP2 analytic gradient theory[46,66] and include all perturbation-independent quantities,

$$\Gamma^\eta_{\mu\nu,P} = \sum_{i}^{\text{occ}} \sum_{a}^{\text{virt}} \underline{L}^\eta_{\mu\underline{i}} \left(\Xi^{\eta,\text{SS}}_{i\overline{a},P} + \Xi^{\eta,\text{OS}}_{i\overline{a},P} \right) \overline{L}^\eta_{\nu\overline{a}},$$
$$\Gamma^\eta_{\lambda\sigma,P} = \sum_{j}^{\text{occ}} \sum_{b}^{\text{virt}} \underline{L}^\eta_{\lambda\underline{j}} \Xi^{\eta,\text{SS}}_{\underline{j}\overline{b},P} \overline{L}^\eta_{\sigma\overline{b}} + \sum_{j}^{\text{occ}} \sum_{b}^{\text{virt}} \underline{L}^{\eta'}_{\lambda\underline{j}} \Xi^{\eta,\text{OS}}_{\underline{j}\overline{b},P} \overline{L}^{\eta'}_{\sigma\overline{b}}. \quad (26)$$

Furthermore, the perturbation of the auxiliary space needs to be considered,

$$\Gamma^{\eta,\xi}(\zeta) = \sum_{P}^{N_{\text{aux}}} \sum_{i,j}^{\text{occ}} \sum_{a,b}^{\text{virt}} D^{\eta,\xi}_{i\overline{a},P} C^\eta_{\underline{j}\overline{b},P} \left(i\overline{a} \parallel \underline{j}\overline{b}\right)_{\eta\eta}$$
$$+ \sum_{P}^{N_{\text{aux}}} \sum_{i,j}^{\text{occ}} \sum_{a,b}^{\text{virt}} D^{\eta,\xi}_{i\overline{a},P} C^{\eta'}_{\underline{j}\overline{b},P} \left(i\overline{a} \mid \underline{j}\overline{b}\right)_{\eta\eta'}, \quad (27)$$

$$\text{with } D^{\eta,\xi}_{i\overline{a},P} = \sum_{Q}^{N_{\text{aux}}} C^\eta_{i\overline{a},Q}(Q \mid P)^\xi.$$

With these matrices, $I^{\eta,\xi}(\zeta)$ is computed by

$$I^{\eta,\xi}(\zeta) = \sum_{\mu\nu}^{N} \sum_{P}^{N_{aux}} \Gamma^{\eta}_{\mu\nu,P} \frac{\partial B_{\mu\nu,P}}{\partial \xi} + \sum_{\lambda\sigma}^{N} \sum_{P}^{N_{aux}} \Gamma^{\eta}_{\lambda\sigma,P} \frac{\partial B_{\lambda\sigma,P}}{\partial \xi}$$
$$- \Gamma^{\eta,\xi}(\zeta). \tag{28}$$

C. QQR-type integral estimation

To exploit the asymptotic scaling behavior, significant integral contributions to the **R**-matrices and to the $I^{\eta,\xi}$ term need to be determined. The standard estimates for the two-electron integrals in quantum chemistry are the classical Schwarz estimates.[41,67] In order to obtain the asymptotic scaling behavior, the distance-dependency of the two-electron integrals needs to be incorporated in the estimation technique. This can be achieved by the so-called QQR-estimation.[35,36]

The decay behavior of the two-electron integral with the separation of the charge distributions Ω_A and Ω_B in bra and in ket can be concluded from a multipole expansion,[36]

$$(\Omega_A \mid \Omega_B) = \frac{q^A_{00} q^B_{00}}{R_{AB}} + \frac{q^A_{00} \left(\sum_{j=-1}^{1} T'_{00,1j} q^B_{1j} \right)}{R^2_{AB}}$$
$$+ \frac{\left(\sum_{i=-1}^{1} q^A_{1i} T'_{1i,00} \right) q^B_{00}}{R^2_{AB}} + \mathcal{O}(R^{-3}_{AB}). \tag{29}$$

Transforming the two-electron integral with PDs as in Eq. (2), the monopoles q_{00} are zero due to the orthogonality of the occupied and the virtual subspace.[36] This also applies to the transformation with the Cholesky decomposed matrices, as proven in Appendix B of Ref. 40. Therefore, fully transformed two-electron integrals exhibit a $1/(R')^3$ decay behavior, as both fully transformed monopoles q^A_{00} and q^B_{00} are zero and the decay behavior is controlled by the $\mathcal{O}(R^{-3}_{AB})$ term in Eq. (29). This leads to the following QQR-type estimate:

$$\left(\underline{i\overline{a}} \mid \underline{j\overline{b}} \right)_{\eta\eta'} \approx \frac{Z^{\eta}_{\underline{i\overline{a}}} Z^{\eta'}_{\underline{j\overline{b}}}}{\left(R - \text{ext}^{\eta}_{\underline{i\overline{a}}} - \text{ext}^{\eta'}_{\underline{j\overline{b}}} \right)^3}, \tag{30}$$

where the **Z**-matrices are the common pseudo-Schwarz matrices introduced by Häser.[31] The following equation gives the definition of the fully and half-transformed pseudo-Schwarz matrices, where the transformation is performed with the Cholesky decomposed PDs:

$$X^{\eta}_{\underline{i\nu}} = \left(\underline{i\nu} \mid \underline{i\nu} \right)^{\frac{1}{2}}_{\eta\eta},$$

$$Y^{\eta}_{\mu\overline{a}} = (\mu\overline{a} \mid \mu\overline{a})^{\frac{1}{2}}_{\eta\eta}, \tag{31}$$

$$Z^{\eta}_{\underline{i\overline{a}}} = \left(\underline{i\overline{a}} \mid \underline{i\overline{a}} \right)^{\frac{1}{2}}_{\eta\eta}.$$

The multipole expansion used to determine the decay behavior of the two-electron integrals in Eq. (29) is only valid when the charge distributions Ω_A and Ω_B are well separated.[36] The criterion chosen in the present work is given by

$$R_{A\to B} - \text{ext}_A - \text{ext}_B \geq 1, \tag{32}$$

where the distance between Ω_A and Ω_B is corrected by their extents. The extents of the untransformed integrals are computed as shown in Appendix B of Ref. 35. The extents for

fully transformed bra or ket are also used for RI-CDD-MP2 energies.[40] Their formulation, as well as the computation of the extents of half-transformed charge distributions required in Eqs. (34) and (35), is shown in the Appendix of our present work. If the criterion in Eq. (32) is not met, common Schwarz estimates are used.

With this, the QQR-type estimate for $I^{\eta,\xi}$ in Eq. (20) can be written as

$$\left(\underline{i\overline{a}} \mid \underline{j\overline{b}} \right)_{\eta\eta'} \left(\underline{i\overline{a}} \mid \underline{j\overline{b}} \right)^{(\xi)}_{\eta\eta'} \approx \frac{\left(Z^{\eta}_{\underline{i\overline{a}}} \right)^2 \left(Z^{\eta'}_{\underline{j\overline{b}}} \right)^2}{\left(R - \text{ext}^{\eta}_{\underline{i\overline{a}}} - \text{ext}^{\eta'}_{\underline{j\overline{b}}} \right)^6}. \tag{33}$$

Hereby, we exploited the fact that the perturbed integral can be screened using the estimates of the nonperturbed integral.[68,69]

In the computation of **R**-matrices in Eqs. (18) and (20), the QQR-type estimates read as follows:

$$\left(\mu'\overline{a} \mid \underline{j\overline{b}} \right)_{\eta\eta'} \left(\mu\overline{a} \mid \underline{j\overline{b}} \right)_{\eta\eta'}$$
$$\approx \frac{Y^{\eta}_{\mu'\overline{a}} Z^{\eta'}_{\underline{j\overline{b}}}}{\left(R - \text{ext}^{\eta}_{\mu'\overline{a}} - \text{ext}^{\eta'}_{\underline{j\overline{b}}} \right)^2} \frac{Y^{\eta}_{\mu\overline{a}} Z^{\eta'}_{\underline{j\overline{b}}}}{\left(R - \text{ext}^{\eta}_{\mu\overline{a}} - \text{ext}^{\eta'}_{\underline{j\overline{b}}} \right)^2} \tag{34}$$

and

$$\left(\underline{i\nu'} \mid \underline{j\overline{b}} \right)_{\eta\eta'} \left(\underline{i\nu} \mid \underline{j\overline{b}} \right)_{\eta\eta'}$$
$$\approx \frac{X^{\eta}_{\underline{i\nu'}} Z^{\eta'}_{\underline{j\overline{b}}}}{\left(R - \text{ext}^{\eta}_{\underline{i\nu'}} - \text{ext}^{\eta'}_{\underline{j\overline{b}}} \right)^2} \frac{X^{\eta}_{\underline{i\nu}} Z^{\eta'}_{\underline{j\overline{b}}}}{\left(R - \text{ext}^{\eta}_{\underline{i\nu}} - \text{ext}^{\eta'}_{\underline{j\overline{b}}} \right)^2}. \tag{35}$$

These triple-transformed integrals have a decay behavior of $1/(R')^2$ because only the monopoles of $q^{\underline{j\overline{b}}}_{00}$ of the multipole expansion are zero.

The computational cost of the screening procedure can be reduced by pre-screenings. In the computation of $I^{\eta,\xi}$ using the RI-CDD approach, significant electron pairs (i, j) are pre-selected by

$$\left(\underline{i\overline{a}} \mid \underline{j\overline{b}} \right)_{\eta\eta'} \left(\underline{i\overline{a}} \mid \underline{j\overline{b}} \right)^{(\xi)}_{\eta\eta'} \lessapprox \frac{\left(\max_{\overline{a}} Z^{\eta}_{\underline{i\overline{a}}} \right)^2 \left(\max_{\overline{b}} Z^{\eta'}_{\underline{j\overline{b}}} \right)^2}{\left(R^{\eta\eta'}_{\underline{i,j}} - \text{ext}^{\eta}_{\underline{i}} - \text{ext}^{\eta'}_{\underline{j}} \right)^6}. \tag{36}$$

The respective centers and extents can be found in the Appendix.

In the computation of **R** using the RI-CDD approach, significant electron pairs (i, j) are preselected, followed by a determination of the significant (i, j, b) combinations, while for $\overline{\textbf{R}}$, significant (a, b) pairs are determined first, for which significant (a, b, j) combinations are selected afterwards.

D. First-order properties: Closed-shell nuclear gradients and hyperfine coupling constants

With the introduced theory for the computation of energy gradients at the MP2-level using the RI-CDD approach, nuclear gradients and first-order properties can be computed. In this work we present both the computation of closed-shell nuclear gradients and of HFCCs.

In the case of a closed-shell system, the above introduced equations simplify considerably. In the RI-CDD approach, the closed-shell gradient contribution at Laplace point ζ is given by

$$\frac{\partial \mathcal{E}^\eta(\zeta)}{\partial \xi} = 2I^\xi(\zeta) + 2\sum_{\mu'\mu} \overline{R}_{\mu'\mu}(\zeta)\frac{\partial \underline{P}^\eta_{\mu'\mu}}{\partial \xi} + 2\sum_{\nu'\nu} \underline{R}_{\nu'\nu}(\zeta)\frac{\partial \overline{P}^\eta_{\nu'\nu}}{\partial \xi},$$

(37)

with the closed-shell I^ξ term,

$$I^\xi(\zeta) = \sum_{i,j}^{\text{occ}} \sum_{a,b}^{\text{virt}} \left(\underline{i\overline{a}} \mid \underline{j\overline{b}} \right)^{(\xi)} \left[2\left(\underline{i\overline{a}} \mid \underline{j\overline{b}} \right) - \left(\underline{i\overline{b}} \mid \underline{j\overline{a}} \right) \right], \quad (38)$$

and the closed-shell **R**-matrices,

$$\overline{R}_{\mu'\mu}(\zeta) = \sum_{j}^{\text{occ}} \sum_{a,b}^{\text{virt}} \left(\mu'\overline{a} \mid \underline{j\overline{b}} \right) \left[2\left(\mu\overline{a} \mid \underline{j\overline{b}} \right) - \left(\mu\overline{b} \mid \underline{j\overline{a}} \right) \right],$$

$$\underline{R}_{\nu'\nu}(\zeta) = \sum_{i,j}^{\text{occ}} \sum_{b}^{\text{virt}} \left(\underline{i\nu'} \mid \underline{j\overline{b}} \right) \left[2\left(\underline{i\nu} \mid \underline{j\overline{b}} \right) - \left(\underline{i\overline{b}} \mid \underline{j\nu} \right) \right].$$

(39)

The integral estimates introduced in Eqs. (33)–(35) can be adapted straightforwardly.

HFCCs in the absence of spin-orbit coupling can be obtained as a first-derivative of the unrestricted MP2 equation with respect to the nuclear magnetic moment M_k of nucleus k, given by the isotropic Fermi-contact term and the anisotropic spin-dipole interaction.[2] As the basis functions are independent of the nuclear magnetic moment, only the **R**-matrices need to be considered in Eq. (7), leading to

$$\frac{\partial \mathcal{E}^\eta(\zeta)}{\partial M_k} = 2\sum_{\mu'\mu} \overline{R}^\eta_{\mu'\mu}(\zeta)\frac{\partial \underline{P}^\eta_{\mu'\mu}}{\partial M_k} + 2\sum_{\nu'\nu} \underline{R}^\eta_{\nu'\nu}(\zeta)\frac{\partial \overline{P}^\eta_{\nu'\nu}}{\partial M_k}, \quad (40)$$

where the **R**-matrices can be computed using the RI-CDD approach as in Eqs. (18) and (19).

In both the computation of closed-shell nuclear gradients and of HFCCs, the **R**-matrices are not directly contracted with perturbed pseudodensities, but instead the Z-vector technique outlined in Eq. (14) is used to compute the contribution per Laplace point as in Eq. (12).

III. COMPUTATIONAL DETAILS

The RI-CDD MP2 energy gradients and the HFCCs were implemented in the program package FermiONs++.[70,71] The reference RI-MP2 nuclear gradients[66] were computed using the implementation in Q-Chem 4.0.[72] The reference HFCCs at the RI-MP2 level were obtained with the program package ORCA.[73] In the non-canonical computations, the PDs are scaled with the coefficients of the Laplace expansion. The selection of linear and exponential Laplace expansion coefficients is performed according to the minimax-approximation by Hackbusch and co-workers.[74] Closed-shell nuclear gradients were obtained using five Laplace expansion points, based on the study by Schweizer *et al.*[48] The extents of the QQR-type integral estimates are determined with the same thresholds as in Ref. 36, i.e., the threshold of the untransformed AO extent is 0.1 and the threshold of the transformed extents is 10^{-3}. The SCF equations were converged at least to a threshold of

10^{-8}. The DL-CPSCF equations for the determination of the Z-vector are converged to a threshold of 10^{-4} for the nuclear gradients, whereas a tighter threshold of 10^{-6} is chosen for the HFCCs. The basis sets def2-SVP[75] and cc-pVTZ[76,77] are used. For the auxiliary space in the RI-approximation, optimized basis sets def2-SVP-RI[47,78] and cc-pVTZ-RI[79,80] are used.

The sparsity of the **L**-matrices is exploited by the use of the efficient block-compressed sparse row (BCSR) matrix multiplications[63,81] with a sparsity criterion of 10^{-7}. A reordering of the atoms in the molecule according to the reverse Cuthill-McKee algorithm[82] is performed by reducing the bandwidth of the connectivity matrix. Subsequently, the structure of the Cholesky decomposed matrices is reestablished as described in Ref. 40. The computation of the right hand side of the DL-CPSCF equation for the nuclear gradients in Eq. (15) involves a contraction of the term $\mathbf{PS}^\xi \mathbf{P}$ with the exchange two-electron integral. Here, the sparsity of $\mathbf{PS}^\xi \mathbf{P}$ was exploited by the BCSR scheme.

The scaling behavior is computed based on the number of basis functions. The scaling behavior at the first Laplace point is assumed to represent the overall behavior.

The timings were performed on a single core of an Intel Xeon E5-2620 using up to 128 GB of RAM.

IV. RESULTS

The following analyzes the accuracy and efficiency of both our closed-shell nuclear gradients and HFCC computations using the RI-CDD ansatz. We start by investigating the accuracy with respect to the canonical formulation, followed by an analysis of the scaling behavior of the rate-determining steps. Finally, we compare our computational cost to the canonical variants for molecules of increasing size.

A. Accuracy

We will first start by analyzing the accuracy of the computation of closed-shell nuclear gradients. Therefore, we computed a set of molecules comprising linear alkanes, glycine chains, and DNA base pairs with our RI-CDD ansatz using different screening thresholds.

Preliminary studies suggested the use of separate screening thresholds for the **R**-matrices and for the I^ξ term in Eqs. (18)–(20). The deviation from the RI-MP2 gradients is shown in Table I, highlighting a fully controllable accuracy. We furthermore compare the accuracy of our RI-CDD gradients of Gly_5 to the canonical MP2 gradient without the RI-approximation, which confirms the correct QQR-screening. A screening threshold of 10^{-6} for the **R**-matrices and 10^{-8} for I^ξ is sufficient for sub-mhartree/bohr accuracy.

Following this analysis of the accuracy of closed-shell nuclear gradients, we investigated the computation of HFCCs. The accuracy of the computation of HFCCs at the RI-CDD MP2-level with respect to its canonical variant depends on different factors: In addition to the QQR-type screening threshold, whose influence was investigated above for the closed-shell nuclear gradients, the dependency of the accuracy on the number of Laplace expansion points needs to be taken into

TABLE I. Root mean square deviation in mhartree/bohr of RI-CDD MP2 energy gradients with respect to conventional RI-MP2 energy gradients for different QQR-type screening thresholds (for $\mathbf{R}$-matrices/for I^ξ) with the basis set def2-SVP. Furthermore, the results of Gly$_5$ are compared to the MO-MP2 result and the DNA$_1$ computation is additionally performed using seven Laplace expansion points.

Molecule	$10^{-5}/10^{-7}$	$10^{-6}/10^{-8}$	$10^{-7}/10^{-9}$
C$_5$H$_{12}$	0.50	0.01	0.02
C$_{10}$H$_{22}$	0.71	0.03	<0.01
C$_{20}$H$_{42}$	1.15	0.08	0.04
Gly$_5$	0.56	0.07	0.10[a]
Gly$_5$ (MO-MP2)	0.56	0.08	0.10[a]
DNA$_1$ (5 LP)	3.74	0.48	0.38
DNA$_1$ (7 LP)	3.92	0.41	0.29

[a]The remaining small deviation can be attributed to the number of Laplace expansion points. Using seven Laplace points, an error of 0.04 mhartree/bohr to RI-MP2 and 0.05 mhartree/bohr to MO-MP2 using screening thresholds of $10^{-7}/10^{-9}$ is achieved.

account. In contrast to the nuclear gradients, whose dependency of the accuracy on the Laplace expansion has been studied in previous work,[48] a similar investigation is not available for the computation of HFCCs, as a Laplace-transform based ansatz for HFCCs has not yet been presented to our knowledge. Additionally, the use of a Laplace expansion based on a generalized range may result in the need for more expansion points. Owing to the open-shell nature, the convergence of the CPSCF equations is also critical.

In order to single out the effect of the different contributors to the accuracy, we varied only one parameter and set the others to their tightest value. These were chosen to be ten Laplace expansion points, a QQR-type screening threshold of 10^{-10}, and a CPSCF-convergence criterion of 10^{-8}. For demonstration purposes, we chose a set of radicals: CHO, NF$_2$, CH$_3$, C$_5$H$_{11}$, C$_{10}$H$_{21}$, and a *p*-benzosemiquinone radical coordinated with four water molecules abbreviated by BQ$^{\bullet-}$. The HFCCs of BQ$^{\bullet-}$ have previously been studied using both DFT and MP2-approaches.[50,83] We used the reference structure from Ref. [83] for our benchmark. Furthermore, we included two small triplet systems in our benchmark, namely, ^{3}O and ^{3}SO. As we compare to the canonical results, the structures used in this work were not extensively optimized using advanced optimization techniques. We want to point out that those radicals and triplets do not represent all the different challenges to the theoretical computation of HFCCs as at this stage we want to focus on the feasibility of our approach and aim to present a proof of principle.

Our comparison of the isotropic HFCCs to their canonical implementation can be found in Table II. The results indicate that the dependency on the different parameters varies between the different systems. The small molecular systems converged faster with a tighter QQR-threshold. QQR-type integral estimation is only performed when the charge distributions are sufficiently separated, else a common Schwarz screening is performed. Thus, the convergence with the QQR-threshold for the small systems is in fact a convergence

TABLE II. Root mean square deviation in MHz of the computed isotropic RI-CDD MP2 HFCCs with respect to the canonical RI-MP2 result for radicals, triplets, and anions using the basis set def2-SVP for different number of Laplace points, QQR-type screening thresholds, and CPSCF convergence criteria.

Molecule	Atom	QQR-threshold[a]			Number of Laplace points[a]			CPSCF-threshold[a]		
		10^{-6}	10^{-8}	10^{-10}	5	7	10	10^{-3}	10^{-5}	10^{-7}
NF$_2$	^{14}N	0.0	0.0	0.0	0.1	0.0	0.0	0.1	0.0	0.0
	^{19}F	0.1	0.0	0.0	1.2	0.2	0.0	13.3	0.0	0.0
CH$_3$	^{13}C	0.0	0.0	0.0	0.0	0.0	0.0	6.8	0.0	0.0
	^{1}H	0.0	0.0	0.0	0.0	0.0	0.0	2.4	0.0	0.0
CHO[b]	^{13}C	0.5	0.5	0.5	23.0	5.4	0.5	16.6	0.5	0.5
	^{1}H	0.1	0.1	0.1	6.4	1.5	0.1	4.3	0.1	0.1
	^{17}O	0.3	0.4	0.4	16.8	3.8	0.4	14.2	0.4	0.4
^{3}O$_2$	^{17}O	0.0	0.0	0.0	0.0	0.0	0.0	1.4	0.0	0.0
^{3}SO	^{33}S	0.0	0.0	0.0	0.1	0.0	0.0	0.1	0.0	0.0
	^{17}O	0.0	0.0	0.0	0.3	0.1	0.0	1.2	0.0	0.0
C$_5$H$_{11}$	^{13}C	0.2	0.0	0.0	0.1	0.0	0.0	2.7	0.0	0.0
	^{1}H	0.2	0.0	0.0	0.0	0.0	0.0	0.9	0.0	0.0
C$_{10}$H$_{21}$	^{13}C	2.6	0.0	0.0	0.1	0.0	0.0	1.9	0.0	0.0
	^{1}H	2.5	0.0	0.0	0.0	0.0	0.0	0.7	0.0	0.0
BQ$^{\bullet-}$	^{13}C	9.4	0.0	0.0	0.5	0.1	0.0	0.4	0.0	0.0
	^{17}O	1.8	0.3	0.0	1.2	0.6	0.0	0.5	0.0	0.0
	^{1}H	2.0	0.3	0.0	0.1	0.1	0.0	0.0	0.0	0.0

[a]The other parameters were each set to their maximum value, being ten Laplace points, a QQR-threshold of 10^{-10}, and a CPSCF-threshold of 10^{-8}.

[b]This system requires more than ten Laplace points for a converged description, resulting in the remaining errors, which can be removed completely with 14 Laplace points.

with respect to a Schwarz screening. This is not the case for our larger systems, namely, C_5H_{11}, $C_{10}H_{21}$, and $BQ^{\bullet-}$. As expected, a systematic deviation to the canonical result can be observed when loose QQR-thresholds are chosen, which is overcome with tighter thresholds. All computations show a fully controllable accuracy with the QQR-threshold and their deviation is in the sub-MHz region for a QQR-threshold of 10^{-8}.

Most of our studied systems are sufficiently described with at most seven Laplace expansion points, except for CHO. This electronically challenging system requires a high number of Laplace points for a sub-MHz description. When 14 Laplace expansion points are chosen, the remaining errors of 0.5 MHz for ^{13}C, of 0.1 MHz for ^{1}H, and of 0.4 MHz for ^{17}O in Table II are removed completely, confirming the fully controllable accuracy. In all our computations a CPSCF-threshold of 10^{-5} was sufficient to ensure a converged result with sub-MHz accuracy.

B. Scaling behavior

In order to analyze the performance of our new RI-CDD MP2 ansatz for first-order properties, we take a closer look at the rate-determining steps. For large-scale systems, the scaling behavior of these steps is of great importance. The scaling behavior can be determined with respect to the number of significant two-electron integrals and with respect to the wall time needed. The former allows the asymptotic limit to be determined, whereas the wall time provides the effective scaling behavior and indicates computational overheads. As discussed in Ref. 84, quasi-one-dimensional systems are ideally suited for a systematic study of the formal asymptotic scaling behavior and possibly remaining higher-order scaling steps. Throughout this work, the scaling behavior is computed based on the number of basis functions.

In the computation of closed-shell nuclear gradients, the evaluation of both the **R**-matrices and of I^ξ in Eqs. (38) and (39) has to be taken into account. The number of significant integrals based on a QQR-type integral estimation is shown in Table III for linear alkanes with increasing length and show that linear alkanes larger than $C_{40}H_{82}$ exhibit a scaling behavior below $N^{1.3}$.

This optimal scaling behavior with respect to the number of integrals is not directly reflected in the scaling of the

wall times, which is usually shifted to larger system sizes. In addition, when the Coulomb-metric RI-approximation is employed, the sum over the auxiliary space increases the asymptotic scaling behavior of the wall times to quadratic. This increase in the scaling behavior could be circumvented by, e.g., the use of local fitting by Werner *et al.*[85] by restricting the auxiliary space. This has been shown to lead to fast computations of the local canonical MP2 energy gradient[18] with good agreement to the exact MP2 energy gradient. Nonetheless, local domains need to be defined which may introduce additional errors.[86] Therefore, we follow the different pathway by removing zero-elements instead of constructing local fitting domains. For RI-CDD MP2 energies, it was shown that an effective reduction of the computational cost by local fitting can only be achieved for larger systems such as $C_{160}H_{322}$.[40] Therefore, local fitting was omitted in this implementation.

In Table IV, the wall times of the computation of the **R**-matrices, of I^ξ, and of the transformation of the **B**-matrices as in Eq. (23) for nuclear gradients of linear alkanes ranging from C_5H_{12} to $C_{100}H_{202}$ are shown. An asymptotic quadratic scaling behavior for the computation of the **R**-matrices is confirmed. The sub-quadratic scaling for $C_{40}H_{82}$ and $C_{60}H_{122}$ is due to the large contribution of the linear-scaling screening to the total wall time. The precontraction of the perturbation-independent quantities in Eq. (25) involves cubic scaling steps, which results in the overall cubic scaling behavior of the computation of I^ξ. The same scaling behavior is to be expected and confirmed for the transformation of the **B**-matrices, due to the formation of $B_{\mu\bar{a},P}$. The apparent fluctuation of the overall scaling behavior of the **B**-matrices and of I^ξ is caused by contributions of manifold steps of different scalings.

Turning to the computation of HFCCs, the previous statements relating to the **R**-matrices are relevant since the I^ξ terms in Eq. (20) are not necessary. The only cubic-scaling step of interest is the formation of $B_{\mu\bar{a},P}$. As can be seen in Table IV, the transformation is cheaper than the contraction of the **R**-matrices for all investigated systems. Therefore, an overall quadratic scaling is to be expected (see Sec. IV C).

TABLE III. Number of selected integral products (10^9) at the first Laplace point according to the QQR-type integral estimation in the computation of I^ξ and of the **R**-matrices in the RI-CDD-MP2 approach using a QQR-type integral screening threshold of 10^{-6} for the **R** matrices and 10^{-8} for I^ξ with the basis set def2-SVP. The respective scaling exponent is indicated in brackets.

Molecule	$\left(i\bar{a}\mid j\bar{b}\right)\left(i\bar{a}\mid\mid j\bar{b}\right)$	**R**	$\overline{\textbf{R}}$
C_5H_{12}	0.003	0.776	3.553
$C_{10}H_{22}$	0.020 (3.1)	10.192 (3.9)	25.815 (3.0)
$C_{20}H_{42}$	0.071 (1.9)	37.627 (1.9)	74.596 (1.6)
$C_{40}H_{82}$	0.174 (1.3)	92.023 (1.3)	172.470 (1.2)
$C_{60}H_{122}$	0.276 (1.2)	146.886 (1.2)	272.162 (1.1)
$C_{80}H_{162}$	0.379 (1.1)	202.224 (1.1)	372.143 (1.1)
$C_{100}H_{202}$	0.481 (1.1)	256.655 (1.1)	468.989 (1.0)

TABLE IV. Wall time in minutes of the transformation of the three-center **B**-matrices, of the QQR-estimation and of the computation of **R**-matrices, and of the total computational time needed for the calculation of I^ξ in the RI-CDD MP2 energy gradient calculation for increasing linear alkanes using the basis set def2-SVP. The wall times are taken from the first Laplace point and the scaling exponent is shown in brackets. The QQR-type integral screening thresholds were set at 10^{-6} for the **R**-matrices and at 10^{-8} for I^ξ.

Molecule	Transform **B**-matrices	**R**-matrices QQR-est.	**R**-matrices Total	I^ξ Total
C_5H_{12}	0.01	0.60	1.07	0.16
$C_{10}H_{22}$	0.11 (3.2)	3.85 (2.8)	7.55 (3.0)	1.77 (3.7)
$C_{20}H_{42}$	1.10 (3.5)	13.42 (1.9)	30.05 (2.1)	13.40 (3.0)
$C_{40}H_{82}$	9.48 (3.2)	30.95 (1.2)	90.82 (1.6)	95.08 (2.9)
$C_{60}H_{122}$	30.25 (2.9)	50.95 (1.2)	175.45 (1.6)	391.30 (3.5)
$C_{80}H_{162}$	67.38 (2.8)	72.22 (1.2)	312.37 (2.0)	993.33 (3.3)
$C_{100}H_{202}$	123.27 (2.7)	97.82 (1.4)	480.10 (1.9)	2070.00 (3.3)

C. Computational efficiency

To finally assess the computational efficiency of our method, the wall times of the gradient computations of linear alkanes, glycine chains, and DNA double strands are compared to the respective RI-MP2 computations in Tables V and VI. At the double-ζ level in Table V, a crossover to the canonical implementation can be seen for systems larger than $C_{60}H_{122}$, as visualized in Fig. 1(a). This matches the corresponding crossover of the RI-CDD MP2 energies with respect to their canonical variant.[40] In the case of the glycine chains, the crossover to the canonical computation lies between a chain length of 10 and of 20 glycine monomers, as can be seen in Fig. 1(b). The DNA fragments consisting of one and of two base pairs were both computed faster using the canonical approach. Using our present single-core implementation, neither the RI-CDD nor the canonical computations of a DNA fragment consisting of four base pairs are possible within a reasonable time frame. However, based on the scaling behavior, it is to be expected that a fragment consisting of four base pairs will be computed faster using the new RI-CDD approach.

By a double-logarithmic plot of the results in Fig. 1, the asymptotic scaling behavior can be shown. This is done in Fig. 2 for both the linear alkanes and the glycine chains, including a linear fit for the three largest systems, highlighting the asymptotic cubic scaling for our RI-CDD method and the $\mathcal{O}(N^5)$ scaling for the canonical RI-MP2 approach. Furthermore, the curvature demonstrates that our ansatz leads to a reduced scaling behavior with the increasing size of the molecule until the asymptotic cubic scaling behavior is reached. In the canonical formulation this is not the case, where the scaling behavior increases, as the dominant low scaling steps at small molecule sizes are outweighed by the high scaling contributions for larger systems.

TABLE V. Comparison of the wall time in hours for different molecules for RI-CDD MP2 and RI-MP2 energy gradient computations using the basis set def2-SVP and a screening threshold of 10^{-6} for **R**-matrices and 10^{-8} for I^{ξ}. The scaling exponent with respect to the next smaller molecular system is indicated in parentheses. Values marked by an asterisk are extrapolated based on the scaling behavior of the wall time of the next smaller system.

Molecule	RI-CDD-MP2	RI-MP2
C_5H_{12}	0.11	0.02
$C_{10}H_{22}$	1.02 (3.4)	0.12 (3.0)
$C_{20}H_{42}$	5.95 (2.6)	0.78 (2.8)
$C_{40}H_{82}$	28.24 (2.3)	10.13 (3.8)
$C_{60}H_{122}$	72.61 (2.3)	62.64 (4.5)
$C_{80}H_{162}$	151.55 (2.6)	242.42 (4.7*)
$C_{100}H_{202}$	282.43 (2.7)	715.74 (4.8*)
Gly_1	0.04	0.01
Gly_3	0.83 (3.4)	0.12 (3.00)
Gly_5	2.61 (2.4)	0.43 (2.8)
Gly_{10}	13.32 (2.5)	4.40 (3.5)
Gly_{20}	70.89 (2.5)	90.34 (4.5)
Gly_{30}	204.81 (2.7)	613.47 (4.7*)
DNA_1	16.93	3.11
DNA_2	295.81 (3.8)	78.69 (4.3)

TABLE VI. Comparison of the wall time in hours for different molecules for RI-CDD MP2 and RI-MP2 energy gradient computations using the cc-pVTZ basis and a screening threshold of 10^{-6} for **R**-matrices and 10^{-8} for I^{ξ}. In addition to the scaling exponent (SE) with respect to the size of the molecular system, shown in brackets, the SE when increasing the basis set at constant size of the molecular system is shown. This SE is computed by comparing the wall time using cc-pVTZ (TZ) with the calculation of the same molecule using def2-SVP (SVP).

	RI-CDD-MP2		RI-MP2	
Molecule	Wall time (SE)	SE basis set size (SVP → TZ)	Wall time (SE)	SE basis set size (SVP → TZ)
C_5H_{12}	2.80	3.6	0.32	3.2
$C_{10}H_{22}$	20.96 (3.1)	3.4	2.17 (2.9)	3.2
$C_{20}H_{42}$	111.11 (2.5)	3.3	13.09 (2.7)	3.2

The performance is similar at the triple-ζ level which can be seen in Table VI. The scaling behavior with respect to the size of the molecular system is analogous to the respective molecular system sizes at the double-ζ level. When the basis set is increased at a constant size of the molecule, a cubic scaling behavior is confirmed. Furthermore, the computation of $C_{20}H_{42}$ with the basis set cc-pVTZ is 8.7 times more expensive with the new RI-CDD approach than with canonical
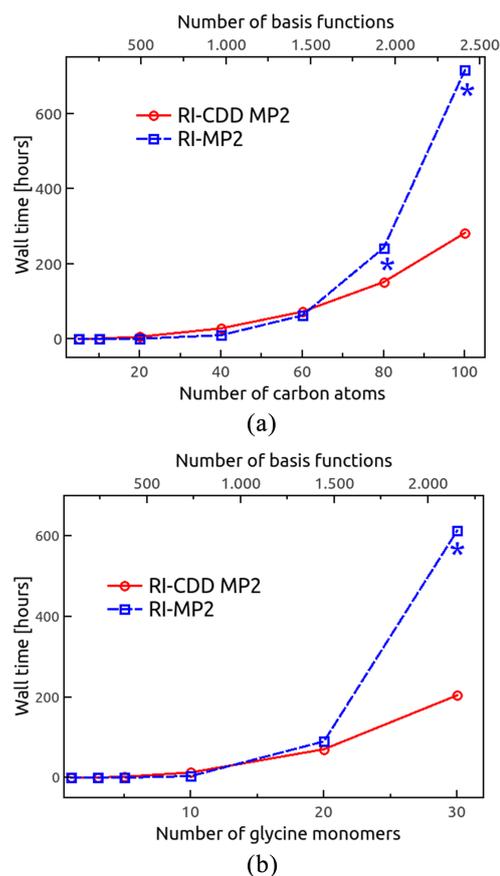
FIG. 1. Comparison of the wall times for canonical RI-MP2 and RI-CDD MP2 energy gradient computations of linear alkanes and glycine chains using the basis set def2-SVP. Different values have been conservatively extrapolated, indicated by an asterisk. The wall time behavior of the RI-MP2 gradient computation of $C_{100}H_{202}$, Gly_{30}, and $C_{80}H_{162}$ was based on the scaling behavior of the previous point. (a) Linear alkanes and (b) glycine chains.

(a)

(b)
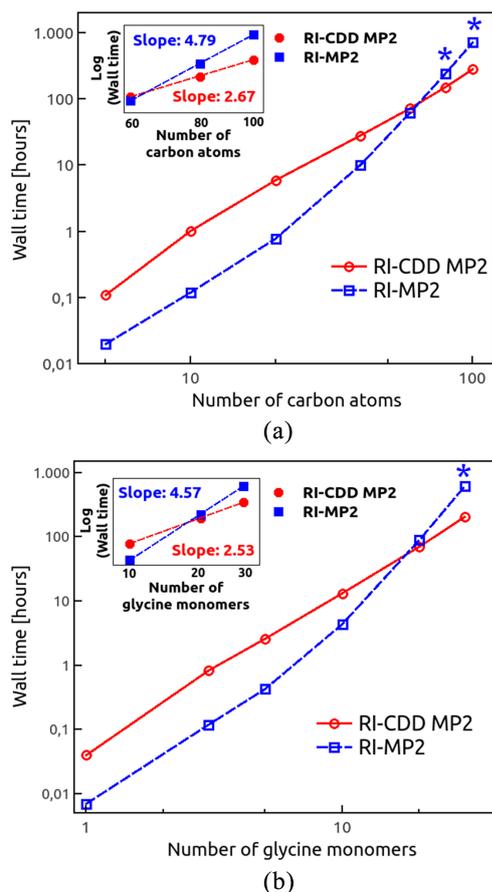
FIG. 2. Double logarithmic comparison of the wall times for canonical RI-MP2 and RI-CDD MP2 energy gradient computations of linear alkanes and glycine chains using the basis set def2-SVP to highlight the asymptotic scaling behavior. The scaling behavior of the respective three largest systems was obtained by a linear fit. The wall time of the RI-MP2 gradient computation of $C_{100}H_{202}$, Gly_{30}, and $C_{80}H_{162}$ (indicated by an asterisk) was extrapolated based on the scaling behavior of the previous point. (a) Linear alkanes and (b) glycine chains.

TABLE VII. Comparison of the wall time in hours of RI-CDD MP2 and RI-MP2 HFCC computations of different molecules with the basis set def2-SVP using seven Laplace points and a screening threshold of 10^{-8}. The scaling exponent is indicated in parentheses. Values marked by an asterisk are conservatively extrapolated based on the scaling behavior of the wall time of the next smaller system.

Molecule	RI-CDD-MP2	RI-MP2
C_5H_{11}	0.65	0.03
$C_{10}H_{21}$	5.90 (3.3)	0.22 (3.0)
$C_{20}H_{41}$	30.57 (2.4)	1.76 (3.1)
$C_{40}H_{81}$	104.75 (1.8)	33.59 (4.3)
$C_{60}H_{121}$	209.28 (1.7)	226.83 (4.7)
$C_{80}H_{161}$	375.15 (2.0)	880.95 (4.7*)
$C_{100}H_{201}$	535.69 (1.8)	2525.31 (4.7*)

overall asymptotic quadratic scaling behavior is confirmed. The efficiency crossover for HFCCs occurs below $C_{60}H_{121}$ with respect to the canonical implementation. In both the RI-CDD and the canonical approach, the computations are more expensive than the nuclear gradient computations of systems of comparable molecular size. In the case of the RI-CDD approach, this is due to the fact that **R**-matrices for the different spin cases need to be screened and contracted separately for the HFCCs. All computations were performed on a single core. We have not yet implemented a parallel version. The effectiveness of the parallelization of the RI-CDD MP2 energies[40] suggests that with our new RI-CDD ansatz, energy gradients and HFCCs can now be computed for large molecular systems in a reasonable time frame. This can furthermore be improved in the future by a migration to graphics processing units (GPU), which sped up the computation of RI-CDD MP2 energies.[87] This will equally be of importance when aiming for molecular properties, which we want to focus on in the future and will make our new approach preferable to its canonical variant for medium- to large-sized molecular systems.

V. CONCLUSION AND OUTLOOK

In this work, we presented a new approach to the computation of MP2 energy gradients, using the RI approximation and a Cholesky decomposition of the pseudodensity matrices in an AO-based formulation. This approach significantly reduces the computational effort with respect to its canonical variant and thus provides an efficient calculation of nuclear energy gradients and HFCCs at the MP2-level. Low scaling with respect to the size of the molecular system is achieved by an extension of the QQR integral estimation method to screen significant integral contributions in a fully controlled manner. With the use of QQR-type integral estimates, the number of two-electron integrals necessary in the computation of the gradients is reduced to linear, whilst the contraction step of the two-electron integrals is reduced to quadratic. Overall an asymptotic N^3 scaling behavior is expected and shown due to the formation and transformation of the three-center integrals as well as the transformation of the perturbed two-electron integral.

RI-MP2. This corresponds to the ratio of 7.6 at the double-ζ level. Accordingly, a crossover in analogy to the double-ζ level is to be expected at systems larger than $C_{60}H_{122}$.

The higher computational cost for small system sizes in comparison to the canonical approach can be attributed to various factors. The gradient of a MP2 energy expression needs to be evaluated per Laplace point. The locality of the AOs will not lead to particularly sparse matrices for small systems; thus, the effort per Laplace point is not considerably reduced with respect to the canonical implementation. In addition, QQR-type and Schwarz screening has to be performed. Besides the computational cost of the screening itself, this also results in a more expensive contraction per element, as per contraction a list of significant elements has to be accessed. This overhead is largely outweighed by the fact that for larger systems the number of significant elements is drastically smaller than the number of all elements, but renders canonical formulations more suited for gradient computations of small molecules.

Similar results can be obtained when analyzing the overall performance of the computation of HFCCs in Table VII. An

A crossover to the conventional RI-MP2 energy gradient computation is shown, which, combined with a future parallelization of the existing code and a migration to GPUs, will render large-scale gradient computations feasible. HFCCs were computed with a reduced computational effort and scaling behavior in comparison to the canonical approach, making our RI-CDD ansatz the method of choice for the computation of HFCCs at the MP2-level for medium- to large-sized molecular systems. Furthermore, our RI-CDD ansatz for analytical MP2 gradients forms the basis for efficient second-order molecular properties such as the calculation of NMR chemical shieldings, for which we presented an AO-based formulation.[88] As NMR shieldings can be obtained reliably at the MP2-level[89] and can be further improved by the use of spin-component scaling,[90] we plan to extend the RI-CDD approach to the computation of NMR shieldings in future work.

ACKNOWLEDGMENTS

APPENDIX: CENTERS AND EXTENTS

The description of the centers and extents in the computation of RI-CDD-MP2 energy gradients follows the corresponding work for MP2 energies.[36,40] To determine the extents of transformed charge distributions, the relative weights of the AO contribution need to be computed. For every fully transformed pair, the relative weights of the AO contribution are given by

$$c_{\underline{i}\overline{a},\eta}^{MN} = \frac{|\underline{L}_{M\underline{i}}^{\eta} S_{MN} \overline{L}_{N\overline{a}}^{\eta}|}{\sum_{KL} |\underline{L}_{K\underline{i}}^{\eta} S_{KL} \overline{L}_{L\overline{a}}^{\eta}|}. \tag{A1}$$

Here, K, L, M, and N denote shells, and the corresponding matrix elements are defined as the maximum of all atomic orbitals within the shell,

$$S_{MN} = \max_{\{\mu \in M, \nu \in N\}} |S_{\mu\nu}|, \qquad \underline{L}_{M\underline{i}}^{\eta} = \max_{\{\mu \in M\}} |\underline{L}_{\mu\underline{i}}^{\eta}|. \tag{A2}$$

In the computation of the extents, solely AO contributions with a relative weight exceeding a threshold ϑ_t are used,

$$\text{ext}_{\underline{i}\overline{a}}^{\eta} = \max_{\{MN \,|\, c_{\underline{i}\overline{a},\eta}^{MN}\rangle \vartheta_t\}} \left\{ r_{\underline{i}\overline{a},MN}^{\eta} + c_{\underline{i}\overline{a},\eta}^{MN} \text{ext}_{MN} \right\}. \tag{A3}$$

The centers of the LPMO products are defined as

$$\vec{r}_{\underline{i}\overline{a}}^{\eta} = \frac{\sum_{MN} |\underline{L}_{M\underline{i}}^{\eta} S_{MN} \overline{L}_{N\overline{a}}^{\eta}| \vec{r}_{MN}}{\sum_{MN} |\underline{L}_{M\underline{i}}^{\eta} S_{MN} \overline{L}_{N\overline{a}}^{\eta}|}. \tag{A4}$$

In contrast to the computation of energies, the QQR-type integral screening for RI-CDD-MP2 energy gradients requires the computation of the extents and centers of half-transformed charge distributions. Therefore, the transformation with the

Cholesky decomposed pseudodensity is replaced by a transformation matrix $T_{\mu M}$, connecting the shells to the atomic orbitals.

For pre-screenings as in Eq. (36), extents assigned to a single orbital are required. Based on the respective work by Maurer *et al.*,[40] the following extents are defined:

$$\text{ext}_{\underline{i}}^{\eta} = \max_{\{M,N,\overline{a} \,|\, c_{\underline{i}\overline{a},\eta}^{MN}\rangle \vartheta_t\}} \left\{ r_{\underline{i},MN}^{\eta} + c_{\underline{i}\overline{a},\eta}^{MN} \text{ext}_{MN} \right\},$$

$$\text{ext}_{\overline{a}}^{\eta} = \max_{\{M,N,\underline{i} \,|\, c_{\underline{i}\overline{a},\eta}^{MN}\rangle \vartheta_t\}} \left\{ r_{\overline{a},MN}^{\eta} + c_{\underline{i}\overline{a},\eta}^{MN} \text{ext}_{MN} \right\}. \tag{A5}$$

[1] P. Pulay, Mol. Phys. **17**, 197 (1969).
[2] J. Gauss, "Molecular properties," in *Modern Methods and Algorithms of Quantum Chemistry* (John von Neumann Institute for Computing, Jülich, 2000), pp. 541–592.
[3] J. Čížek, J. Chem. Phys. **45**, 4256 (1966).
[4] J. Čížek, J. Paldus, and L. Šroubková, Int. J. Quantum Chem. **3**, 149 (1969).
[5] J. Čížek and J. Paldus, Int. J. Quantum Chem. **5**, 359 (2007).
[6] R. J. Bartlett and M. Musiał, Rev. Mod. Phys. **79**, 291 (2007).
[7] C. Møller and M. S. Plesset, Phys. Rev. **46**, 618 (1934).
[8] D. Cremer, Wiley Interdiscip. Rev.: Comput. Mol. Sci. **1**, 509 (2011).
[9] P. Pulay, Chem. Phys. Lett. **100**, 151 (1983).
[10] P. Pulay and S. Saebø, Theor. Chim. Acta **69**, 357 (1986).
[11] S. Sæbø and P. Pulay, Chem. Phys. Lett. **113**, 13 (1985).
[12] S. Saebø and P. Pulay, J. Chem. Phys. **86**, 914 (1987).
[13] S. Saebø and P. Pulay, Annu. Rev. Phys. Chem. **44**, 213 (1993).
[14] T. Korona, D. Kats, M. Schütz, T. B. Adler, Y. Liu, and H.-J. Werner, *Local Approximations for an Efficient and Accurate Treatment of Electron Correlation and Electron Excitations in Molecules* (Springer, The Netherlands, 2011), pp. 345–407.
[15] M. Schütz, G. Hetzer, and H.-J. Werner, J. Chem. Phys. **111**, 5691 (1999).
[16] G. Hetzer, M. Schütz, H. Stoll, and H.-J. Werner, J. Chem. Phys. **113**, 9443 (2000).
[17] A. E. Azhary, G. Rauhut, P. Pulay, and H.-J. Werner, J. Chem. Phys. **108**, 5185 (1998).
[18] M. Schütz, H.-J. Werner, R. Lindh, and F. R. Manby, J. Chem. Phys. **121**, 737 (2004).
[19] K. Kitaura, T. Sawai, T. Asada, T. Nakano, and M. Uebayasi, Chem. Phys. Lett. **312**, 319 (1999).
[20] K. Kitaura, E. Ikeo, T. Asada, T. Nakano, and M. Uebayasi, Chem. Phys. Lett. **313**, 701 (1999).
[21] Y. Mochizuki, K. Yamashita, T. Murase, T. Nakano, K. Fukuzawa, K. Takematsu, H. Watanabe, and S. Tanaka, Chem. Phys. Lett. **457**, 396 (2008).
[22] D. G. Fedorov, T. Ishida, M. Uebayasi, and K. Kitaura, J. Phys. Chem. A **111**, 2722 (2007).
[23] T. Ishikawa, N. Yamamoto, and K. Kuwata, Chem. Phys. Lett. **500**, 149 (2010).
[24] T. Ishikawa and K. Kuwata, J. Phys. Chem. Lett. **3**, 375 (2012).
[25] I.-M. Høyvik, K. Kristensen, B. Jansik, and P. Jørgensen, J. Chem. Phys. **136**, 014105 (2012).
[26] K. Kristensen, M. Ziółkowski, B. Jansik, T. Kjærgaard, and P. Jørgensen, J. Chem. Theory Comput. **7**, 1677 (2011).
[27] P. Baudin, P. Ettenhuber, S. Reine, K. Kristensen, and T. Kjærgaard, J. Chem. Phys. **144**, 054102 (2016).
[28] K. Kristensen, T. Kjærgaard, I.-M. Høyvik, P. Ettenhuber, P. Jørgensen, B. Jansik, S. Reine, and J. Jakowski, Mol. Phys. **111**, 1196 (2013).
[29] K. Kristensen, P. Jørgensen, B. Jansik, T. Kjærgaard, and S. Reine, J. Chem. Phys. **137**, 114102 (2012).
[30] D. Bykov, K. Kristensen, and T. Kjærgaard, J. Chem. Phys. **145**, 024106 (2016).
[31] M. Häser, Theor. Chim. Acta **87**, 147 (1993).
[32] M. Häser and J. Almlöf, J. Chem. Phys. **96**, 489 (1992).
[33] J. Almlöf, Chem. Phys. Lett. **181**, 319 (1991).
[34] P. Y. Ayala and G. E. Scuseria, J. Chem. Phys. **110**, 3660 (1999).
[35] S. A. Maurer, D. S. Lambrecht, D. Flaig, and C. Ochsenfeld, J. Chem. Phys. **136**, 144107 (2012).
[36] S. A. Maurer, D. S. Lambrecht, J. Kussmann, and C. Ochsenfeld, J. Chem. Phys. **138**, 014101 (2013).

[37]H. Koch, A. Sánchez de Meras, and T. B. Pedersen, J. Chem. Phys. **118**, 9481 (2003).

[38]F. Aquilante, L. Boman, J. Boström, H. Koch, R. Lindh, A. S. de Merás, and T. B. Pedersen, "Cholesky decomposition techniques in electronic structure theory," in *Linear-Scaling Techniques in Computational Chemistry and Physics: Methods and Applications*, edited by R. Zalesny, M. G. Papadopoulos, P. G. Mezey, and J. Leszczynski (Springer, Netherlands, 2011), pp. 301–343.

[39]J. Boström, M. Pitoňák, F. Aquilante, P. Neogrády, T. B. Pedersen, and R. Lindh, J. Chem. Theory Comput. **8**, 1921 (2012).

[40]S. A. Maurer, L. Clin, and C. Ochsenfeld, J. Chem. Phys. **140**, 224112 (2014).

[41]J. L. Whitten, J. Chem. Phys. **58**, 4496 (1973).

[42]B. I. Dunlap, J. W. D. Connolly, and J. R. Sabin, J. Chem. Phys. **71**, 3396 (1979).

[43]M. Feyereisen, G. Fitzgerald, and A. Komornicki, Chem. Phys. Lett. **208**, 359 (1993).

[44]O. Vahtras, J. Almlöf, and M. Feyereisen, Chem. Phys. Lett. **213**, 514 (1993).

[45]D. E. Bernholdt and R. J. Harrison, Chem. Phys. Lett. **250**, 477 (1996).

[46]F. Weigend and M. Häser, Theor. Chem. Acc. **97**, 331 (1997).

[47]F. Weigend, M. Häser, H. Patzelt, and R. Ahlrichs, Chem. Phys. Lett. **294**, 143 (1998).

[48]S. Schweizer, B. Doser, and C. Ochsenfeld, J. Chem. Phys. **128**, 154101 (2008).

[49]S. Grimme, J. Chem. Phys. **118**, 9095 (2003).

[50]S. Kossmann and F. Neese, J. Phys. Chem. A **114**, 11768 (2010).

[51]S. Grimme, J. Chem. Phys. **124**, 034108 (2006).

[52]A. S. Menon and L. Radom, J. Phys. Chem. A **112**, 13225 (2008).

[53]F. Neese, T. Schwabe, and S. Grimme, J. Chem. Phys. **126**, 124115 (2007).

[54]S. Kossmann, B. Kirchner, and F. Neese, Mol. Phys. **105**, 2049 (2007).

[55]P. R. Surján, Chem. Phys. Lett. **406**, 318 (2005).

[56]M. Frisch, Chem. Phys. **141**, 189 (1990).

[57]C. Ochsenfeld and M. Head-Gordon, Chem. Phys. Lett. **270**, 399 (1997).

[58]C. Moler and C. Van Loan, SIAM Rev. **20**, 801 (1978).

[59]T. C. Fung, Int. J. Numer. Methods Eng. **59**, 1273 (2004).

[60]N. C. Handy and H. F. Schaefer III, J. Phys. Chem. **81**, 5031 (1984).

[61]Y. Yamaguchi, J. D. Goddard, Y. Osamura, and H. F. Schaefer III, *New Dimension to Quantum Chemistry: Analytic Derivative Methods in Ab Initio Molecular Electronic Structure Theory* (Oxford University Press, Inc., Oxford, 1994).

[62]M. Beer and C. Ochsenfeld, J. Chem. Phys. **128**, 221102 (2008).

[63]J. Kussmann and C. Ochsenfeld, J. Chem. Phys. **127**, 054103 (2007).

[64]N. J. Higham, Wiley Interdiscip. Rev.: Comput. Stat. **1**, 251 (2009).

[65]H. Harbrecht, M. Peters, and R. Schneider, Appl. Numer. Math. **62**, 428 (2012).

[66]R. A. DiStasio, Jr., R. P. Steele, Y. M. Rhee, Y. Shao, and M. Head-Gordon, J. Comput. Chem. **28**, 839 (2007).

[67]M. Beer and R. Ahlrichs, J. Comput. Chem. **10**, 104 (1989).

[68]H. Horn, H. Weiß, M. Häser, M. Ehrig, and R. Ahlrichs, J. Comput. Chem. **12**, 1058 (1991).

[69]C. Ochsenfeld, Chem. Phys. Lett. **327**, 216 (2000).

[70]J. Kussmann and C. Ochsenfeld, J. Chem. Phys. **138**, 134114 (2013).

[71]J. Kussmann and C. Ochsenfeld, J. Chem. Theory Comput. **11**, 918 (2015).

[72]Y. Shao, Z. Gan, E. Epifanovsky, A. T. Gilbert, M. Wormit, J. Kussmann, A. W. Lange, A. Behn, J. Deng, X. Feng, D. Ghosh, M. Goldey, P. R. Horn, L. D. Jacobson, I. Kaliman, R. Z. Khaliullin, T. Ku, A. Landau, J. Liu, E. I. Proynov, Y. M. Rhee, R. M. Richard, M. A. Rohrdanz, R. P. Steele, E. J. Sundstrom, H. L. Woodcock III, P. M. Zimmerman, D. Zuev, B. Albrecht, E. Alguire, B. Austin, G. J. O. Beran, Y. A. Bernard, E. Berquist, K. Brandhorst, K. B. Bravaya, S. T. Brown, D. Casanova, C.-M. Chang, Y. Chen, S. H. Chien, K. D. Closser, D. L. Crittenden, M. Diedenhofen, R. A. DiStasio, Jr., H. Do, A. D. Dutoi, R. G. Edgar, S. Fatehi, L. Fusti-Molnar, A. Ghysels, A. Golubeva-Zadorozhnaya, J. Gomes, M. W. Hanson-Heine, P. H. Harbach, A. W. Hauser, E. G. Hohenstein, Z. C. Holden, T.-C. Jagau, H. Ji, B. Kaduk, K. Khistyaev, J. Kim, J. Kim, R. A. King, P. Klunzinger, D. Kosenkov, T. Kowalczyk, C. M. Krauter, K. U. Lao, A. D. Laurent, K. V. Lawler, S. V. Levchenko, C. Y. Lin, F. Liu, E. Livshits, R. C. Lochan, A. Luenser, P. Manohar, S. F. Manzer, S.-P. Mao, N. Mardirossian, A. V. Marenich, S. A. Maurer, N. J. Mayhall, E. Neuscamman, C. M. Oana, R. Olivares-Amaya, D. P. O'Neill, J. A. Parkhill, T. M. Perrine, R. Peverati, A. Prociuk, D. R. Rehn, E. Rosta, N. J. Russ, S. M. Sharada, S. Sharma, D. W. Small, A. Sodt, T. Stein, D. Stck, Y.-C. Su, A. J. Thom, T. Tsuchimochi, V. Vanovschi, L. Vogt, O. Vydrov, T. Wang, M. A. Watson, J. Wenzel, A. White, C. F. Williams, J. Yang, S. Yeganeh, S. R. Yost, Z.-Q. You, I. Y. Zhang, X. Zhang, Y. Zhao, B. R. Brooks, G. K. Chan, D. M. Chipman, C. J. Cramer, W. A. Goddard III, M. S. Gordon, W. J. Hehre, A. Klamt, H. F. Schaefer III, M. W. Schmidt, C. D. Sherrill, D. G. Truhlar, A. Warshel, X. Xu, A. Aspuru-Guzik, R. Baer, A. T. Bell, N. A. Besley, J.-D. Chai, A. Dreuw, B. D. Dunietz, T. R. Furlani, S. R. Gwaltney, C.-P. Hsu, Y. Jung, J. Kong, D. S. Lambrecht, W. Liang, C. Ochsenfeld, V. A. Rassolov, L. V. Slipchenko, J. E. Subotnik, T. V. Voorhis, J. M. Herbert, A. I. Krylov, P. M. W. Gill, and M. Head-Gordon, Mol. Phys. **113**, 184 (2015).

[73]F. Neese, Wiley Interdiscip. Rev.: Comput. Mol. Sci. **2**, 73 (2012).

[74]A. Takatsuka, S. Ten-no, and W. Hackbusch, J. Chem. Phys. **129**, 044112 (2008).

[75]F. Weigend and R. Ahlrichs, Phys. Chem. Chem. Phys. **7**, 3297 (2005).

[76]T. H. Dunning, J. Chem. Phys. **90**, 1007 (1989).

[77]D. E. Woon and T. H. Dunning, J. Chem. Phys. **98**, 1358 (1993).

[78]A. Hellweg, C. Hättig, S. Höfener, and W. Klopper, Theor. Chem. Acc. **117**, 587 (2007).

[79]F. Weigend, A. Köhn, and C. Hättig, J. Chem. Phys. **116**, 3175 (2002).

[80]C. Hättig, Phys. Chem. Chem. Phys. **7**, 59 (2005).

[81]F. G. Gustavson, ACM Trans. Math. Software **4**, 250 (1978).

[82]E. Cuthill and J. McKee, in *Proceedings of the 1969 24th National Conference, ACM 1969* (ACM, New York, NY, USA, 1969), pp. 157–172.

[83]S. Sinnecker, E. Reijerse, F. Neese, and W. Lubitz, J. Am. Chem. Soc. **126**, 3280 (2004).

[84]H. F. Schurkus and C. Ochsenfeld, J. Chem. Phys. **144**, 031101 (2016).

[85]H.-J. Werner, F. R. Manby, and P. J. Knowles, J. Chem. Phys. **118**, 8149 (2003).

[86]M. S. Frank, G. Schmitz, and C. Hättig, Mol. Phys. **115**, 343 (2017).

[87]S. A. Maurer, J. Kussmann, and C. Ochsenfeld, J. Chem. Phys. **141**, 051106 (2014).

[88]M. Maurer and C. Ochsenfeld, J. Chem. Phys. **138**, 174104 (2013).

[89]D. Flaig, M. Maurer, M. Hanni, K. Braunger, L. Kick, M. Thubauville, and C. Ochsenfeld, J. Chem. Theory Comput. **10**, 572 (2014).

[90]M. Maurer and C. Ochsenfeld, J. Chem. Theory Comput. **11**, 37 (2015).

7.2 Paper II: Selected-Nuclei MP2 hyperfine coupling constants

Selected-Nuclei Method for the Computation of Hyperfine Coupling Constants within Second-Order Møller−Plesset Perturbation Theory

Sigurd Vogler,[†] Gökcen Savasci,[†,‡] Martin Ludwig,[†] and Christian Ochsenfeld*[,†,‡]

[†]Chair of Theoretical Chemistry and Center for Integrated Protein Science Munich (CIPSM), Department of Chemistry, University of Munich (LMU), Butenandtstrasse 7, 81377 Munich, Germany

[‡]Max Planck Institute for Solid State Research, Heisenbergstrasse 1, 70569 Stuttgart, Germany
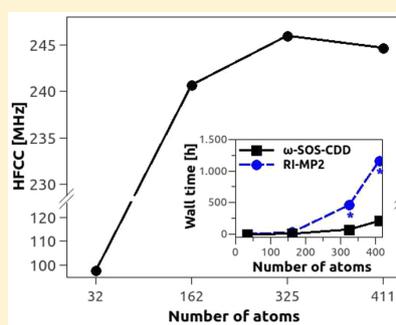
ABSTRACT: We introduce a new ansatz to compute hyperfine coupling constants of selected nuclei at the level of second-order Møller−Plesset perturbation (MP2) and double-hybrid density functional theory with reduced computational effort, opening the route to the analiys of hyperfine coupling constants of large molecular structures. Our approach is based on a reformulation of the canonical MP2 term in atomic orbitals, thus exploiting the locality of electron correlation. We show that a perturbation-including integral screening reduces the scaling behavior of the number of significant two-electron integrals to sublinear. This selected-nuclei approach allows for an efficient computation within scaled-opposite spin (SOS) RI-MP2 on massively parallelized architectures such as graphical processor units (GPUs), thus enabling studies on the influence of the environment on hyperfine coupling constants.

1. INTRODUCTION

Molecules with nonzero spins, such as organic radicals, can be studied by electron paramagnetic resonance (EPR) spectroscopy.[1] One important quantity in EPR spectroscopy is the hyperfine coupling, which can be seen as the analogon to the *J*-coupling in nuclear magnetic resonance (NMR) spectrometry and which is based on the interaction of the electronic spin with the nuclear magnetic moment. Thus, it is directly related to the spin density in the vicinity of the nucleus and can therefore be described in the nonrelativistic limit by the Fermi contact and an anisotropic magnetic dipole−dipole interaction of the nuclear-electronic spins.[2,3]

The theoretical description and computation of hyperfine coupling constants (HFCCs) is a challenging task, especially in regards to accuracy and reliability. For accurate results, both EPR-specific basis sets and electron correlation need to be taken into account.[4−7] Due to the large computational cost and the scaling behavior of canonical electron correlation methods, these are still limited to small molecules. This is especially true for elaborate methods such as coupled-cluster theory,[8,9] multireference configuration interaction,[10] and multireference perturbation theory.[11] Second-order Møller−Plesset perturbation theory (MP2), as the cheapest wave function-based

correlation method, is often a good compromise between computational cost and accuracy. On the other hand, density functional theory (DFT) often already provides reliable results[12,13] that can be improved upon by using the restricted-unrestricted approach[14,15] or by combining with perturbation theory in double-hybrid DFT (DH-DFT),[16] which was shown to improve the reliability.[17] The latter includes a term analogous to the unrestricted MP2 equations, which is the focus in our present work. It has to be noted that although an unrestricted formulation includes spin polarization and delocalization, the resulting spin contamination introduces an error that may lead to meaningless results in severe cases. This is especially true for transition metal compounds.[18,19]

We recently introduced an efficient method for computing HFCCs using MP2,[20] where we exploit the locality of atomic orbitals (AOs) and reduce the computational overhead by a Cholesky decomposition of the required (pseudo)densities (CDD)[21,22] and by the resolution-of-the-identity (RI) approximation.[23−28] Significant two-electron integral contributions are selected based on a QQR-type integral screening.[29,30] This

enables the computation of larger molecular systems than previously accessible.

In this work, we push the limit of computable molecular systems further. For this, we exploit the fact that the spin density of large molecular systems is often highly localized, especially for systems that are well described by single-reference methods. Thus, only a small number of nuclei will have a HFCC deviating considerably from zero. For systems with localized spin densities, it can be sufficient to compute the HFCCs of these selected nuclei by exploiting the locality of the perturbation. In addition, this opens the route to study the effect of solvation and the environment on the HFCCs, e.g., of spin labels in the study of distances in large molecular systems.[31−34] Similarly, the effect of long-range spin polarization on the HFCCs of specific nuclei can be computed. Studies on the influence of the protein environment on flavoprotein radicals[35] and on plastocyanin[36] have previously been conducted using a hybrid quantum mechanical and molecular mechanical (QM/MM) approach at the DFT-level. The influence of the solvent on the hyperfine coupling has for example been studied theoretically on methyl and ClO_2 radicals.[37]

In order to enable such selected-nuclei studies, we need to explicitly compute the response (pseudo)density matrix with respect to the perturbation. When the perturbation is local, this results in a sublinear scaling number of significant two-electron integrals in an AO-based formulation using a perturbation-including integral screening. The locality can be further exploited in the transformation of the three-center quantities in an RI-based algorithm when only the opposite-spin term in MP2 is computed and an attenuated Coulomb-metric is employed.[38−40] So-called scaled opposite-spin (SOS) MP2 was introduced by Jung et al. for the computation of correlation energies[41] and was shown to provide accurate results in the computation of the energy at reduced computational cost. To our knowledge, SOS-MP2 has not been studied in the context of HFCCs, whereas the promising results of SOS-MP2 for the computation of NMR shieldings[42] motivate the use of the SOS-ansatz to compute HFCCs as well. We thus introduce a new attenuated Coulomb-metric SOS-MP2 ansatz to compute the HFCCs of selected nuclei and present their computation employing an extended version of our adapted J-engine-based ansatz[43] on massively parallel architectures such as GPUs.

With our new methodoloy, we furthermore analyze the effect of the environment on the HFCCs in a typical molecular system, namely that of a glycyl radical enzyme, by increasing the size of the protein environment surrounding the catalytic site.

2. THEORY

At the nonrelativistic level, i.e., in the absence of spin−orbit coupling, the HFCC can be computed as first derivative of the energy with respect to the nuclear magnetic moment M_k of nucleus k.[44] This perturbation is separable into the isotropic Fermi-contact term and the anisotropic spin-dipole interaction.

At the MP2-level, the unrestricted MP2 equation needs to be perturbed. To achieve low scaling and to exploit the locality of electron correlation, the use of either local molecular orbitals (MOs) or AOs is beneficial. The latter requires a Laplace transformation of the energy denominator, as introduced by Almlöf and Häser.[45−47] In a Laplace-transform AO-based ansatz, computational overhead and the large dependency on the basis set size can be reduced by the RI approximation[23−28]

and a Cholesky decomposition,[21,22] resulting in an efficient ansatz to compute the HFCCs of all nuclei.[20] The MP2 energy perturbed with respect to the nuclear magnetic moment M_k of nucleus k using a generalized Laplace range[20] reads as

$$
\begin{aligned}
\frac{\partial E_{\text{AO-MP2}}}{\partial M_k} &= -\frac{1}{2} \sum_{\zeta}^{\tau} w_\zeta \sum_{\eta}^{\alpha,\beta} \frac{\partial \mathcal{E}^\eta(\zeta)}{\partial M_k} \\
&= -\frac{1}{2} \sum_{\zeta}^{\tau} w_\zeta \sum_{\eta}^{\alpha,\beta} \left[2 \sum_{\mu'\mu} \bar{R}^\eta_{\mu'\mu}(\zeta) \frac{\partial \underline{P}^\eta_{\mu'\mu}}{\partial M_k} \right. \\
&\quad \left. + 2 \sum_{\nu'\nu} \underline{R}^\eta_{\nu'\nu}(\zeta) \frac{\partial \bar{P}^\eta_{\nu'\nu}}{\partial M_k} \right]
\end{aligned}
\tag{1}
$$

where τ is the number of Laplace expansion points ζ with corresponding linear expansion coefficient w_ζ and the exponent t_ζ. $\bar{P}$ and $\underline{P}$ are local pseudodensities (PDs)

$$
\underline{P}^\eta_{\mu\mu'} = \sum_i^{\text{occ}} c^\eta_{\mu i} e^{\varepsilon^\eta_i t_\zeta} c^\eta_{\mu'i}
$$

$$
\bar{P}^\eta_{\nu\nu'} = \sum_a^{\text{virt}} c^\eta_{\nu a} e^{-\varepsilon^\eta_a t_\zeta} c^\eta_{\nu'a}
\tag{2}
$$

with the molecular orbital coefficients $c^\eta_{\mu i}$ and the orbital energies ε^η_i. The **R**-matrices in eq 1 are defined as

$$
\begin{aligned}
\bar{R}^\eta_{\mu'\mu}(\zeta) &= \sum_j^{\text{occ}} \sum_{a,b}^{\text{virt}} (\mu'\bar{a}|\underline{j}\,\bar{b})_{\eta\eta}(\mu\bar{a}\|\underline{j}\,\bar{b})_{\eta\eta} \\
&\quad + \sum_j^{\text{occ}} \sum_{a,b}^{\text{virt}} (\mu'\bar{a}|\underline{j}\,\bar{b})_{\eta\eta'}(\mu\bar{a}|\underline{j}\,\bar{b})_{\eta\eta'}
\end{aligned}
\tag{3}
$$

$$
\begin{aligned}
\underline{R}^\eta_{\nu'\nu}(\zeta) &= \sum_{i,j}^{\text{occ}} \sum_b^{\text{virt}} (\underline{i}\nu'|\underline{j}\,\bar{b})_{\eta\eta}(\underline{i}\nu\|\underline{j}\,\bar{b})_{\eta\eta} \\
&\quad + \sum_{i,j}^{\text{occ}} \sum_b^{\text{virt}} (\underline{i}\nu'|\underline{j}\,\bar{b})_{\eta\eta'}(\underline{i}\nu|\underline{j}\,\bar{b})_{\eta\eta'}
\end{aligned}
\tag{4}
$$

where $\eta' \neq \eta$.

Here, the two-electron integrals are transformed with local Cholesky orbital matrices $\underline{L}$ and $\bar{L}$, obtained by a Cholesky decomposition of the PD matrices:

$$
(\underline{i}\bar{a}|\underline{j}\,\bar{b})_{\eta\eta'} = \sum_{\mu\nu\lambda\sigma}^{N} \underline{L}^\eta_{\mu\underline{i}} \bar{L}^\eta_{\nu\bar{a}}(\mu\nu|\lambda\sigma) \underline{L}^{\eta'}_{\lambda\underline{j}} \bar{L}^{\eta'}_{\sigma\bar{b}}
\tag{5}
$$

The local Cholesky orbitals inherit the locality of the AOs, while the number of occupied and virtual local Cholesky orbitals is equal to or smaller than the size of the canonical MO space.[48] Thus, the dependency on the basis set size is reduced.

For an efficient implementation for all nuclei in the molecular system, the contraction of the **R**-matrices with the perturbed PDs in eq 1 is performed for all perturbations simultaneously using the Z-vector method[49−51] in a density matrix-based Laplace-transform unrestricted coupled-perturbed self-consistent field (DL-UCPSCF) algorithm.[20,52]

Furthermore, the prefactor in the scaling behavior is reduced by the RI-approximation. The two-electron integral is hereby described as a contraction of two three-center quantities:

$$(\underline{i}\,\overline{a}|j\,\overline{b})_{\eta\eta'} = \sum_P^{N_{aux}} B^{\eta}_{\underline{i}\,\overline{a},P} B^{\eta'}_{j\,\overline{b},P}$$

$$\text{with: } B^{\eta}_{\underline{i}\,\overline{a},P} = \sum_Q^{N_{aux}} (\underline{i}\,\overline{a}|Q)_{\eta}(Q|P)^{-1/2} \tag{6}$$

As shown in our previous work, an effectively quadratic scaling computation of all nuclei can be achieved using this approach.[20] The number of integrals needed for the computation, based on a QQR-type integral screening,[29,30] is linear. Due to the auxiliary space of the RI approximation, this results in a quadratic scaling behavior.

For our new ansatz, we aim to compute only the HFCCs of selected nuclei while exploiting the locality of the perturbation. Herefore, the contraction of the perturbed PD matrix with the respective **R**-matrix is performed directly, resulting in the following contribution per Laplace point ζ:

$$\sum_{\eta}^{\alpha,\beta} \frac{\partial \mathcal{E}^{\eta}(\zeta)}{\partial M_k} = \sum_{\eta}^{\alpha,\beta} \underline{A}_{\eta}(\zeta) + \overline{A}_{\eta}(\zeta) \tag{7}$$

with

$$\underline{A}_{\eta}(\zeta) = \sum_{\nu}^{N} \sum_{i,j}^{occ} \sum_{b}^{virt} (\underline{i}\,\overline{\nu}^{M_k}|j\,\overline{b})_{\eta\eta}(\underline{i}\,\nu||j\,\overline{b})_{\eta\eta}$$
$$+ \sum_{\nu}^{N} \sum_{i,j}^{occ} \sum_{b}^{virt} (\underline{i}\,\overline{\nu}^{M_k}|j\,\overline{b})_{\eta\eta'}(\underline{i}\,\nu|j\,\overline{b})_{\eta\eta'} \tag{8}$$

and

$$\overline{A}_{\eta}(\zeta) = \sum_{\mu}^{N} \sum_{j}^{occ} \sum_{a,b}^{virt} (\underline{\mu}^{M_k}\,\overline{a}|j\,\overline{b})_{\eta\eta}(\mu\overline{a}||j\,\overline{b})_{\eta\eta}$$
$$+ \sum_{\mu}^{N} \sum_{j}^{occ} \sum_{a,b}^{virt} (\underline{\mu}^{M_k}\,\overline{a}|j\,\overline{b})_{\eta\eta'}(\mu\overline{a}|j\,\overline{b})_{\eta\eta'} \tag{9}$$

where the two-electron integrals are transformed with the perturbed PDs as, e.g.

$$(\underline{\mu}^{M_k}\,\overline{a}|j\,\overline{b})_{\eta\eta'} = \sum_{\mu'} P^{\eta,M_k}_{\mu\mu'}(\zeta)(\mu'\overline{a}|j\,\overline{b})_{\eta\eta'} \tag{10}$$

The perturbed PDs can be obtained as outlined in eqs 15 and 16 in ref 53. Provided that the perturbation is local, i.e., that the perturbed PDs are local, the number of significant integrals in both $\underline{A}$ and $\overline{A}$ is constant with increasing size of the molecule. This can be exploited by a QQR-type integral estimation which includes the perturbation in the screening procedure. Based on the distance-including QQR-type integral estimates by Maurer et al.,[29,30] the perturbation-including estimate of the Coulomb-type integral contraction in eq 8 reads as

$$(\underline{i}\,\overline{\nu}^{M_k}|j\,\overline{b})_{\eta\eta'}(\underline{i}\,\nu|j\,\overline{b})_{\eta\eta'}$$
$$\approx \frac{Z^{\eta}_{\underline{i}\,\overline{\nu}^{M_k}} Z^{\eta'}_{j\,\overline{b}}}{(R - \text{ext}^{\eta}_{\underline{i}\,\overline{\nu}^{M_k}} - \text{ext}^{\eta'}_{j\,\overline{b}})^2} \frac{X^{\eta}_{\underline{i}\,\nu} Z^{\eta'}_{j\,\overline{b}}}{(R - \text{ext}^{\eta}_{\underline{i}\,\nu} - \text{ext}^{\eta'}_{j\,\overline{b}})^2} \tag{11}$$

Here, Häser's pseudo-Schwarz matrices are employed,[47] with

$$X^{\eta}_{\underline{i}\,\nu} = (\underline{i}\,\nu|\underline{i}\,\nu)^{1/2}_{\eta\eta}$$

$$Y^{\eta}_{\mu\overline{a}} = (\mu\overline{a}|\mu\overline{a})^{1/2}_{\eta\eta}$$

$$Z^{\eta}_{\underline{i}\,\overline{a}} = (\underline{i}\,\overline{a}|\underline{i}\,\overline{a})^{1/2}_{\eta\eta}$$

$$Z^{\eta}_{\underline{i}\,\overline{\nu}^{M_k}} = (\underline{i}\,\overline{\nu}^{M_k}|\underline{i}\,\overline{\nu}^{M_k})^{1/2}_{\eta\eta} \tag{12}$$

and the distance between the two charge distributions is determined taking into consideration their extents (ext). The centers and extents are computed as described in refs 20, 30, and 48. The exponent in the denominator is based on a multipole expansion and on the orthogonality of the virtual and the occupied subspace of the unperturbed PDs.

In analogy, the Coulomb-type integral contraction in eq 9 leads to

$$(\underline{\mu}^{M_k}\,\overline{a}|j\,\overline{b})_{\eta\eta'}(\mu\overline{a}|j\,\overline{b})_{\eta\eta'}$$
$$\approx \frac{Z^{\eta}_{\underline{\mu}^{M_k}\,\overline{a}} Z^{\eta'}_{j\,\overline{b}}}{(R - \text{ext}^{\eta}_{\underline{\mu}^{M_k}\,\overline{a}} - \text{ext}^{\eta'}_{j\,\overline{b}})^2} \frac{Y^{\eta}_{\mu\overline{a}} Z^{\eta'}_{j\,\overline{b}}}{(R - \text{ext}^{\eta}_{\mu\overline{a}} - \text{ext}^{\eta'}_{j\,\overline{b}})^2} \tag{13}$$

The additional summation over the auxiliary space in eq 6 will result in an overall asymptotic linear scaling behavior of the contraction. The formation and transformation of the three-center quantities in eq 6 are only linear-scaling when only integrals significant in the final contraction are computed. This requires prior screening and bookkeeping of significant two-electron integrals, which is computationally demanding and memory intensive.

In the Coulomb-type contraction of the two-electron integrals, this can be circumvented by the use of an attenuated Coulomb-metric.[38−40,54] The idea of scaling the Coulomb-type opposite spin contribution and neglecting the same spin contribution that includes both Coulomb- and exchange-type contractions was shown to be an economical but still reliable ansatz in the realm of correlation energies.[41] This was similarly shown to be beneficial in the computation of NMR shieldings[42] which motivates the computation of SOS-MP2 HFCCs. Only the opposite spin-terms are computed in SOS-MP2 approaches and scaled accordingly with c_{os}. In analogy we propose the computation of erfc-attenuated Coulomb-metric SOS-MP2 HFCCs by

$$\frac{\partial E_{AO-MP2}}{\partial M_k} = -c_{os} \sum_{\zeta}^{\tau} \sum_{\eta}^{\alpha,\beta} (\overline{\mathbf{Z}}^{\omega}_{\eta}(\zeta)\mathbf{C}^{\omega}\mathbf{Z}^{\omega}_{\eta'}(\zeta)\mathbf{C}^{\omega}$$
$$+ \underline{\mathbf{Z}}^{\omega}_{\eta}(\zeta)\mathbf{C}^{\omega}\mathbf{Z}^{\omega}_{\eta'}(\zeta)\mathbf{C}^{\omega}) \tag{14}$$

with

$$(C^{\omega})_{PQ} = \sum_{P'Q'}^{N_{aux}} (P|P')^{-1}_{\omega}(P'|Q')(Q'|Q)^{-1}_{\omega} \tag{15}$$

where

$$(P|Q)_{\omega} = \left(P\left|\frac{\text{erfc}(\omega r_{12})}{r_{12}}\right|Q\right) \tag{16}$$

Hereby, the attenuation of the Coulomb-metric is controlled by ω, for which in work by Luenser et al.[54] a value of 0.1 was shown to lead to no loss in accuracy with respect to the full Coulomb-metric while still providing the computational

efficiency of the local overlap metric. We confirmed this behavior for a small selection of test cases.

The **Z**-matrices in eq 14 are defined as

$$(\underline{Z}_\eta^\omega(\zeta))_{PP'} = \sum_\mu^N \sum_a^{virt} (\underline{\mu}^{M_k}\overline{a}|P)_\omega^{\zeta,\eta}(\mu\overline{a}|P')_\omega^{\zeta,\eta}$$

$$(\overline{Z}_\eta^\omega(\zeta))_{PP'} = \sum_i^{occ} \sum_\nu^N (\underline{i}\overline{\nu}^{M_k}|P)_\omega^{\zeta,\eta}(\underline{i}\nu|P')_\omega^{\zeta,\eta}$$

$$(Z_\eta^\omega(\zeta))_{PP'} = \sum_j^{occ} \sum_b^{virt} (\underline{j}\overline{b}|P)_\omega^{\zeta,\eta}(\underline{j}\overline{b}|P')_\omega^{\zeta,\eta} \tag{17}$$

where the overline or underline indicate the transformation with the perturbed PD (virtual or occupied respectively) and the subscript ω indicates the erfc-metric as in eq 16.

To reduce the computational effort in the formation of the **Z**-matrices and to efficiently exploit the sparsity using the block-compressed sparse row (BCSR) format,[55,56] we follow the ansatz by Maurer et al.[43] and formulate the transformed three-center quantities as

$$\overline{\underline{\mathcal{B}}}^P(\omega) = \underline{\mathbf{L}}\underline{\mathbf{L}}^T\mathcal{B}^P(\omega)\overline{\mathbf{P}}$$

$$\overline{\underline{\mathcal{B}}}^P_{\overline{M}_k}(\omega) = \underline{\mathbf{L}}\underline{\mathbf{L}}^T\mathcal{B}^P(\omega)\overline{\mathbf{P}}^{M_k}$$

$$\overline{\underline{\mathcal{B}}}^P_{\underline{M}_k}(\omega) = \underline{\mathbf{P}}^{M_k}\mathcal{B}^P(\omega)\overline{\mathbf{P}} \tag{18}$$

with

$$\mathcal{B}_{\mu\nu}^P(\omega) = (\mu\nu|P)_\omega \tag{19}$$

We thus omit the Cholesky decomposition of the virtual PD and exploit the sparsity of significant basis function pairs. The occupied perturbed PD cannot be Cholesky decomposed as it is skew-symmetric and not positive semidefinite by definition as the PD matrices. The drastic locality of the perturbed PDs renders the transformation of the three-center integral with the perturbed PDs computationally efficient even without exploiting their rank deficiency. This leads to the following **Z**-matrices:

$$(\underline{Z}_\eta^\omega(\zeta))_{PP'} = \overline{\underline{\mathcal{B}}}^P_{\underline{M}_k}(\omega)\mathcal{B}^{P'}(\omega)$$

$$(\overline{Z}_\eta^\omega(\zeta))_{PP'} = \overline{\underline{\mathcal{B}}}^P_{\overline{M}_k}(\omega)\mathcal{B}^{P'}(\omega)$$

$$(Z_\eta^\omega(\zeta))_{PP'} = \overline{\underline{\mathcal{B}}}^P(\omega)\mathcal{B}^{P'}(\omega) \tag{20}$$

It has to be noted, that this formulation does not rely on a screening of significant integral products but directly exploits the sparsity via the BCSR format.[55,56] This can also be extended to the use on GPUs by employing the J-engine as outlined in ref 43.

Furthermore, the presented MP2 approaches can straightforwardly be extended to DH-DFT:[16] For this approach the perturbative contribution according to Møller–Plesset to the second order is constructed by unrestricted Kohn–Sham orbitals instead of unresticed Hartree–Fock orbitals. This has been shown to be especially advantageous in the computation of open-shell properties, as it is less susceptible to spin contamination, as shown in work on bond dissociation energies.[57] Reliable HFCCs can be computed with the use of analytical gradients of the DH-DFT energy.[17,58]

3. COMPUTATIONAL DETAILS

The algorithms to compute the HFCCs at the MP2-level using the RI approximation and CDD are implemented in the program package FermiONs++.[59,60] The Laplace expansion coefficients are selected based on the minimax-approximation.[61] The extents of the QQR-type integral estimates are determined with the same thresholds as in ref 30, i.e., the threshold of the untransformed AO extent is 0.1 and the threshold of the transformed extents is 10^{-3}. The QQR-screening threshold was set to 10^{-8}, and seven Laplace expansion points were chosen, based on the study of the accuracy in ref 20. The DL-UCPSCF was converged to a threshold of 10^{-4} in the case of the linear alkane systems and to 10^{-2} in the case of the PFL systems. A ω-value of 0.1 was chosen for the attenuated Coulomb-metric throughout. Deviations of less than 1 MHz can be expected with these thresholds.[20] The basis set def2-SVP and its respective auxiliary basis set has been used throughout.[28,62] The SOS-computations are scaled with $c_{OS} = 1.3$, as suggested for SOS-MP2 energy computations,[41] except if stated otherwise. Reference canonical RI-MP2 computations were performed with the program package ORCA.[63] Values that could not be obtained with the canonical method or the all-nuclei ansatz were extrapolated conservatively based on the scaling behavior of the next smaller computable system.

4. RESULTS

4.1. Locality of the Perturbation. The idea of the present ansatz to compute HFCCs of selected nuclei is based on the locality of the perturbation, i.e, of the Fermi-Contact term and of the anisotropic contribution. The locality of the Fermi-Contact operator for nuclear spin–spin coupling at the DFT-level has been shown by Luenser et al.[64] Two quantities can be analyzed to support this approach for open-shell systems at the MP2-level: (i) the sparsity of the perturbed PD matrices and (ii) the number of significant integrals in the computation of $\underline{A}$ and $\overline{A}$ in eqs 8 and 9.

Figure 1 shows the PDs perturbed with respect to the Fermi-Contact term and to the anisotropic spin-dipole interaction for $C_{100}H_{201}$ obtained by DL-UCPSCF. In our case, $C_{100}H_{201}$ has the radical center at the C_1 position and the perturbed PD of α spin with respect to this atom is shown. It clearly highlights the locality for both the occupied and the virtual perturbed PD. Figure 2 shows the number of significant integrals for the computation of the isotropic HFCC at the MP2-level of the C_1 atom in linear alkanes based on a perturbation-including QQR-type integral estimation as in eqs 11 and 13. The $O(1)$ scaling behavior of the integrals with the system size can clearly be seen.

4.2. Selected-Nuclei HFCCs. To investigate the computational performance, we calculated the isotropic HFCCs at the MP2-level of linear alkanes where the radical center was located at the C_1 position. Linear alkanes in general are well-suited to demonstrate the asymptotic scaling behavior and additionally show locality of the spin density which makes them an ideal test case for a selected-nuclei approach.

The isotropic HFCCs of the linear alkane chains were computed using the all-nuclei approach with the canonical implementation in ORCA,[63] our RI-CDD variant,[20] and the new selected-nuclei ansatz with the Coulomb-metric RI as in eq 7 on one central processing unit (CPU) node using 12 cores. A perturbation-including QQR-type integral screening as in eqs 11 and 13 was performed in the contraction in eqs 8 and 9. To

(a) Fermi-Contact term (occupied)

(b) Fermi-Contact term (virtual)

(c) Anistropic spin-dipole (occupied)

(d) Anistropic spin-dipole (virtual)
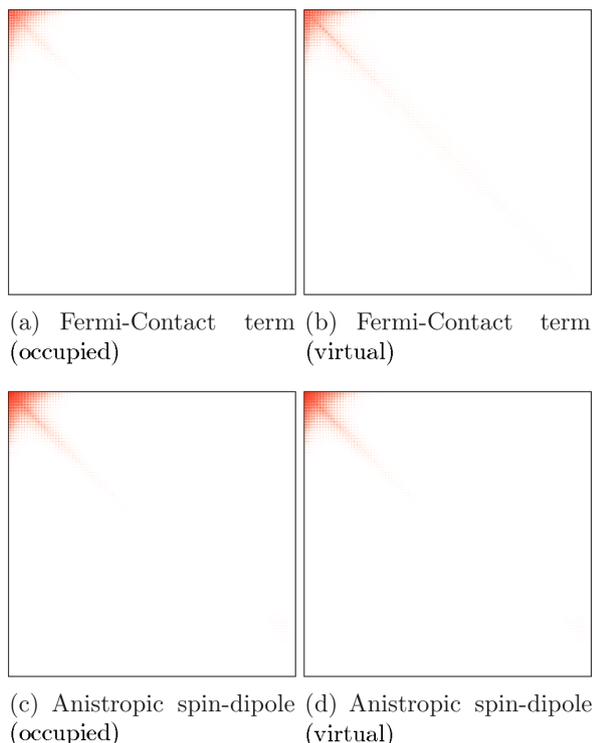
Figure 1. Sparsity patterns for the perturbed occupied and virtual PDs of α spin with respect to the Fermi-Contact term in (a) and (b) and to the anisotropic spin-dipole interaction in (c) and (d) obtained from DL-UCPSCF for $C_{100}H_{201}$. The absolute values of the different spatial contributions to the anisotropic term are summed. The sparsity patterns are taken from the first Laplace expansion point, and elements smaller than 10^{-5} are discarded (white).
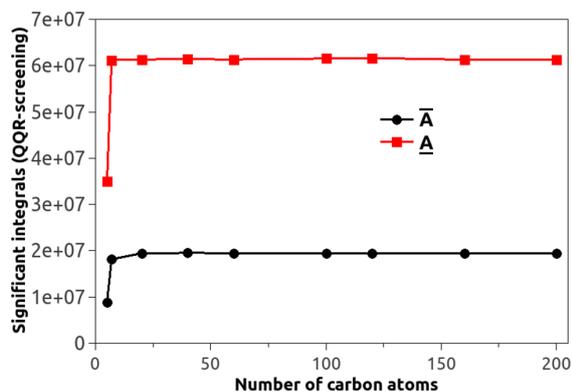
Figure 2. Number of significant integrals in the computation of $\overline{A}$ and $\underline{A}$ in eqs 8 and 9 for linear alkanes with increasing chain lengths using a perturbation-including QQR-type integral threshold of 10^{-8}. Numbers are shown for the first Laplace point.

reduce the necessary bookkeeping we omitted the external QQR-type integral screening to determine significant transformed three-center quantities. Instead, a shell-pair-based screening is used to compute only significant untransformed three-center integrals, and then the sparsity and locality of the PDs in the transformation procedure are exploited.

The wall times are shown in Figure 3 and clearly highlight a reduction in the computational cost for large systems with
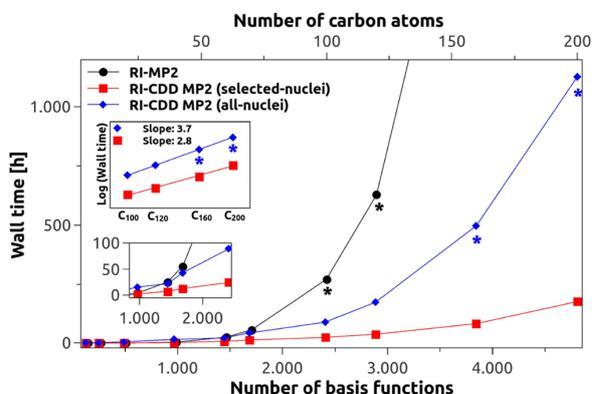
Figure 3. Computational wall time for the calculation of isotropic HFCCs using the all-nuclei RI- and RI-CDD MP2 ansatz and the selected-nuclei RI-CDD MP2 method on a CPU node using 12 cores. Values indicated by an asterisk were conservatively extrapolated. The upper inset shows a double-logarithmic plot for the four largest systems and a linear fit to obtain the scaling behavior at those system sizes. The lower inset shows a detailed view of the crossover of the canonical method to the all-nuclei computation.

respect to the canonical implementation for both AO-based approaches. Furthermore, the scaling behavior for the four largest systems is obtained by a linear fit of the double-logarithmic plot. It clearly shows a reduced scaling behavior for the selected-nuclei approach. The overall nonquadratic scaling behavior can be attributed to expensive I/O operations. Every selected-nuclei computation was faster than the all-nuclei ansatz. For $C_{20}H_{41}$ this amounts to a 10-fold decrease in wall time. We show the MP2 contribution to the HFCC of $C_{20}H_{41}$ in Figure 4 for a detailed analysis. The total HFCCs and the MP2 contribution decay fast with the distance to the C_1 position. Only eight nuclei in total have a MP2 contribution to the HFCC larger than 5 MHz, with the largest contribution being from the selected C_1 nucleus. Thus, if one is, for example, only interested in HFCCs larger than 5 MHz, the resulting computational cost using the selected-nuclei approach would still be smaller than computing the HFCCs of all nuclei. The speed-up will be even more significant for nuclei with small MP2 contributions, as the smaller MP2 contributions also lead to less significant integrals and thus a faster computation than that of the C_1 nucleus.

For an intermediate size of molecules larger than $C_{20}H_{41}$, the computation of the selected-nuclei HFCCs is dominated by the transformation of the three-center **B**-matrices so that the speed-up does not increase. As we do not perform an external screening, the fully transformed matrix $B_{i\bar{a},P}$ needs to be computed irrespective of the perturbation. However, for larger systems, the computational cost of the transformation in the case of the all-nuclei approach is dominated by the formation of the half-transformed $B_{\mu\bar{a},P}$, which is not necessary in the case of the selected-nuclei ansatz. Thus, as can be seen in Figure 3, the speed-up increases again for very large systems.

4.3. SOS-MP2 HFCCs. *4.3.1. Accuracy.* SOS-MP2 is an economical and often accurate variant to obtain the correlation contributions and allows for the use of an attenuated Coulomb-metric in RI-based methods. This can be exploited in the

(a) Isotropic HFCCs of carbon atoms

(b) Isotropic HFCCs of hydrogen atoms

Figure 4. Isotropic HFCCs at the HF- and the RI-MP2-level of the a) carbon and the b) hydrogen atoms with increasing distance to the radical center of $C_{20}H_{41}$.

Table 1. Average Deviation ($\overline{\Delta}$) of Isotropic SOS-MP2 HFCCs with c_{OS} = 1.0 and c_{OS} = 1.3 to RI-MP2/DH-DFT HFCCs (PT2) and to SCS-RI-MP2/DH-DFT for NF_2, CH_3, CHO, $^3O_2$, and ^3SO for Different Basis Sets[a]

	RI-MP2 (def2-SVP)	RI-MP2 (cc-pVTZ)	B2-PLYP (cc-pVTZ)
$\overline{\Delta}(SOS_{c_{OS} = 1.0} - PT2)$	23.6 MHz (70%)	16.4 MHz (64%)	6.0 MHz (62%)
$\overline{\Delta}(SOS_{c_{OS} = 1.3} - PT2)$	16.5 MHz (77%)	11.6 MHz (72%)	5.0 MHz (66%)
$\overline{\Delta}(SCS - PT2)$	11.0 MHz	7.7 MHz	3.3 MHz
$\overline{\Delta}(SCS - SOS_{c_{OS} = 1.3})$	5.5 MHz (92%)	3.9 MHz (91%)	1.7 MHz (89%)

[a]The percentage of the recovery of the RI-MP2/DH-DFT and the SCS-RI-MP2/DH-DFT result using the SOS-approximation is shown in brackets.

beneficial, as, e.g., in the case of NMR shieldings.[42] Further extensive benchmarks with respect to higher-level results and the according adjustment of the scaling parameter c_{OS} for different elements might even render the SOS-HFCCs more reliable than nonscaled results.

4.3.2. Computational Efficiency. We computed the SOS-MP2-HFCCs of linear alkanes using our new approach with and without the attenuated Coulomb-metric, indicated by ω, both on CPUs and on GPUs. The GPU computations were performed using the modified J-engine,[43] whereas the Z-matrices were obtained using BCSR-algebra in the case of the CPU computations. For efficiency reasons, BCSR was only used for ω-SOS-CDD computations and only for systems larger than $C_{60}H_{121}$. In the GPU algorithm, only the contraction in eq 20 is performed on GPUs. The timings are shown in Figure 5.

Figure 5. Comparison of the computational wall time for the selected-nuclei calculation of isotropic HFCCs of the C_1 atom of linear alkanes using the RI-CDD MP2 ansatz and the SOS-CDD approaches both on CPUs and on GPUs with and without the attenuated Coulomb-metric indicated by an ω. The scaling behavior for the four largest systems is obtained by a linear fit of the double-logarithmic plot shown in the inset.

computation of the HFCCs of selected nuclei. Furthermore, a migration to GPUs using the efficient modified J-engine algorithm as presented in ref 43 can be achieved. Scaling the opposite spin contribution with respect to more accurate methods using extended basis sets, as was previously for example performed for NMR shieldings,[42] can lead to highly accurate results. In our present work, we do not aim to correct the error due to an incomplete basis set, nor the method error by adjusting the scaling parameter with respect to accurate reference results, as obtained, e.g., by coupled cluster theory. Instead we aim to highlight that the assumption of neglecting the same spin contribution and correspondingly weighting the opposite spin terms is a useful approach in the case of HFCCs.

We computed the isotropic HFCCs for a variety of molecules with both RI-MP2 and DH-DFT and compared the standard unscaled RI-MP2/DH-DFT HFCCs to their SOS variant in Table 1. Most of the correlation can be attributed to the opposite spin (OS) contribution, thus the average deviation of the SOS-computation is small. When the OS contribution is scaled with c_{OS} = 1.3, the value determined for the computation of SOS-MP2 energies,[41] these deviations decrease considerably. Further, when comparing to spin component scaled (SCS) RI-MP2/DH-DFT HFCCs, where both the opposite and the same spin contributions are scaled,[65] the values indicate that SOS-MP2 HFCCs are a good approximate to SCS results, notably better than unscaled RI-MP2/DH-DFT. The same spin contributions are especially important when long-range effects are considered, which is not the case for the local HFCCs.

Despite the fact, that the deviations are small, an element-specific and basis set-dependent scaling factor might be

As expected, the GPU computations outperform those on CPUs for all systems. Only for even more extended systems will the ω-SOS-CDD MP2 implementation on CPUs be more efficient than their GPU variant. This is due to the formation and the contraction of the Z-matrices, which asymptotically scale sublinearly in the CPU code by the use of BCSR algebra. In contrast, the highly efficient J-engine used in the GPU

implementations does not rely on BCSR and scales cubically. The linear fits of the double-logarithmic plot in Figure 5 confirm the reduced scaling behavior of the ω-CPU variant with respect to the ω-GPU implementation. The reported scaling behaviors are not (sub)linear, which is due to computation of all three-center quantities without memory intensive prior screening and bookkeeping and due to expensive I/O-processes.

When comparing to the selected-nuclei RI-CDD ansatz, the SOS-computations on CPUs perform less efficiently than the computation of both the opposite and same spin contributions. While at first this seems contradictory, the reason is that the RI-CDD ansatz is based on the highly efficient perturbation-including QQR-type integral screening, whereas our present SOS-code only resorts to BCSR algebra. The latter is more efficient for more extended systems, as can be seen for $C_{200}H_{401}$ in Figure 5.

4.4. Environment Effects on HFCCs. In order to demonstrate the overall performance and applicability of our new method, we investigated HFCCs of catalytic residues found in pyruvate formate lyase (PFL). This glycyl radical enzyme catalyzes the reversible transformation of pyruvate and coenzyme-A into formate and acetyl-CoA involving three residues:[66,67] The Cys418 thiyl radical acylates the carbon atom of the pyruvate carbonyl, the Cys419 thiyl radical administrates hydrogen-atom transfers, and the Gly734 glycyl radical transfers the radical to and from Cys418 using Cys419.

We analyze the effect of the protein environment on the radical center in PFL based on the crystal structure[68] (PDB code 2PFL) by computing HFCCs of the radical centers taking more and more surrounding residues of the protein into account.

The crystal structure depicts the structure of the nonradical form of PFL in its complex with the substrate analog oxamate and the C418A/C419A double mutant. Gly734 and Cys419 positioned at the tips of opposing hairpin loops meet in the apolar barrel center, whereas oxamate fits into a compact pocket where C2 is juxtaposed with Cys418, which in turn is close to Cys419.[68] The model of the active site is suggestive of a snapshot of the catalytic cycle when the pyruvate-carbonyl awaits attack by the Cys418 thiyl radical proposed as a homolytic radical mechanism for PFL that involves Cys418 and Cys419 both as thiyl radicals with distinct chemical functions.

Starting from the X-ray crystal structure, force field MD minimizations were performed with the NAMD 2.10 package[69] and the AMBER 99SB force field.[70] Parameters were taken from GAFF[71] and 99SB[70] force fields. MD minimizations employed the particle mesh Ewald method periodic boundary conditions.[72]

Four different molecular clusters with different sizes resembling increasing molecular environments for involved catalytic residues were cut out of the minimized crystal structure, using 2, 3, and 4 Å around catalytic residues. Cut structures were then modified to resemble specific radical species on respective catalytic residues by abstracting corresponding hydrogen atoms and then optimized subsequently using the GFN-xTB[73] method as doublets using soft constraints on all hydrogen atoms. We name the systems according to the residue from which we abstracted the hydrogen atom, i.e., Cys418•, Cys419•, and Gly734•. The systems are shown in Figure 6.

In order to confirm the locality of the electronic spin density we analyze the Fermi contact term of all nuclei of the smallest
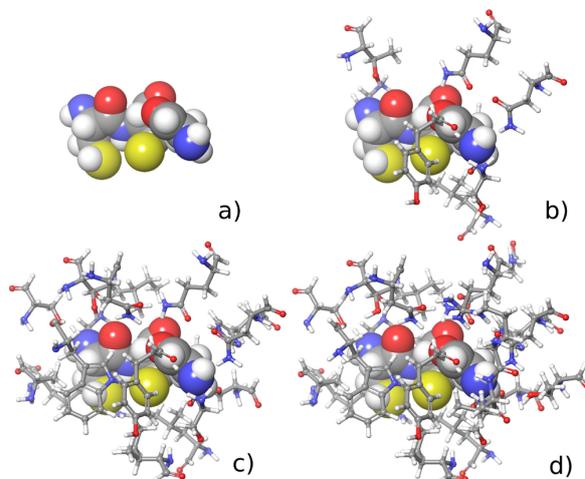
Figure 6. Cys419• in its protein environment. The system in a) consists of the three catalytic residues Cys418, Cys419, and Gly734 (32 atoms), the system in b) contains surrounding residues within 2 Å (162 atoms), c) within 3 Å (325 atoms), and d) within 4 Å (411 atoms).

systems consisting of the three catalytic residues at the DH-DFT level in Figure 7. It confirms the high localization at the sulfur atoms in the case of the radical systems Cys418• and Cys419•. In the case of Gly734•, a significant distribution along the neighboring carbon, oxygen, and nitrogen atoms can be seen.

In the following, we computed the HFCC of the atoms from which we abstracted the neighboring hydrogen atom of the three catalyic residues in their protein environment. We first compare the HFCCs obtained by B3LYP to the HFCCs obtained using B2-PLYP with both our selected-nuclei RI-CDD approach on CPUs and the ω-SOS-CDD ansatz on GPUs for the smallest system and the system with a 2 Å sphere in Table 2. The values highlight that the SOS approximation is valid in this case, as the SOS results mirror the behavior of the unscaled results and the deviations are small.

To further demonstrate the applicability of our new methods for the study of the influence of the protein environment on HFCCs, we computed the HFCCs of the radical centers with ω-SOS-CDD taking more and more of the environment into account. The results are shown in Figure 8 and clearly show a convergence of the HFCCs with an increase in the size of the protein environment. The radicals Cys418• and Cys419• only show a small dependence on the environment of less than 2 MHz. Intriguingly, the small effect of the environment on Cys419• is of opposite sign when correlation effects are taken into account by DH-DFT than with B3LYP. Gly734•, however, is highly dependent on the environment. Including residues within 2 Å leads to a change in the HFCC of 143 MHz. This can be explained by the fact that the smallest system consists of no covalently attached neighboring residues. By including neighboring residues, the electron spin density is localized more on the radical center, in contrast to the result obtained for the smallest system in Figure 7. Nonetheless, when further residues are taken into account, the HFCC still changes by more than 5 MHz.

Using our ω-SOS-CDD the HFCCs of large systems could be analyzed with a reduced computational effort. All three

(a) Cys418●
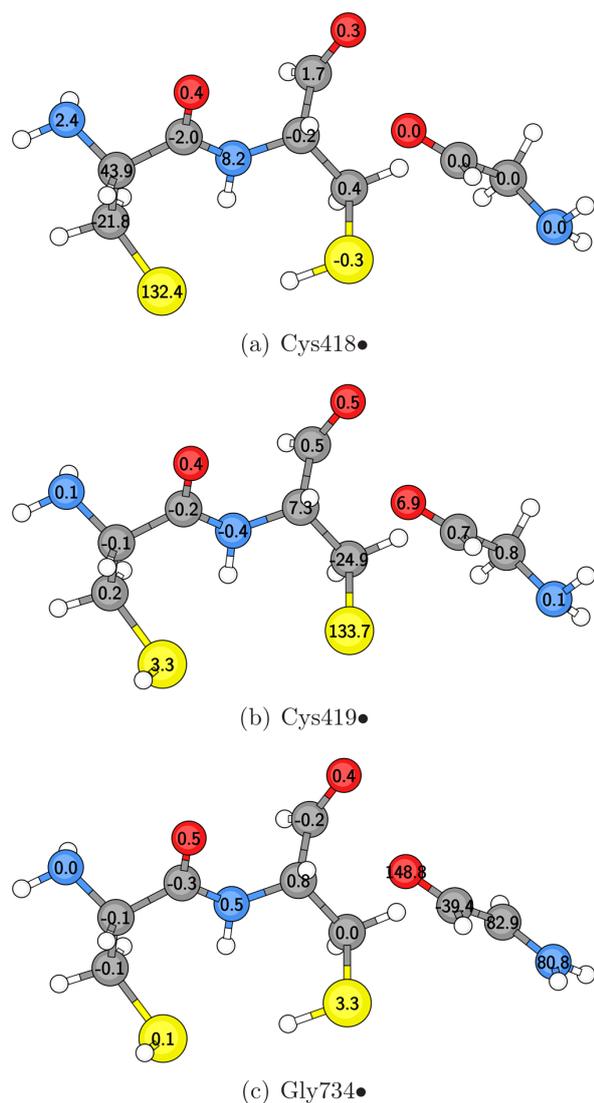
(b) Cys419●

(c) Gly734●

Figure 7. Fermi contact term of the electronic spin density of all nuclei of the three catalytic residues at the DH-DFT-level using B2-PLYP for the radicals Cys418●, Cys419●, and Gly734●. The values are displayed in 10^{-3} au as the HFCCs rely also on the nuclear magnetic moment and are thus less appropriate to highlight the locality of the spin density.

radicals were computed including residues within 3 Å in less than 80 h, the system with residues within 4 Å in less than 9 days. It has to be noted, that the systems with 2 Å spheres are

(a) Sulfur-atom HFCC (Cys418● and Cys419●)
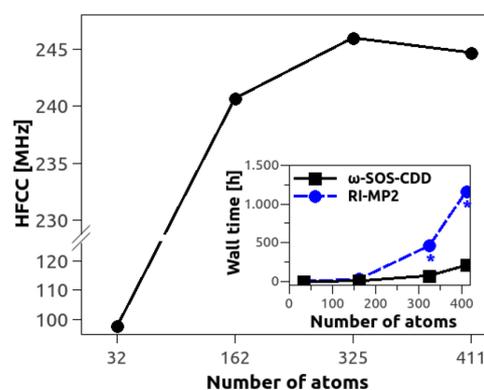
(b) Carbon-atom HFCC (Gly734●)

Figure 8. Dependence of the ω-SOS-CDD DH-DFT isotropic HFCC in three radicals (Cys418●, Cys419●, and Gly734●) as a function of the protein environment. Size convergence was studied by including residues within a 2 Å (162 atoms), 3 Å (325 atoms), and 4 Å (411 atoms) distance around the catalytic residues. The inlay shows the computational wall time of the Gly734 computations as a function of the number of atoms in comparison to the respective timings with the canonical method. Values indicated by an asterisk are extrapolated. The SCF of the largest system of the radical Cys418 could not be converged.

comparable in size with respect to the number of significant shellpairs with $C_{100}H_{201}$, whereas the systems with 3 Å spheres exceed the largest computed linear alkane system $(C_{200}H_{401})$ by a factor of 1.6. The systems with 4 Å spheres have more than twice the amount of significant shellpairs than $C_{200}H_{401}$.

5. CONCLUSION AND OUTLOOK

Based on our previous method to compute HFCCs at the MP2-level using the RI-CDD approach,[20] we have introduced a

Table 2. Isotropic HFCCs of the Spin Centers of the Catalytic Center of Three Different Radical Systems (Cys418●, Cys419●, Gly734●) with and without Surrounding Residues within 2 Å[a]

	Cys418 [MHz]			Cys419 [MHz]			Gly734 [MHz]		
	1	2	3	1	2	3	1	2	3
catalytic center	35.0	45.1	43.9	37.5	45.6	44.3	81.3	92.8	97.7
2 Å sphere	36.8	45.4	44.2	36.5	45.9	44.4	211.1	234.7	240.7

[a]The HFCCs were computed using the functional B3LYP (1), the new selected-nuclei RI-CDD B2-PLYP ansatz (2), and our new ω-SOS-CDD algorithm on GPUs(3).

methodology to compute the HFCCs of selected nuclei with reduced computational effort for systems with a localized spin density. For this, a sublinear number of two-electron integrals are determined to be significant by the use of a perturbation-including QQR-type integral screening. By only contracting screened two-electron integrals, an overall low-scaling and low-prefactor ansatz is obtained, leading to an up to 10-fold decrease in the computational effort with respect to our previous all-nuclei method.

Further, we introduce the computation of SOS-MP2 HFCCs of selected nuclei, where also an asymptotically sublinear scaling in the contraction can be achieved by BCSR algebra.[55,56] The overall efficiency can be improved with an attenuated Coulomb-metric,[38−40,54] as the scaling behavior of the transformation of the three-center integrals is reduced. The SOS-MP2 formulation can also be migrated to GPUs, where a modified J-engine as in ref 43 is employed. This results in a total CPU time spent for the computation of the isotropic HFCC of the terminal C atom in $C_{200}H_{401}$ on a node containing 12 CPU cores and four Nvidia GeForce GTX Titan GPUs of only 35 h. Similarly, the HFCC of a radical system in its protein environment with more than 300 atoms was computed in less than 80 h.

Our new approaches to efficiently compute the HFCCs of selected nuclei can be used to compute, e.g., the effect of the environment on molecular systems with a fully quantum mechanical description of the environment (at the MP2-level) as well as the computation of large molecular systems with high spin localization. Further work could include long-range solvation effects by, e.g., continuum solvation models. Our first promising results motivate future studies on an optimal scaling parameter in the SOS computation, in analogy to similar studies performed for NMR-shieldings,[42] with which even more reliable results at this simplified level are expected.

■ AUTHOR INFORMATION

Corresponding Author
*E-mail: christian.ochsenfeld@uni-muenchen.de.

ORCID ⓘ
Gökcen Savasci: 0000-0002-6183-7715
Christian Ochsenfeld: 0000-0002-4189-6558

Notes
The authors declare no competing financial interest.

■ REFERENCES

(1) Gerson, F.; Huber, W. *Electron Spin Resonance Spectroscopy of Organic Radicals*; Wiley-VCH Verlag GmbH & Co. KGaA: 2003; DOI: 10.1002/3527601627.

(2) Harriman, J. E. *Theoretical Foundations of Electron Spin Resonance*; Academic: 1978.

(3) Neese, F. Quantum chemical calculations of spectroscopic properties of metalloproteins and model compounds: EPR and Mössbauer properties. *Curr. Opin. Chem. Biol.* **2003**, *7*, 125−135.

(4) Chipman, D. M. Gaussian basis sets for calculation of spin densities in first-row atoms. *Theoret. Chim. Acta* **1989**, *76*, 73−84.

(5) Barone, V. *Recent Advances in Density Functional Methods*; World Scientific Singapore: 1995; p 287.

(6) Barone, V.; Cimino, P. Accurate and feasible computations of structural and magnetic properties of large free radicals: The PBE0/N07D model. *Chem. Phys. Lett.* **2008**, *454*, 139−143.

(7) Al Derzi, A. R.; Fau, S.; Bartlett, R. J. Benchmark Study of Isotropic Hyperfine Coupling Constants for Hydrogen: Influence of Geometry, Correlation Method, and Basis Set. *J. Phys. Chem. A* **2003**, *107*, 6656−6667.

(8) Verma, P.; Perera, A.; Morales, J. A. Massively parallel implementations of coupled-cluster methods for electron spin resonance spectra. I. Isotropic hyperfine coupling tensors in large radicals. *J. Chem. Phys.* **2013**, *139*, 174103.

(9) Puzzarini, C.; Barone, V. Toward spectroscopic accuracy for open-shell systems: Molecular structure and hyperfine coupling constants of H2CNH2CN, H2CPH2CP, NH2NH2, and PH2PH2 as test cases. *J. Chem. Phys.* **2010**, *133*, 184301.

(10) Engels, B. *Calculation of NMR and EPR Parameters*; Wiley-VCH: 2004; p 463.

(11) Shiozaki, T.; Yanai, T. Hyperfine Coupling Constants from Internally Contracted Multireference Perturbation Theory. *J. Chem. Theory Comput.* **2016**, *12*, 4347−4351.

(12) Hedegard, E. D.; Kongsted, J.; Sauer, S. P. A. Validating and Analyzing EPR Hyperfine Coupling Constants with Density Functional Theory. *J. Chem. Theory Comput.* **2013**, *9*, 2380−2388.

(13) Rogowska, A.; Kuhl, S.; Schneider, R.; Walcarius, A.; Champagne, B. Theoretical investigation of the EPR hyperfine coupling constants in amino derivatives. *Phys. Chem. Chem. Phys.* **2007**, *9*, 828−836.

(14) Fernández, B.; Jørgensen, P.; Byberg, J.; Olsen, J.; Helgaker, T.; Jensen, H. J. A. Spin polarization in restricted electronic structure theory: Multiconfiguration self-consistent-field calculations of hyperfine coupling constants. *J. Chem. Phys.* **1992**, *97*, 3412−3419.

(15) Rinkevicius, Z.; Telyatnyk, L.; Vahtras, O.; Ågren, H. Density functional theory for hyperfine coupling constants with the restricted-unrestricted approach. *J. Chem. Phys.* **2004**, *121*, 7614−7623.

(16) Grimme, S. Semiempirical hybrid density functional with perturbative second-order correlation. *J. Chem. Phys.* **2006**, *124*, 034108.

(17) Kossmann, S.; Kirchner, B.; Neese, F. Performance of modern density functional theory for the prediction of hyperfine structures: meta-GGA and double hybrid functionals. *Mol. Phys.* **2007**, *105*, 2049−2071.

(18) Munzarova, M.; Kaupp, M. A Critical Validation of Density Functional and Coupled-Cluster Approaches for the Calculation of EPR Hyperfine Coupling Constants in Transition Metal Complexes. *J. Phys. Chem. A* **1999**, *103*, 9966−9983.

(19) Neese, F. Metal and ligand hyperfine couplings in transition metal complexes: The effect of spin-orbit coupling as studied by coupled perturbed Kohn-Sham theory. *J. Chem. Phys.* **2003**, *118*, 3939.

(20) Vogler, S.; Ludwig, M.; Maurer, M.; Ochsenfeld, C. Low-scaling first-order properties within second-order Møller-Plesset perturbation theory using Cholesky decomposed density matrices. *J. Chem. Phys.* **2017**, *147*, 024101.

(21) Koch, H.; Sánchez de Meras, A.; Pedersen, T. B. Reduced scaling in electronic structure calculations using Cholesky decompositions. *J. Chem. Phys.* **2003**, *118*, 9481−9484.

(22) Boström, J.; Pitoňák, M.; Aquilante, F.; Neogrády, P.; Pedersen, T. B.; Lindh, R. Coupled Cluster and Møller-Plesset Perturbation Theory Calculations of Noncovalent Intermolecular Interactions using Density Fitting with Auxiliary Basis Sets from Cholesky Decompositions. *J. Chem. Theory Comput.* **2012**, *8*, 1921−1928.

(23) Whitten, J. L. Coulombic potential energy integrals and approximations. *J. Chem. Phys.* **1973**, *58*, 4496−4501.

(24) Dunlap, B. I.; Connolly, J. W. D.; Sabin, J. R. On some approximations in applications of Xα theory. *J. Chem. Phys.* **1979**, *71*, 3396−3402.

(25) Feyereisen, M.; Fitzgerald, G.; Komornicki, A. Use of approximate integrals in ab initio theory: An application in MP2 energy calculations. *Chem. Phys. Lett.* **1993**, *208*, 359−363.

(26) Vahtras, O.; Almlöf, J.; Feyereisen, M. Integral approximations for LCAO-SCF calculations. *Chem. Phys. Lett.* **1993**, *213*, 514−518.

(27) Weigend, F.; Häser, M. RI-MP2: first derivatives and global consistency. *Theor. Chem. Acc.* **1997**, *97*, 331−340.

(28) Weigend, F.; Häser, M.; Patzelt, H.; Ahlrichs, R. RI-MP2: optimized auxiliary basis sets and demonstration of efficiency. *Chem. Phys. Lett.* **1998**, *294*, 143−152.

(29) Maurer, S. A.; Lambrecht, D. S.; Flaig, D.; Ochsenfeld, C. Distance-dependent Schwarz-based integral estimates for two-electron integrals: Reliable tightness vs. rigorous upper bounds. *J. Chem. Phys.* **2012**, *136*, 144107.

(30) Maurer, S. A.; Lambrecht, D. S.; Kussmann, J.; Ochsenfeld, C. Efficient distance-including integral screening in linear-scaling Møller-Plesset perturbation theory. *J. Chem. Phys.* **2013**, *138*, 014101.

(31) Altenbach, C.; Marti, T.; Khorana, H. G.; Hubbell, W. L. Transmembrane protein structure: spin labeling of bacteriorhodopsin mutants. *Science* **1990**, *248*, 1088−1092.

(32) Jeschke, G.; Polyhach, Y. Distance measurements on spin-labelled biomacromolecules by pulsed electron paramagnetic resonance. *Phys. Chem. Chem. Phys.* **2007**, *9*, 1895−1910.

(33) Shelke, S. A.; Sigurdsson, S. T. Site-Directed Spin Labelling of Nucleic Acids. *Eur. J. Org. Chem.* **2012**, *2012*, 2291−2301.

(34) Gmeiner, C.; Klose, D.; Mileo, E.; Belle, V.; Marque, S. R. A.; Dorn, G.; Allain, F. H. T.; Guigliarelli, B.; Jeschke, G.; Yulikov, M. Orthogonal Tyrosine and Cysteine Site-Directed Spin Labeling for Dipolar Pulse EPR Spectroscopy on Proteins. *J. Phys. Chem. Lett.* **2017**, *8*, 4852−4857.

(35) Pauwels, E.; Declerck, R.; Verstraelen, T.; De Sterck, B.; Kay, C. M. W.; Van Speybroeck, V.; Waroquier, M. Influence of Protein Environment on the Electron Paramagnetic Resonance Properties of Flavoprotein Radicals: A QM/MM Study. *J. Phys. Chem. B* **2010**, *114*, 16655−16665.

(36) Sinnecker, S.; Neese, F. QM/MM Calculations with DFT for Taking into Account Protein Effects on the EPR and Optical Spectra of Metalloproteins. Plastocyanin as a Case Study. *J. Comput. Chem.* **2006**, *27*, 1463−1475.

(37) Fernández, B.; Christiansen, O.; Bludsky, O.; Jørgensen, P.; Mikkelsen, K. V. Theory of hyperfine coupling constants of solvated molecules: Applications involving methyl and ClO₂ radicals in different solvents. *J. Chem. Phys.* **1996**, *104*, 629−635.

(38) Jung, Y.; Sodt, A.; Gill, P. M. W.; Head-Gordon, M. Auxiliary basis expansions for large-scale electronic structure calculations. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 6692−6697.

(39) Jung, Y.; Shao, Y.; Head-Gordon, M. Fast evaluation of scaled opposite spin second-order Møller-Plesset correlation energies using auxiliary basis expansions and exploiting sparsity. *J. Comput. Chem.* **2007**, *28*, 1953−1964.

(40) Reine, S.; Tellgren, E.; Krapp, A.; Kjærgaard, T.; Helgaker, T.; Jansik, B.; Høst, S.; Salek, P. Variational and robust density fitting of four-center two-electron integrals in local metrics. *J. Chem. Phys.* **2008**, *129*, 104101.

(41) Jung, Y.; Lochan, R. C.; Dutoi, A. D.; Head-Gordon, M. Scaled opposite-spin second order Møller-Plesset correlation energy: An economical electronic structure method. *J. Chem. Phys.* **2004**, *121*, 9793.

(42) Maurer, M.; Ochsenfeld, C. Spin Component-Scaled Second-Order Møller-Plesset Perturbation Theory for Calculating NMR Shieldings. *J. Chem. Theory Comput.* **2015**, *11*, 37−44.

(43) Maurer, S. A.; Kussmann, J.; Ochsenfeld, C. Communication: A reduced scaling J-engine based reformulation of SOS-MP2 using graphics processing units. *J. Chem. Phys.* **2014**, *141*, 051106.

(44) Gauss, J. *Modern Methods and Algorithms of Quantum Chemistry*; John von Neumann Institute for Computing: Jülich, 2000; pp 541−592.

(45) Almlöf, J. Elimination of energy denominators in Møller-Plesset perturbation theory by a Laplace transform approach. *Chem. Phys. Lett.* **1991**, *181*, 319−320.

(46) Häser, M.; Almlöf, J. Laplace transform techniques in Møller-Plesset perturbation theory. *J. Chem. Phys.* **1992**, *96*, 489−494.

(47) Häser, M. Møller-Plesset (MP2) perturbation theory for large molecules. *Theor. Chim. Acta* **1993**, *87*, 147−173.

(48) Maurer, S. A.; Clin, L.; Ochsenfeld, C. Cholesky-decomposed density MP2 with density fitting: accurate MP2 and double-hybrid DFT energies for large systems. *J. Chem. Phys.* **2014**, *140*, 224112.

(49) Handy, N. C.; Schaefer, H. F., III On the evaluation of analytic energy derivatives for correlated wave functions. *J. Chem. Phys.* **1984**, *81*, 5031−5033.

(50) Yamaguchi, Y.; Goddard, J. D.; Osamura, Y.; Schaefer, H. F., III *A New Dimension to Quantum Chemistry: Analytic Derivative Methods in Ab Initio Molecular Electronic Structure Theory*; Oxford University Press Inc.: Oxford, 1994.

(51) Schweizer, S.; Doser, B.; Ochsenfeld, C. An atomic orbital-based reformulation of energy gradients in second-order Møller-Plesset perturbation theory. *J. Chem. Phys.* **2008**, *128*, 154101.

(52) Beer, M.; Ochsenfeld, C. Efficient linear-scaling calculation of response properties: Density matrix-based Laplace-transformed coupled perturbed self-consistent field theory. *J. Chem. Phys.* **2008**, *128*, 221102.

(53) Maurer, M.; Ochsenfeld, C. A linear- and sublinear-scaling method for calculating NMR shieldings in atomic orbital-based second-order Møller-Plesset perturbation theory. *J. Chem. Phys.* **2013**, *138*, 174104.

(54) Luenser, A.; Schurkus, H. F.; Ochsenfeld, C. Vanishing Overhead Linear Scaling Random Phase Approximation by Cholesky Decomposition and an Attenuated Coulomb-Metric. *J. Chem. Theory Comput.* **2017**, *13*, 1647−1655.

(55) Gustavson, F. G. Two Fast Algorithms for Sparse Matrices: Multiplication and Permuted Transposition. *ACM Trans. Math. Softw.* **1978**, *4*, 250−269.

(56) Kussmann, J.; Ochsenfeld, C. Linear-scaling method for calculating nuclear magnetic resonance chemical shifts using gauge-including atomic orbitals within Hartree-Fock and density-functional theory. *J. Chem. Phys.* **2007**, *127*, 054103.

(57) Menon, A. S.; Radom, L. Consequences of Spin Contamination in Unrestricted Calculations on Open-Shell Species: Effect of Hartree-Fock and Møller-Plesset Contributions in Hybrid and Double-Hybrid Density Functional Theory Approaches. *J. Phys. Chem. A* **2008**, *112*, 13225−13230.

(58) Neese, F.; Schwabe, T.; Grimme, S. Analytic derivatives for perturbatively corrected "double hybrid" density functionals: Theory, implementation, and application. *J. Chem. Phys.* **2007**, *126*, 124115.

(59) Kussmann, J.; Ochsenfeld, C. Pre-selective screening for matrix elements in linear-scaling exact exchange solutions. *J. Chem. Phys.* **2013**, *138*, 134114.

(60) Kussmann, J.; Ochsenfeld, C. Preselective Screening for Linear-Scaling Exact Exchange-Gradient Calculations for Graphics Processing Units and General Strong-Scaling Massively Parallel Calculations. *J. Chem. Theory Comput.* **2015**, *11*, 918−922.

(61) Takatsuka, A.; Ten-no, S.; Hackbusch, W. Minimax approximation for the decomposition of energy denominators in Laplace-transformed Møller-Plesset perturbation theories. *J. Chem. Phys.* **2008**, *129*, 044112.

(62) Weigend, F.; Ahlrichs, R. Balanced basis setse of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297−3305.

(63) Neese, F. The ORCA program system. *WIREs Comput. Mol. Sci.* **2012**, *2*, 73−78.

(64) Luenser, A.; Kussmann, J.; Ochsenfeld, C. Computation of indirect nuclear spin-spin couplings with reduced complexity in pure and hybrid density functional approximations. *J. Chem. Phys.* **2016**, *145*, 124103.

(65) Grimme, S. Improved second-order Møller-Plesset perturbation theory by separate scaling of parallel- and antiparallel-spin pair correlation energies. *J. Chem. Phys.* **2003**, *118*, 9095.

(66) Becker, A.; Kabsch, W. X-ray structure of pyruvate formate-lyase in complex with pyruvate and CoA. How the enzyme uses the Cys-418 thiyl radical for pyruvate cleavage. *J. Biol. Chem.* **2002**, *277*, 40036−40042.

(67) Plaga, W.; Vielhaber, G.; Wallach, J.; Knappe, J. Modification of Cys-418 of pyruvate formate-lyase by methacrylic acid, based on its radical mechanism. *FEBS Lett.* **2000**, *466*, 45−48.

(68) Becker, A.; Fritz-Wolf, K.; Kabsch, W.; Knappe, J.; Schultz, S.; Wagner, A. F. V. Structure and mechanism of the glycyl radical enzyme pyruvate formate-lyase. *Nat. Struct. Biol.* **1999**, *6*, 969−975.

(69) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781−1802.

(70) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins: Struct., Funct., Genet.* **2006**, *65*, 712−725.

(71) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general AMBER force field. *J. Comput. Chem.* **2004**, *25*, 1157−1174.

(72) Darden, T.; Perera, L.; Li, L.; Pedersen, L. New tricks for modelers from the crystallography toolkit: the particle mesh Ewald algorithm and its use in nucleic acid simulations. *Structure* **1999**, *7*, R55−R60.

(73) Grimme, S.; Bannwarth, C.; Shushkov, P. A Robust and Accurate Tight-Binding Quantum Chemical Method for Structures, Vibrational Frequencies, and Noncovalent Interactions of Large Molecular Systems Parametrized for All spd-Block Elements (Z = 1−86). *J. Chem. Theory Comput.* **2017**, *13*, 1989−2009.

7.3 Paper III: Gauge-origin dependence in electronic g-tensor calculations

The following article is reproduced in agreement with its publisher (AIP Publishing LLC) and can be found online at:

http://aip.scitation.org/doi/pdf/10.1063/1.5028454

Gauge-origin dependence in electronic g-tensor calculations

Michael Glasbrenner, Sigurd Vogler, and Christian Ochsenfeld[a)]
Chair of Theoretical Chemistry and Center for Integrated Protein Science Munich (CIPSM),
Department of Chemistry, University of Munich (LMU), Butenandtstr. 7, 81377 Munich, Germany

We present a benchmark study on the gauge-origin dependence of the electronic g-tensor using data from unrestricted density functional theory calculations with the spin-orbit mean field ansatz. Our data suggest in accordance with previous studies that g-tensor calculations employing a common gauge-origin are sufficiently accurate for small molecules; however, for extended molecules, the introduced errors can become relevant and significantly exceed the basis set error. Using calculations with the spin-orbit mean field ansatz and gauge-including atomic orbitals as a reference, we furthermore show that the accuracy and reliability of common gauge-origin approaches in larger molecules depends strongly on the locality of the spin density distribution. We propose a new pragmatic ansatz for choosing the gauge-origin which takes the spin density distribution into account and gives reasonably accurate values for molecules with a single localized spin center. For more general cases like molecules with several spatially distant spin centers, common gauge-origin approaches are shown to be insufficient for consistently achieving high accuracy. Therefore the computation of g-tensors using distributed gauge-origin methods like gauge-including atomic orbitals is considered as the ideal approach and is recommended for larger molecular systems. *Published by AIP Publishing.*
https://doi.org/10.1063/1.5028454

I. INTRODUCTION

A central issue in the calculation of magnetic molecular properties is the gauge-origin dependence. Unless gauge-origin independence is enforced by appropriate methods, computed values can depend on the positioning of the molecule within the chosen coordinate system.[1] This unphysical effect is only present in calculations with a finite basis set and vanishes in the complete basis set limit. Well-established methods exist that use distributed gauge-origins and assure gauge-origin independent results. The most commonly used approach employs gauge-including atomic orbitals (GIAOs, also called "London atomic orbitals").[2–4] Alternatively, the "individual gauge for localized orbitals" (IGLO)[6,7] or the "localized orbital/local origin" (LORG)[8] approaches have been used in methods employing localized molecular orbitals.[5]

Gauge-origin dependence is also a problem in g-tensor calculations; however, in several studies on the electronic g-tensor, it was found to be smaller than for other magnetic properties.[9–13] Still, some evidence of a non-negligible influence of the gauge in g-tensor calculations can be found in the literature: Lushington *et al.*[13] analyzed the gauge-origin dependence using Hartree-Fock and reported that for some of the studied molecules g-shifts changed considerably upon translation of the molecule by 1.73 bohrs in a Sadlej basis set[14] (30% change for Δg_{xx} of CO_2^-; 34% change for Δg_{xx} of C_3H_5). Kaupp *et al.*[15] observed a rather small dependence of the g-shifts on the chosen gauge-origin in calculations on various phenoxyl radicals, but a pronounced gauge-origin

dependence of the g-tensor orientation. Patchkovskii *et al.*[16] found a pronounced gauge-origin dependence of the (in most cases relatively small) spin-other-orbit contribution to the g-tensor. van Lenthe *et al.*[12] studied the gauge-origin dependence of their relativistic ansatz for g-tensor calculations based on the zeroth-order regular approximation method and observed for NO_2 with triple-ζ Slater type orbital basis set changes of around 20% upon translation of the molecule by 10 Å in the x-, y-, and z-direction. Lushington and Grein[17] reported a relative change of 18% in the g-shifts in multi-reference configuration interaction calculations for MgF with two choices of the gauge-origin [on the Mg atom or in the electronic charge centroid (ECC)[18]].

Despite these examples which show that the gauge-origin dependence can have a non-negligible effect on the g-shifts, many studies on electronic g-tensors employ a common gauge-origin.[10,11,19–35] This can also be explained by the fact that most previous studies focused on g-tensors of small molecules.[17,28,30,33,34,36,37] By contrast, for g-tensors of large molecules, the gauge-origin dependence of the obtained values can be expected to be much more severe. This is because the errors associated with the gauge-origin dependence increase with distance from the gauge-origin;[1] in extended molecules always some parts of the molecule are relatively far apart from the origin — no matter how the molecule is positioned. As several recent studies presented g-tensor calculations on larger molecules,[38–45] a careful evaluation of the influence of the gauge-origin dependence in larger systems appears to be necessary.

To this end we carried out the, to our knowledge, most extensive study on the gauge-origin dependence of the electronic g-tensor to date. In particular, we also discuss the

---

a)Electronic mail: christian.ochsenfeld@uni-muenchen.de

accuracy and reliability of common gauge-origin approaches for larger molecules and highlight the role played by the spin density distribution. Our method of choice is unrestricted density functional theory (DFT) because it provides a reasonable compromise between accuracy and computational efficiency. We use B3LYP[46] as a functional, which has shown a good agreement with coupled-cluster singles and doubles (CCSD) calculations on g-tensors in a recent study by Perera *et al.*[47] Further details on our employed ansatz are provided in Sec. II.

II. THEORY

The components of the electronic g-tensor can be computed in analytical derivative theory by taking a second derivative of the energy E with respect to the magnetic field $\vec{B}$ and the electronic spin $\vec{s}$

$$g_{pq} = \frac{2}{\alpha} \frac{\partial^2 E}{\partial B_p \partial s_q}\bigg|_{B=0,s=0}, \qquad p,q \in \{x,y,z\}, \qquad (1)$$

where α denotes the fine structure constant.[48] Throughout this paper, only the three g-shifts or their isotropic average are shown; the g-shifts are the deviations of the singular values of the g-tensor from the g-value of the free electron g_{el}. The contributions to the g-tensor include the relativistic mass correction g^{rmc}, the diamagnetic one- and two-electron terms (g^{dso1} and g^{dso2}), and the paramagnetic one- and two-electron terms (g^{pso1} and g^{pso2})[48]

$$g_{pq} = \delta_{pq} g_{\text{el}} + g_{pq}^{\text{rmc}} + g_{pq}^{\text{dso1}} + g_{pq}^{\text{dso2}} + g_{pq}^{\text{pso1}} + g_{pq}^{\text{pso2}}. \qquad (2)$$

In the following, we summarize how these contributions are computed in our implementation. We initially show the equations for a basis of regular atomic orbitals (AOs) and then describe the necessary modifications for a basis of GIAOs.

g^{rmc} is readily computed from the kinetic energy integrals $T_{\mu\nu}$ and the spin density $P_{\mu\nu}^{\alpha-\beta}$

$$g_{pq}^{\text{rmc}} = -\frac{\alpha^2 g_{\text{el}}}{2S} \delta_{pq} \sum_{\mu\nu} P_{\mu\nu}^{\alpha-\beta} T_{\mu\nu}. \qquad (3)$$

S is the total spin of the electronic state (e.g., $\frac{1}{2}$ for doublet states and $\frac{2}{2}$ for triplet states). The diamagnetic one-electron contribution is given by the following expression:

$$g_{pq}^{\text{dso1}} = \frac{\alpha^2 g'}{8S} \sum_A Z_A \sum_{\mu\nu} P_{\mu\nu}^{\alpha-\beta}$$
$$\left[\int \chi_\mu(\vec{r}) \frac{\delta_{pq}(\vec{r}-\vec{R}_A)\cdot(\vec{r}-\vec{R}_0)}{|\vec{r}-\vec{R}_A|^3} \chi_\nu(\vec{r}) dr^3 \right.$$
$$\left. - \int \chi_\mu(\vec{r}) \frac{(\vec{r}-\vec{R}_A)_p(\vec{r}-\vec{R}_0)_q}{|\vec{r}-\vec{R}_A|^3} \chi_\nu(\vec{r}) dr^3 \right], \qquad (4)$$

where Z_A and $\vec{R}_A$ are the charge and position of nucleus A, respectively, and g' is the electronic spin-orbit g-factor.[49] $\vec{R}_0$ is the chosen gauge-origin. The diamagnetic two-electron contribution is usually rather small; we only indirectly account for it by replacing the nuclear charges Z_A in Eq. (4) by effective

nuclear charges $\tilde{Z}_A$ as determined by Koseki *et al.*[37] In this way, we obtain for the diamagnetic contributions

$$g_{pq}^{\text{dso1}} + g_{pq}^{\text{dso2}} \approx \frac{\alpha^2 g'}{8S} \sum_A \tilde{Z}_A \sum_{\mu\nu} P_{\mu\nu}^{\alpha-\beta}$$
$$\left[\int \chi_\mu(\vec{r}) \frac{\delta_{pq}(\vec{r}-\vec{R}_A)\cdot(\vec{r}-\vec{R}_0)}{|\vec{r}-\vec{R}_A|^3} \chi_\nu(\vec{r}) dr^3 \right.$$
$$\left. - \int \chi_\mu(\vec{r}) \frac{(\vec{r}-\vec{R}_A)_p(\vec{r}-\vec{R}_0)_q}{|\vec{r}-\vec{R}_A|^3} \chi_\nu(\vec{r}) dr^3 \right]. \qquad (5)$$

For the paramagnetic terms, we employ the spin-orbit mean field (SOMF) operator.[50(a),50(b)] It treats the two-electron spin-orbit coupling in a mean field way and has been shown to be a highly accurate approximation to the exact spin-orbit operators.[50(a)] The SOMF operator $\hat{z}$ has a similar structure to the Fock operator and contains one-electron, Coulomb, and exchange contributions

$$z_{\mu\nu}^q = h_{\mu\nu,q}^{\text{SOMF}} + J_{\mu\nu,q}^{\text{SOMF}} + K_{\mu\nu,q}^{\text{SOMF}}, \qquad q \in \{x,y,z\}, \qquad (6)$$

$$h_{\mu\nu,q}^{\text{SOMF}} = \frac{\alpha^2 g'}{4} \sum_A Z_A \int \chi_\mu(\vec{r}) \frac{\left[(\vec{r}-\vec{R}_A)\times\hat{p}\right]_q}{|\vec{r}-\vec{R}_A|^3} \chi_\nu(\vec{r}) dr^3, \qquad (7)$$

$$J_{\mu\nu,q}^{\text{SOMF}} = -\frac{\alpha^2 g'}{4} \sum_{\lambda\sigma} P_{\lambda\sigma} g_{\mu\nu\lambda\sigma}^{\text{soc, q}}, \qquad (8)$$

$$K_{\mu\nu,q}^{\text{SOMF}} = \frac{3}{8} \alpha^2 g' \sum_{\lambda\sigma} \left(P_{\lambda\sigma} g_{\mu\lambda\sigma\nu}^{\text{soc, q}} + P_{\lambda\sigma} g_{\sigma\nu\mu\lambda}^{\text{soc, q}} \right), \qquad (9)$$

$$g_{\mu\nu\lambda\sigma}^{\text{soc, q}} = \int \chi_\mu(\vec{r}) \frac{\left[(\vec{r}-\vec{r}')\times\hat{p}\right]_q}{|\vec{r}-\vec{r}'|^3} \chi_\nu(\vec{r}) \chi_\lambda(\vec{r}') \chi_\sigma(\vec{r}') dr^3 dr'^3. \qquad (10)$$

Finally, we further apply the 1X-approximation, which only introduces minor errors, to the SOMF operator as suggested by Neese.[50(a)] The SOMF-1X operator is obtained from the full SOMF-operator by neglecting all multi-center integrals in Eq. (9).

With the SOMF operator, the paramagnetic contribution to the g-tensor can be obtained as follows:

$$g_{pq}^{\text{pso1}} + g_{pq}^{\text{pso2}} \approx \frac{1}{\alpha S} \sum_{\mu\nu} \frac{\partial P_{\mu\nu}^{\alpha-\beta}}{\partial B_p}\bigg|_{s=0} z_{\mu\nu}^q. \qquad (11)$$

The magnetic field derivative of the spin density is obtained from the difference of the perturbed α and β densities

$$\frac{\partial P_{\mu\nu}^{\alpha-\beta}}{\partial B_p}\bigg|_{s=0} = \frac{\partial P_{\mu\nu}^{\alpha}}{\partial B_p}\bigg|_{s=0} - \frac{\partial P_{\mu\nu}^{\beta}}{\partial B_p}\bigg|_{s=0}, \qquad (12)$$

which in turn can be computed by unrestricted coupled-perturbed self-consistent field (CPSCF) as described, e.g., in Ref. 51.

In addition to the gauge-origin dependent theory in a regular AO basis described so far, we also implemented the SOMF ansatz in combination with GIAOs, which are obtained from a product of a regular AO basis function $\chi_\mu(\vec{r})$ and a magnetic field dependent phase factor[4]

$$\phi_\mu(\vec{r}) = \exp\left[-i\frac{\alpha}{2}(\vec{B}\times(\vec{R}_\mu-\vec{R}_0))\cdot\vec{r}\right] \chi_\mu(\vec{r}). \qquad (13)$$

GIAOs have previously been used in several methods for g-tensor computations such as the coupled-cluster approach from Gauss *et al.*[69] and the DFT ansatz from Schreckenbach and Ziegler.[48] If GIAOs are employed the expression for the diagmagnetic contributions changes as follows:

$$
g_{pq}^{dso1} + g_{pq}^{dso2} \approx \frac{\alpha^2 g'}{8S} \sum_A \tilde{Z}_A \sum_{\mu\nu} P_{\mu\nu}^{\alpha-\beta}
$$
$$
\left[\int \chi_\mu(\vec{r}) \frac{\delta_{pq}(\vec{r} - \vec{R}_A) \cdot (\vec{r} - \vec{R}_\nu)}{|\vec{r} - \vec{R}_A|^3} \chi_\nu(\vec{r}) dr^3 \right.
$$
$$
\left. - \int \chi_\mu(\vec{r}) \frac{(\vec{r} - \vec{R}_A)_p (\vec{r} - \vec{R}_\nu)_q}{|\vec{r} - \vec{R}_A|^3} \chi_\nu(\vec{r}) dr^3 \right].
$$
(14)

In a GIAO basis, the magnetic field derivatives of the integrals contributing to the Fock matrix and of the exchange-correlation potential need to be included in the CPSCF equations as described in Ref. 52 for restricted CPSCF. The use of GIAOs also leads to an additional paramagnetic contribution from the B-field derivative of the SOMF matrix

$$
g_{pq}^{pso1} + g_{pq}^{pso2} \approx \frac{1}{\alpha S} \sum_{\mu\nu} \left[\frac{\partial P_{\mu\nu}^{\alpha-\beta}}{\partial B_p} \bigg|_{s=0} z_{\mu\nu}^q + P_{\mu\nu}^{\alpha-\beta} \left(\frac{\partial z_{\mu\nu}^q}{\partial B_p} \right) \right],
$$
(15)

$$
\frac{\partial z_{\mu\nu}^q}{\partial B_p} = -\frac{\alpha}{2} \left(\vec{R}_\mu \times \vec{R}_\nu \right)_p z_{\mu\nu}^q - \frac{\alpha}{2} \left[\left(\vec{R}_\mu - \vec{R}_\nu \right) \times z_{\mu\nu^+}^q \right]_p.
$$
(16)

Here, $\vec{R}_\mu$ and $\vec{R}_\nu$ are the positions of the atomic centers of χ_μ and χ_ν, respectively. The index ν^+ indicates the basis function χ_ν, whose angular momentum has been incremented in either x-, y-, or z-direction as necessitated by the cross product. Note that the basis functions χ_λ and χ_σ in Eqs. (8)–(10) are not substituted by GIAOs because the SOMF operator enters the Hamiltonian as a one-electron operator. Considerable simplification occurs if the 1X-approximation is invoked; in this case, no exchange terms need to be computed during the evaluation of Eq. (16) because both $\vec{R}_\mu \times \vec{R}_\nu$ and $\vec{R}_\mu - \vec{R}_\nu$ vanish if all basis functions are centered on the same atom.

III. COMPUTATIONAL DETAILS

The theory described in Sec. II was implemented both for a basis of GIAOs and a basis of regular AOs in a development version of the FermiONs++ program.[53,54] The reference state is obtained via unrestricted Kohn-Sham DFT and the SCF energy converged to a threshold of 10^{-8} a.u. Direct inversion of the iterative subspace (DIIS)[55] is employed for updating the perturbed density during the CPSCF. The perturbed densities were converged below 10^{-7}. Shell pairs with a maximum basis function overlap of less than 10^{-12} were omitted from the calculation. The molecular integration grid used for DFT was generated as a product of a spherical Lebedev/Laikov[56] grid with 590 angular points and a Treutler-Ahlrichs "M4" grid[57] with 99 radial points. We employ basis sets from the def2-series[58] (def2-SVP, def2-TZVP, and def2-QZVPPD[59]); these basis sets are well-suited for DFT calculations and have also been used in several recent DFT studies on g-tensors.[38,40–44]

The structures of the small molecules used in Secs. IV A and IV B were taken from two sources: a set of molecules from Schreckenbach and Ziegler used for benchmarking their g-tensor ansatz[48] and a set of small main group radicals from Ref. 60 with structures optimized by unrestricted CCSD with perturbative triples [UCCSD(T)]. The used test set contains 45 molecules, including different spin states (doublets and triplets) as well as molecular charges (anions, cations, and neutral species). The geometries of the molecules used for Secs. IV C and IV D were optimized with the HF-3c method[61] using the ORCA program.[62,63] All computed g-shifts are shown in the supplementary material. The employed geometries are available for download at http://www.cup.uni-muenchen.de/pc/ochsenfeld/download/.

IV. RESULTS AND DISCUSSION

A. Significance of the gauge-origin dependence in g-tensor calculations

In this section, we present calculations on a test set of 45 main group radicals for the purpose of benchmarking the gauge-origin dependence of the electronic g-tensor at the DFT level of theory. We quantify the gauge-origin dependence in our calculations by performing two calculations per molecule with different positioning of the molecule relative to the gauge-origin: in one of these calculations the gauge-origin is positioned in the center of mass of the molecule; in the second calculation the molecule is translated away from the gauge-origin by 10.0 Å in the x-, y-, and z-direction. The difference in the g-shifts in these calculations is in the following denoted as Δgauge. In order to assess the relevance of the observed gauge-origin dependence and its practical implications, we try to compare the gauge error to the basis set error. Both of these errors represent different aspects of the basis set incompleteness and vanish for a complete basis set. Our estimate for basis set error (which we denote as Δbasis) is obtained from the difference to the g-shifts computed in a def2-QZVPPD basis set, which we assume to be close to the complete basis set limit (in both calculations the gauge-origin is at the center of mass). Based on the comparison between gauge error and basis set error, we then discuss to what extent the use of distributed gauge-origin methods like GIAOs can improve the accuracy of g-tensor calculations in a given basis.

Although our main interest focuses on the gauge errors that can occur in medium-sized and larger molecules, we exclusively employ small molecules (less than eight atoms) in the calculations of this section. We avoid larger molecules as the calculations with the def2-QZVPPD basis set would be computationally very demanding. However, as we will show in Secs. IV C and IV D in calculations on extended molecules, the contributions to the molecular g-tensor are often local and originate from small parts of the molecule. We therefore think that Δgauge provides a good estimate of the errors that can be expected when contributions to the g-tensor from a group of atoms inside a large molecule are computed—assuming that this group of atoms does not happen to be spatially close to the gauge-origin. The employed translation distance of 17.3 Å (10.0 Å in each spatial direction) is not unreasonably large for this purpose as the numerous studies
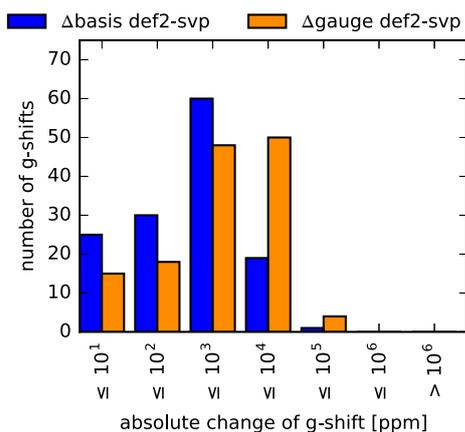
FIG. 1. Comparison of the influence of the basis set and gauge-origin on g-shifts in our test set. Δbasis is the absolute deviation of the value computed with the def2-SVP basis set from the def2-QZVPPD result obtained with the same nuclear coordinates. Δgauge is the absolute change in the individual g-shifts upon a translation of the molecule by 10.0 Å in the x-, y-, and z-direction (using def2-SVP as a basis set in both calculations).

FIG. 3. Comparison of the influence of the basis set and gauge-origin on g-shifts in our test set. Δbasis is the unsigned relative deviation of the value computed with the def2-SVP basis set from the def2-QZVPPD result obtained with the same nuclear coordinates. Δgauge is the unsigned relative change in the individual g-shifts upon a translation of the molecule by 10.0 Å in the x-, y-, and z-direction (using def2-SVP as a basis set in both calculations).

using EPR spectroscopy on biological macromolecules including, e.g., spin-labeled proteins,[64,65] lipids,[66] or nucleic acid polymers,[67] show. For such molecules, inter-atomic distances can easily exceed 17.3 Å. Nevertheless, we will also discuss the implications of our findings for calculations on smaller molecules.

In Figs. 1–4, we present how large Δgauge and Δbasis are in absolute and relative terms for the molecules in our test set; for this purpose, we sorted the computed deviations into categories in order to show how these errors vary among the test set. The calculation with the same basis set and with the gauge-origin in the center of mass is taken as a reference value for the relative deviations due to gauge-origin dependence. The relative deviations due to Δbasis are computed relative to the def2-QZVPPD values (also with the center of mass as gauge-origin). Statistical values on the calculations are presented in

Table I. As g-shifts can vary over several orders of magnitude for different molecules, it is worth mentioning that the average isotropic g-shift of our test set amounts to 3495 ppm (def2-QZVPPD basis).

For the def2-SVP basis set and the employed translation distance of 17.3 Å, the average value of Δgauge exceeds the average of Δbasis significantly—both for individual g-shifts (2318 ppm compared to 693 ppm) and for isotropic g-shifts (1967 compared to 642 ppm). In general, we observe that for molecules whose three g-shifts have significantly different magnitudes, the smallest shifts of these g-shifts show the highest sensitivity to the gauge. This effect explains why the mean relative deviations in the individual g-shifts due to Δgauge (346.3% for def2-SVP; 95.9% for def2-TZVP) are much higher than the mean relative deviations in the isotropic g-shifts (79.3% for def2-SVP; 25.5% for def2-TZVP). These

FIG. 2. Comparison of the influence of the basis set and gauge-origin on g-shifts in our test set. Δbasis is the absolute deviation of the value computed with the def2-TZVP basis set from the def2-QZVPPD result obtained with the same nuclear coordinates. Δgauge is the absolute change in the individual g-shifts upon a translation of the molecule by 10.0 Å in the x-, y-, and z-direction (using def2-TZVP as a basis set in both calculations).

FIG. 4. Comparison of the influence of the basis set and gauge-origin on g-shifts in our test set. Δbasis is the unsigned relative deviation of the value computed with the def2-TZVP basis set from the def2-QZVPPD result obtained with the same nuclear coordinates. Δgauge is the unsigned relative change in the individual g-shifts upon a translation of the molecule by 10.0 Å in the x-, y-, and z-direction (using def2-TZVP as a basis set in both calculations).

TABLE I. Mean and standard deviations (s.d.) from the calculations on our test set with either def2-SVP or def2-TZVP basis set. Δbasis is the deviation of the computed value from the def2-QZVPPD result obtained with the same nuclear coordinates. Δgauge is the change in the individual g-shifts upon a translation of the molecule by 10.0 Å in the x-, y-, and z-direction (using the same basis set in both calculations). Shown are either statistics on the individual g-shifts ("all"), the isotropic g-shifts ("iso"), or the isotropic g-shifts between 500 and 5000 ppm ["iso (500–5000 ppm)"].

Shifts		def2-SVP		def2-TZVP	
		Absolute (ppm)	Relative (%)	Absolute (ppm)	Relative (%)
All	Δgauge (mean)	2318	346.3	1193	95.9
	Δbasis (mean)	693	15.5	175	7.3
	Δgauge (s.d.)	5245	1690.4	3041	463.8
	Δbasis (s.d.)	1564	23.5	368	15.9
Iso	Δgauge (mean)	1967	79.3	1041	25.5
	Δbasis (mean)	642	14.1	156	6.3
	Δgauge (s.d.)	3580	121.1	2228	36.7
	Δbasis (s.d.)	990	14.5	253	10.8
Iso (500–5000 ppm)	Δgauge (mean)	1290	102.9	387	21.2
	Δbasis (mean)	274	13.1	102	4.4
	Δgauge (s.d.)	1041	152.1	302	17.1
	Δbasis (s.d.)	477	17.1	263	7.2

large relative deviations can also be seen from Figs. 3 and 4.

Δbasis is reduced by roughly a factor of four upon changing from a def2-SVP basis to a def2-TZVP basis (693 ppm compared to 175 ppm); Δgauge is on average only reduced by roughly a factor of two (2318 ppm compared to 1193 ppm). If only isotropic g-shifts between 500 and 5000 ppm are included in the statistics, Δgauge is diminished to a somewhat larger extent upon changing from a def2-SVP to a def2-TZVP basis set (reduction from 1290 ppm to 387 ppm) than Δbasis (reduction from 274 ppm to 102 ppm). Thus the gauge-error is of similar magnitude relative to the basis set error both for the def2-SVP and the def2-TZVP basis sets, even though its absolute size decreases by about 50% for the larger basis set.

The shown data imply that the "total basis set incompleteness error" (including the gauge error and the regular basis set error) in contributions to the g-tensor from a group of atoms displaced 17.3 Å from the gauge-origin can be reduced considerably if distributed gauge-origin methods are employed as the gauge error exceeds the standard basis set error significantly (for def2-TZVP/individual g-shifts by a factor of 6.8). Thus, the use of distributed gauge-origin approaches would be beneficial in situations where the gauge-origin cannot be positioned closer to sites of the molecule with significant contributions to the g-tensor — especially considering that distributed gauge-origin approaches increase the computational cost to a much smaller extent than an increase of the basis set.

In general, the components of the g-tensor depend linearly on the displacement from the gauge-origin.[13] The same linear dependence also holds for the closely related NMR shielding tensors.[68] This does not imply a strictly linear dependence of the individual g-shifts; only the isotropic g-shift changes in a linear fashion with increasing distance from the gauge-origin. Therefore one can extrapolate how large the gauge-error in the isotropic g-shifts would be on average for smaller or larger displacements than the employed 17.3 Å. For def2-SVP, the gauge error in the isotropic g-shift amounts to 113.7 ppm/Å on average; for def2-TZVP, it amounts to 60.2 ppm/Å. For a distance of 1.0 Å to the gauge-origin, the gauge-error is clearly below Δbasis suggesting that distributed gauge-origin methods would not significantly improve the accuracy in g-tensor calculations on very small molecules. On average, the gauge error in the isotropic g-shifts exceeds Δbasis for displacements larger than 5.6 Å (def2-SVP) and 2.6 Å (def2-TZVP). If only values between 500 and 5000 ppm are included, Δgauge is on average larger than Δbasis for displacements above 3.7 Å (def2-SVP) and 4.6 Å (def2-TZVP). These results imply that the usage of distributed gauge-origin methods can improve the accuracy considerably for extended molecules.

We also found rather large variations of the gauge-origin dependence among the molecules in our test set. In Table II, data from several molecules are shown, whose g-shifts display a quite pronounced dependence on the gauge. For all these molecules, the gauge error after displacement significantly exceeds the changes that are caused by an enlargement of the basis set. With the def2-SVP basis set, the isotropic g-shifts of the molecules in Table II deviate on average 934 ppm from the def2-QZVPPD numbers; after displacement this deviation grows to 6374 ppm. For the def2-TZVP basis, the deviations relative to the def2-QZVPPD values are increased by more than one order of magnitude by the translation away from the gauge-origin (from 155 ppm to 2197 ppm). The gauge error in the calculations exceeds Δbasis on average for more than 2.5 Å (def2-SVP) and 1.2 Å (def2-TZVP) distance to the gauge-origin. This shows even more impressively than the statistics on the entire test set how large influences of the gauge in g-tensor calculations can be. It

TABLE II. Examples of molecules in our test set with strong gauge-origin dependence. The column "transl." indicates whether the molecule was positioned close to the origin or translated away from it by 10.0 Å in the x-, y- and z- direction.

Molecule	Basis	Transl.	Δg_1 (ppm)	Δg_2 (ppm)	Δg_3 (ppm)
ClO_3	def2-SVP	No	10 056	10 056	847
	def2-SVP	Yes	16 099	15 221	−25
	def2-TZVP	No	7 581	7 581	1058
	def2-TZVP	Yes	10 185	9 892	766
	def2-QZVPPD	No	7 316	7 316	1105
GeH_3	def2-SVP	No	14 613	14 613	−211
	def2-SVP	Yes	55 387	46 245	−9106
	def2-TZVP	No	16 196	16 196	−61
	def2-TZVP	Yes	29 764	27 576	−2217
	def2-QZVPPD	No	16 762	16 762	−107
MgF	def2-SVP	No	−5	−1 905	−1905
	def2-SVP	Yes	877	−5 169	−6048
	def2-TZVP	No	−3	−1 705	−1705
	def2-TZVP	Yes	119	−2 504	−2625
	def2-QZVPPD	No	−3	−1 742	−1742
SO_3^-	def2-SVP	No	4 379	4 379	73
	def2-SVP	Yes	8 564	7 852	−636
	def2-TZVP	No	2 761	2 761	366
	def2-TZVP	Yes	4 709	4 399	57
	def2-QZVPPD	No	2 534	2 534	169
SiH_3	def2-SVP	No	2 062	2 062	−89
	def2-SVP	Yes	7 360	6 205	−1239
	def2-TZVP	No	2 266	2 266	−78
	def2-TZVP	Yes	2 619	2 599	−99
	def2-QZVPPD	No	2 285	2 285	−78

TABLE III. Illustrative examples of the basis set convergence in some of the molecules in our test set. "+ giao" indicates the use of a GIAO basis; in all other cases, a regular AO basis set was employed and the gauge-origin was positioned in the center of mass.

Molecule	Basis	Δg_1 (ppm)	Δg_2 (ppm)	Δg_3 (ppm)
BS	def2-SVP	−83	−8594	−8594
	def2-SVP + giao	−83	−8358	−8358
	def2-TZVP	−84	−9338	−9338
	def2-TZVP + giao	−84	−9322	−9322
	def2-QZVPPD	−84	−9099	−9099
CO^+	def2-SVP	−136	−2336	−2336
	def2-SVP + giao	−136	−2255	−2255
	def2-TZVP	−135	−2534	−2534
	def2-TZVP + giao	−135	−2514	−2514
	def2-QZVPPD	−134	−2535	−2535
NH_3^+	def2-SVP	1449	1449	−162
	def2-SVP + giao	1475	1475	−156
	def2-TZVP	1595	1595	−158
	def2-TZVP + giao	1600	1600	−153
	def2-QZVPPD	1657	1657	−156

the DFT level only to a very small extent. This result agrees with the findings from Gauss *et al.*, who reported only marginal acceleration of the basis set convergence by GIAOs in their CCSD approach.[69]

In Secs. IV C and IV D, we will present calculations with different choices of a common gauge-origin and compare them to calculations with GIAOs. Based on the findings of this section, this comparison is suitable for estimating the gauge error in the calculations without GIAOs, as GIAOs ensure gauge-origin independence but only have a small impact on the basis set convergence.

C. Gauge-origin dependence in molecules with a single localized spin center

As the calculations on small molecules shown in Sec. IV A suggest, errors due to gauge-origin dependence can significantly deteriorate the achievable accuracy in g-tensor calculations on larger molecules. In the current section and Sec. IV D, we present calculations on medium-sized molecules. In order to assess the influence of the gauge-origin dependence, we compare the g-shifts computed with our SOMF-GIAO ansatz to g-shifts obtained with different choices of a common

strongly suggests that the gauge-origin dependence cannot be ignored for larger molecules because the introduced errors might render the computed values meaningless. Distributed gauge-origin methods offer one elegant way to completely remove the gauge error at moderately increased computational cost and should be employed for larger molecules if possible. In Secs. IV C and IV D, we will present calculations on larger molecules and analyze the gauge-origin dependence in these systems.

B. Influence of GIAOs on the basis set convergence

Most importantly, GIAOs ensure gauge-origin independent results in magnetic property calculations. Apart from that, the use of GIAOs also has the desirable advantage of an accelerated basis set convergence. In this section, we analyze to what extent the GIAOs accelerate basis set convergence in g-tensor computations at the DFT level. We present some example data from our test set (Table III) and statistical values (Table IV). For the def2-SVP basis set, the use of GIAOs reduces the absolute deviation to the def2-QZVPPD result from 693 ppm to 645 ppm on average while the relative deviation decreases from 15.5% to 14.6%. For the def2-TZVP basis, mean absolute deviations of 196 ppm and 175 ppm absolute deviation with and without GIAOs are observed; the mean relative deviations amount to 6.5% and 7.3%. It shows that GIAOs accelerate the basis set convergence in g-tensor calculations at

TABLE IV. Deviations of the def2-SVP or def2-TZVP data from the def2-QZVPPD reference values for our test set. Mean absolute deviation (MAD), absolute standard deviation (ASD), mean relative deviation (MRD), and relative standard deviation (RSD) are given. The calculations were carried out in a regular AO basis ("AO") or in a GIAO basis ("GIAO").

	def2-SVP		def2-TZVP	
	AO	GIAO	AO	GIAO
MAD (ppm)	693	645	175	196
ASD (ppm)	1564	1523	368	437
MRD (%)	15.5	14.6	7.3	6.5
RSD (%)	23.5	17.0	15.9	9.4

gauge-origin. One of them is the electronic charge centroid (ECC) which is the most commonly employed gauge-origin in the literature on g-tensors.[10,11,22–35] Other choices for the gauge-origin that have been employed in the literature are the charge center of the nuclei[20] or the center of mass.[21]

All of the mentioned choices lead to a gauge-origin that is quite centrally located within the molecule. However, we claim that such gauge-origins are in general not optimal as they do not take into account the local nature of the g-tensor. As one can see from Eqs. (3)–(16), the g-tensor is determined by the distribution of the ground state spin density and the first order response of the spin density to the applied magnetic field perturbation. In many larger open-shell molecules, the spin density is highly local. The perturbed spin density [Eq. (12)] is exactly zero for a closed-shell reference state as all spin-dependent operators in the Hamiltonian are set to zero for $\vec{s} = \vec{0}$, and α and β electrons are therefore affected in the same way by the magnetic field perturbation. In large molecules with an electronic structure that only locally displays open-shell character, one can accordingly expect the perturbed spin density to be a local quantity as well. We also observed this in our calculations; in Fig. 5, plots of the ground state and perturbed spin densities in the $\cdot$O$-$[CH$_2$]$_{18}$$-CH_3$ radical are presented, which show their similar distribution and locality.

Given a spin density and perturbed spin densities that are all well localized within one small part of the molecule, it is clear that significant contributions to the g-tensor only arise from this spatial region. In order to accurately compute these contributions, a common gauge-origin should be positioned

within the region with significant spin density and perturbed spin densities. In analogy to the ECC whose position r_{ECC} can be computed using the following equation:[18]

$$r_{\text{ECC},p} = \frac{1}{n_{\text{el}}} \sum_{\mu\nu} P_{\mu\nu} \langle \mu | \hat{r}_p | \nu \rangle, \quad p \in \{x, y, z\}, \quad (17)$$

where n_{el} is the number of electrons, we therefore propose the use of the spin density center (SDC). The equation used for computing the position of the SDC is given by

$$r_{\text{SDC},p} = \frac{1}{\text{Tr}(|P^{\alpha-\beta}|S)} \sum_{\mu\nu} |P_{\mu\nu}^{\alpha-\beta}| \langle \mu | \hat{r}_p | \nu \rangle, \quad p \in \{x, y, z\}. \quad (18)$$

The density matrix P in Eq. (17) is substituted by the absolute values of the spin density matrix $P^{\alpha-\beta}$. The trace of the matrix product of $|P^{\alpha-\beta}|$ with the overlap matrix S ensures the correct normalization. In the following, we compare the g-shifts obtained with GIAOs to the g-shifts obtained with both the ECC and the SDC as common gauge-origins.

In Fig. 6, computed isotropic g-shifts of alkane radicals with increasing chain length are shown. The radical center is at the end of the alkane chain. Both the isotropic g-shift obtained with GIAOs and with the SDC as common gauge-origin quickly converge with chain length; contrarily, the isotropic g-shifts computed with the ECC as gauge-origin show a wrong behavior with a linear increase for alkane radicals with more than five carbon atoms. This qualitatively different behavior between the gauge-origins ECC and SDC can be interpreted as follows: The alkane radicals have a strongly localized spin density at the radical center (which is at the end of the chain); in the calculation employing the SDC, the gauge-origin is located close to the terminal carbon atom at the radical center, and the region containing the spin density is well described. The rapid convergence of the isotropic g-shifts with chain length in the calculations with GIAOs and with the SDC reflects the locality of the contributions to the g-tensor in the alkane radicals. The agreement of the isotropic g-shifts computed with GIAOs and with the SDC as common gauge-origin is quite good in these calculations; the small remaining deviations can originate from the accelerated basis convergence in the GIAO calculation or

FIG. 5. Absolute values of the entries in the spin density matrix and the perturbed spin density matrices for the calculation of the $\cdot$O$-$[CH$_2$]$_{18}$$-CH_3$ molecule (def2-SVP basis set). Only matrix elements above a threshold of 10^{-5} are shown. The basis functions centered on the oxygen atom are located in the upper left corner.
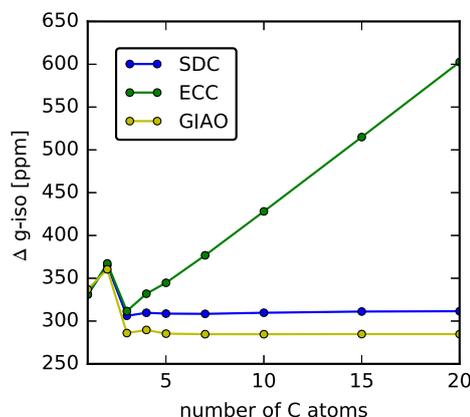
FIG. 6. Isotropic g-shifts of alkane radicals of increasing chain length with a terminal CH$_2$ group. The computations were carried out at the UDFT/B3LYP-level with a def2-TZVP basis. The GIAO ansatz is compared to the common gauge-origins SDC and ECC.

from a minor gauge error within the SDC calculation. In contrast to the SDC, the ECC is by construction always located centrally within the molecule (the ECC can even be viewed as one possible definition of the center of the molecule); for these calculations, it means that the ECC is for growing chain length increasingly far apart from the region with significant spin density. Hence larger and larger gauge errors are introduced in the calculation leading to a linear increase in the isotropic g-shift. For the $C_{20}H_{41}$ radical, the isotropic g-shift obtained with the ECC deviates 318 ppm (111.6%) from the GIAO value.

In Table V, computed g-shifts for other molecules with a single localized spin center are presented. The g-shifts obtained with the SDC as gauge-origin are also for these molecules significantly closer to the GIAO reference values than the ECC g-shifts. For the calculations shown in Table V, the mean absolute deviations in the isotropic g-shifts compared to the GIAO values are 289 ppm for the ECC and 26 ppm for the SDC. The average relative deviations in the isotropic g-shifts amount to 18.8% for the ECC and 2.9% for the SDC. The improvement is especially significant for the smaller g-shifts like in the ·NH—[CH$_2$—O]$_3$—CH$_3$—OH or the ·O—[CH$_2$]$_{18}$—CH$_3$ molecule where the smallest g-shifts obtained with the ECC have opposite sign. For even larger molecules, one can expect even more pronounced improvements by the SDC compared to the ECC as gauge-origin. Irrespective of the good agreement between the g-shifts obtained

with the SDC and the g-shifts in a GIAO basis, the results show that the influence of the gauge can be substantial in these calculations on medium-sized molecules and significant errors are introduced with a sub-optimal gauge origin such as the ECC. This can be seen from several examples in which the ECC leads to rather large errors like for Δg_1 of the ·NH—[CH$_2$—O]$_7$—CH$_3$—OH molecules with a deviation of 1561 ppm (37.9%).

For the substituted toluene radical in Table V, the deviation of the SDC values from the GIAO values is larger than for the other systems (14.7% deviation in the isotropic g-shift). This is expected due to the delocalization of the spin density over the aromatic ring and the attached CH$_2$ group. Nevertheless, the differences of the SDC values to the GIAO results are still smaller than those of the ECC values with a deviation of 50.2% in the isotropic g-shift.

With LiH$^+$ and NaF$^+$, Table V also includes two molecules which have a relatively large separation of ECC and SDC despite their small extent (1.338 Å separation for LiH$^+$; 1.196 Å separation for NaF$^+$). Also for these molecules, the SDC values are closer to the g-shifts obtained in a basis of GIAOs than the ECC values. In general, for small molecules, the differences between SDC and ECC can be expected to be small as they are often close to each other; for symmetric molecules, they might even coincide (e.g., for the O$_2$ molecule).

TABLE V. g-tensor calculations with different gauge-origins (SDC or ECC) or GIAOs on radicals with a single spin-center. The employed basis set is def2-TZVP in all cases.

Molecule	Gauge	Δg_1 (ppm)	Δg_2 (ppm)	Δg_3 (ppm)	Δg_{iso} (ppm)
MTSL	ECC	7 147	4 295	−44	3 799
	SDC	7 139	3 950	−163	3 642
	GIAO	7 084	3 805	−226	3 554
·NH−[CH$_2$−O]$_7$−CH$_2$−OH	ECC	5 680	1 505	207	2 464
	SDC	4 137	1 509	−132	1 838
	GIAO	4 119	1 540	−151	1 836
·O−[CH$_2$]$_{18}$−CH$_3$	ECC	65 816	7 382	292	24 497
	SDC	65 935	6 494	−143	24 095
	GIAO	66 021	6 443	−175	24 096
·Cys−Gly$_4$	ECC	219 431	17 158	83	78 891
	SDC	220 569	17 268	15	79 284
	GIAO	220 519	17 231	17	79 256
	ECC	774	491	274	513
	SDC	557	553	−28	361
	GIAO	552	549	−81	340
	ECC	786	593	45	475
	SDC	619	520	−52	362
	GIAO	566	504	−122	316
LiH$^+$	ECC	−36	−36	−36	−36
	SDC	−36	−39	−39	−38
	GIAO	−36	−39	−39	−38
NaF$^+$	ECC	99 018	69 225	−265	55 993
	SDC	99 018	68 179	−291	55 635
	GIAO	99 018	68 231	−288	55 654
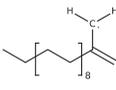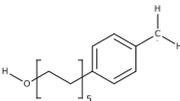
In summary, we have shown that the choice of gauge has a large influence on the g-shifts in medium-sized molecules, and common gauge-origin approaches can lead to significant errors compared to the gauge-origin independent GIAO values. A common gauge-origin that is chosen according to the spin density distribution like our proposed SDC is better able to account for the local nature of the g-tensor than the ECC. The g-shifts obtained with the SDC as gauge-origin agree much better with GIAO values than the g-shifts obtained with the ECC. Therefore the SDC can be used as a pragmatic common gauge-origin in larger molecules with a single localized spin center if no suitable GIAO implementation is available. However, we want to stress that the SDC does not represent a generally applicable solution to the gauge-origin problem as shown in Sec. IV D in calculations on molecules with multiple spin centers.

One minor advantage of a common gauge-origin approach involving the SDC is a small speed-up of the calculation. The gauge-origin independence provided by the GIAO approach comes at the price of slightly increased computational cost because it requires the computation of several integral derivatives with respect to the magnetic field; the integrals needed for the magnetic field derivative of the SOMF matrix [Eq. (16)] are especially costly. In our preliminary implementation, the computation of the SOMF matrix requires 35.5% of the total calculation time (calculation on alkane-20 radical/def2-TZVP basis); with GIAOs, this amounts to 53.4% (including the magnetic field derivative of the SOMF matrix). In these calculations, the use of GIAOs increased the total computation time by 71.1%. However, this additional effort might be substantially reduced by the resolution-of-the-identity (RI) approximation[50(a),70–72] which we currently do not employ. In a fully optimized implementation, we expect a significantly smaller overhead due to GIAOs. Therefore we recommend to use distributed gauge-origin methods if possible also for larger molecules with a single spin center as they provide fully gauge-origin independent results.

D. Gauge-origin dependence in molecules with multiple spin centers

As shown in Sec. IV C, a suitably chosen common gauge-origin can allow for reasonably accurate g-tensor computations also for larger molecules with a single, well localized spin center. However, there are also many examples for molecules with multiple spin centers or significantly delocalized spin density. Much larger errors due to gauge-origin dependence can be expected for molecules of this kind as no common gauge-origin can be positioned close to all sites of the molecule with significant (perturbed) spin density. This is confirmed by the illustrative calculations shown on molecules with two or three spin centers (Fig. 7 and Table VI).

In Fig. 7, isotropic g-shifts of the $\cdot NH-[CH_2]_x-\dot{C}H_2$ (x varied from 1 to 15) molecules are shown. Both the SDC and the ECC as gauge-origins lead to significant deviations from the GIAO values; the agreement with the GIAO g-shifts deteriorates with increased spacing between the two spin centers. For x = 15, the deviations in the isotropic g-shift are 392 ppm/34.1% (ECC) and 403 ppm/35.0% (SDC). In these molecules, the SDC is — similarly as the ECC — located
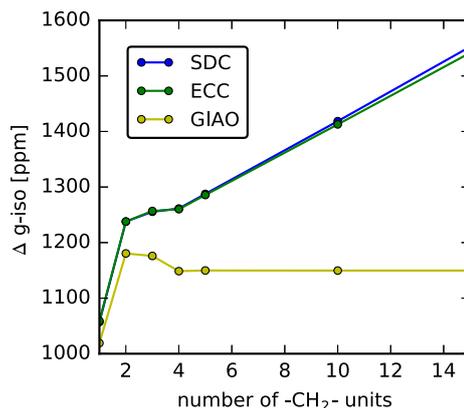
FIG. 7. Isotropic g-shifts of radicals of the type $\cdot NH-[CH_2]_x-\dot{C}H_2$ with triplet spin state, where x is varied from 1 to 15. The computations were carried out at the UDFT/B3LYP-level and with a def2-TZVP basis. The GIAO ansatz is compared to the common gauge-origins SDC and ECC.

relatively centrally between the two spin centers. For both gauge-origins, the increase of the distance between the gauge-origin and the spin centers with increasing chain length causes an increase in the gauge error which explains the large deviations from the GIAO numbers. After an irregular behavior for the smallest chain lengths, the deviations are growing linearly; one could therefore easily extrapolate the size of the error for even larger spacing between the spin centers. By contrast, the isotropic g-shift converges rapidly in a GIAO basis due to the locality of the contributions from the two spin centers.

In Table VI, results from g-tensor calculations on several other molecules with two or three spin centers are shown. For the presented molecules with radical centers on sulfur atoms, the gauge-origin dependence is mostly negligible. The same holds for the calculations on the $\cdot Cys-Gly_4$ radical shown in Table V. We suspect that due to the dominance of the atomic contributions of the sulfur the g-shifts are both significantly larger and less sensitive to the gauge than for the other analyzed systems; in the limiting case of a single atom, the g-shifts are independent of the gauge as discussed in Ref. 73 for the related theory of NMR shielding tensors. More pronounced differences between the GIAO values and the common gauge-origin approaches can be seen for the g-shifts of the other molecules. For the largest g-shift of $\cdot Ala-Gly-Lys\cdot$ di-radical, the difference to the GIAO g-shifts amounts to 620 ppm/23.6% (ECC) and 517 ppm/19.6% (SDC). The Δg_2 value of the $\cdot O-[CH_2]_{18}-NH$ molecule differs by 721 ppm/19.0% (ECC) and 718 ppm/18.9% (SDC) from the GIAO result.

Because of the demonstrated low reliability of common-gauge-origin approaches for molecules with several, spatially distant spin centers, we recommend to use distributed gauge-origin methods like GIAOs for molecules of this kind.

We have on purpose not shown calculations on molecules with extensively delocalized spin density, as in many cases multi-reference methods might be necessary to properly describe these systems, and the reliability of the DFT ansatz is questionable. Still, the findings of this section should be transferable to this class of molecules because it is not possible to position the entire region with significant spin density and perturbed spin densities in close vicinity of a common

TABLE VI. g-tensor calculations on molecules with two or three spin centers. All calculations were done with UDFT/B3LYP and the def2-TZVP basis set. Values obtained with the SDC or ECC as gauge-origins are compared to the GIAO results.

Molecule	Gauge	Δg_1 (ppm)	Δg_2 (ppm)	Δg_3 (ppm)	Δg_{iso} (ppm)
$O{=}\dot{C}{-}[CH_2]_{17}{-}C{\equiv}C\cdot$	ECC	1 550	−92	−3268	−603
	SDC	1 487	−96	−3248	−619
	GIAO	1 353	−118	−3253	−673
$\cdot NF{-}CH_2{-}[CH{=}CH{-}CH_2]_5{-}CH_2{-}CH{=}\dot{C}H$	ECC	4 167	1 788	−35	1 973
	SDC	4 197	1 786	−28	1 985
	GIAO	3 704	1 518	−155	1 689
$\cdot Ala{-}Gly_2{-}Lys\cdot$	ECC	3 252	1 031	113	1 465
	SDC	3 149	1 032	111	1 431
	GIAO	2 632	887	−32	1 162
$\cdot NH{-}[CH_2{-}O]_7{-}CH_2{-}S\cdot$	ECC	108 751	9 710	−100	39 454
	SDC	108 742	9 710	−100	39 451
	GIAO	108 747	9 710	−100	39 452
$\cdot O{-}[CH_2]_{18}{-}\dot{N}H$	ECC	35 012	4 516	1 108	13 545
	SDC	35 002	4 513	1 110	13 542
	GIAO	34 749	3 795	561	13 035
	ECC	138 717	20 958	2 980	54 218
	SDC	138 718	20 968	2 975	54 220
	GIAO	138 658	20 575	2 805	54 013
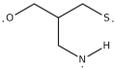
gauge-origin. Consequently we also recommend the use of a distributed gauge-origin ansatz for molecules with delocalized spin density.

V. CONCLUSIONS

We provided a detailed study of the gauge-origin dependence in g-tensor calculations at the DFT level of theory using the spin-orbit mean field ansatz. Our findings show that the influence of the gauge is only negligible for small molecules and not for larger molecular systems. For our test set of small molecules, we found a pronounced dependence of the g-shifts on the gauge; only a few Ångström displacements of the molecules from the gauge-origin are necessary to introduce errors that significantly exceed the basis set errors. We also analyzed to what extent GIAOs accelerate the basis set convergence in g-tensor calculations and found the effect to be negligible for our test set. It was shown in further calculations on medium-sized molecules that the influence of the gauge can be large in these systems and substantial errors can be introduced by common gauge-origin approaches. This renders distributed gauge-origin methods like GIAOs the preferred approach for larger systems. For extended molecules with a single localized spin center, the spin density center (SDC) was proposed as a common gauge-origin that takes the locality of the g-tensor into account; in contrast to the commonly employed electronic charge centroid (ECC) as gauge-origin, the SDC was shown to give reasonable agreement with the GIAO values for molecules of this kind. For more general situations such as molecules with multiple spin centers, neither the ECC nor the SDC as common gauge-origins lead to reliable and accurate g-shifts; for these cases, distributed gauge-origin approaches like GIAOs were

shown to be essential. Although this study was based entirely on DFT calculations, it is highly likely that the presented findings are transferable to other quantum-chemical methods like wave-function based correlation methods. One can expect that gauge-origin dependence of very similar magnitude occurs in these methods as it is caused by basis set incompleteness and not by method specifics.

SUPPLEMENTARY MATERIAL

See supplementary material for all computed g-shifts.

[1] J. Gauss, in *Modern Methods and Algorithms of Quantum Chemistry*, NIC Series Vol. 3, 2nd ed., edited by J. Grotendorst (John von Neumann Institute for Computing, Jülich, 2000), p. 541.
[2] F. London, J. Phys. Radium **8**, 397 (1937).
[3] R. Ditchfield, J. Chem. Phys. **65**, 3123 (1976).
[4] K. Wolinski, J. F. Hinton, and P. Pulay, J. Am. Chem. Soc. **112**, 8251 (1990).
[5] C. Edmiston and K. Ruedenberg, Rev. Mod. Phys. **35**, 457 (1963).
[6] W. Kutzelnigg, Isr. J. Chem. **19**, 193 (1980).
[7] M. Schindler and W. Kutzelnigg, J. Chem. Phys. **76**, 1919 (1982).
[8] A. E. Hansen and T. D. Bouman, J. Chem. Phys. **82**, 5035 (1985).
[9] O. L. Malkina, J. Vaara, B. Schimmelpfennig, M. Munzarová, V. G. Malkin, and M. Kaupp, J. Am. Chem. Soc. **122**, 9206 (2000).
[10] G. Lushington and F. Grein, Theor. Chim. Acta **93**, 259 (1996).

[11]D. Jayatilaka, J. Chem. Phys. **108**, 7587 (1998).

[12]E. van Lenthe, P. E. Wormer, and A. van der Avoird, J. Chem. Phys. **107**, 2488 (1997).

[13]G. Lushington, P. Bündgen, and F. Grein, Int. J. Quantum Chem. **55**, 377 (1995).

[14]A. J. Sadlej, Collect. Czech. Chem. Commun. **53**, 1995 (1988).

[15]M. Kaupp, T. Gress, R. Reviakine, O. L. Malkina, and V. G. Malkin, J. Phys. Chem. B **107**, 331 (2003).

[16]S. Patchkovskii, R. Strong, C. Pickard, and S. Un, J. Chem. Phys. **122**, 214101 (2005).

[17]G. Lushington and F. Grein, Int. J. Quantum Chem. **60**, 1679 (1996).

[18]A. Luzanov, E. Babich, and V. Ivanov, J. Mol. Struct. **311**, 211 (1994).

[19]H. Bolvin, ChemPhysChem **7**, 1575 (2006).

[20]I. Ciofini, R. Reviakine, A. Arbuznikov, and M. Kaupp, Theor. Chem. Acc. **111**, 132 (2004).

[21]M. Witwicki, A. R. Jaszewski, J. Jezierska, M. Jerzykiewicz, and A. Jezierski, Chem. Phys. Lett. **462**, 300 (2008).

[22]B. Frecus, Z. Rinkevicius, and H. Ågren, Phys. Chem. Chem. Phys. **15**, 10466 (2013).

[23]M. Witwicki and J. Jezierska, J. Phys. Chem. B **115**, 3172 (2011).

[24]M. Bühl, S. E. Ashbrook, D. M. Dawson, R. A. Doyle, P. Hrobárik, M. Kaupp, and I. A. Smellie, Chem. - Eur. J. **22**, 15328 (2016).

[25]S. M. Mattar, J. Phys. Chem. B **108**, 9449 (2004).

[26]M. Engström, R. Owenius, and O. Vahtras, Chem. Phys. Lett. **338**, 407 (2001).

[27]Q. Daniel, P. Huang, T. Fan, Y. Wang, L. Duan, L. Wang, F. Li, Z. Rinkevicius, F. Mamedov, M. Ahlquist, S. Styring, and L. Sun, Coord. Chem. Rev. **346**, 206 (2017).

[28]G. Lushington and F. Grein, J. Chem. Phys. **106**, 3292 (1997).

[29]O. Vahtras, B. Minaev, and H. Ågren, Chem. Phys. Lett. **281**, 186 (1997).

[30]P. J. Bruna, G. H. Lushington, and F. Grein, Chem. Phys. **225**, 1 (1997).

[31]P. J. Bruna and F. Grein, Int. J. Quantum Chem. **77**, 324 (2000).

[32]M. Engström, O. Vahtras, and H. Ågren, Chem. Phys. **243**, 263 (1999).

[33]M. Engström, B. Minaev, O. Vahtras, and H. Ågren, Chem. Phys. **237**, 149 (1998).

[34]P. Bruna, G. Lushington, and F. Grein, J. Mol. Struct.: THEOCHEM **527**, 139 (2000).

[35]S. Brownridge, F. Grein, J. Tatchen, M. Kleinschmidt, and C. M. Marian, J. Chem. Phys. **118**, 9552 (2003).

[36]P. J. Bruna and F. Grein, Chem. Phys. **276**, 1 (2002).

[37]S. Koseki, M. W. Schmidt, and M. S. Gordon, J. Phys. Chem. **96**, 10768 (1992).

[38]S. Pelties, D. Herrmann, B. de Bruin, F. Hartl, and R. Wolf, Chem. Commun. **50**, 7014 (2014).

[39]S. Schott, E. R. McNellis, C. B. Nielsen, H.-Y. Chen, S. Watanabe, H. Tanaka, I. McCulloch, K. Takimiya, J. Sinova, and H. Sirringhaus, Nat. Commun. **8**, 15200 (2017).

[40]A. Komarovskikh, A. Dmitriev, V. Nadolinny, and Y. Palyanov, Diamond Relat. Mater. **76**, 86 (2017).

[41]T. U. Nick, K. R. Ravichandran, J. Stubbe, M. Kasanmascheff, and M. Bennati, Biochemistry **56**, 3647 (2017).

[42]S. A. Thomson, J. Niklas, K. L. Mardis, C. Mallares, I. D. Samuel, and O. G. Poluektov, J. Phys. Chem. C **121**, 22707 (2017).

[43]A. Giard, B. Xerri, I. Ciofini, and D. Berthomieu, Chem. Phys. Lett. **614**, 226 (2014).

[44]O. Gromov, E. Golubeva, G. Zhidomirov, and M. Y. Mel'nikov, Dokl. Phys. Chem. **451**, 154–156 (2013).

[45]A. Massolle, T. Dresselhaus, S. Eusterwiemann, C. Doerenkamp, H. Eckert, A. Studer, and J. Neugebauer, Phys. Chem. Chem. Phys. **20**, 7661–7675 (2018).

[46]K. Kim and K. Jordan, J. Phys. Chem. **98**, 10089 (1994).

[47]A. Perera, J. Gauss, P. Verma, and J. A. Morales, J. Chem. Phys. **146**, 164104 (2017).

[48]G. Schreckenbach and T. Ziegler, J. Phys. Chem. A **101**, 3388 (1997).

[49]J. E. Harriman, *Theoretical Foundations of Electron Spin Resonance: Physical Chemistry: A Series of Monographs* (Academic Press, 2013).

[50](a) F. Neese, J. Chem. Phys. **122**, 034107 (2005); (b) B. A. Hess, C. M. Marian, U. Wahlgren, and O. Gropen, Chem. Phys. Lett. **251**, 365 (1996).

[51]F. Neese, J. Chem. Phys. **115**, 11080 (2001).

[52]J. Kussmann and C. Ochsenfeld, J. Chem. Phys. **127**, 054103 (2007).

[53]J. Kussmann and C. Ochsenfeld, J. Chem. Phys. **138**, 134114 (2013).

[54]J. Kussmann and C. Ochsenfeld, J. Chem. Theory Comput. **11**, 918 (2015).

[55]P. Pulay, J. Comput. Chem. **3**, 556 (1982).

[56]V. I. Lebedev, USSR Comput. Math. Math. Phys. **16**, 10 (1976).

[57]O. Treutler and R. Ahlrichs, J. Chem. Phys. **102**, 346 (1995).

[58]F. Weigend and R. Ahlrichs, Phys. Chem. Chem. Phys. **7**, 3297 (2005).

[59]D. Rappoport and F. Furche, J. Chem. Phys. **133**, 134105 (2010).

[60]D. J. Henry, C. J. Parkinson, and L. Radom, J. Phys. Chem. A **106**, 7927 (2002).

[61]R. Sure and S. Grimme, J. Comput. Chem. **34**, 1672 (2013).

[62]F. Neese, Wiley Interdiscip. Rev.: Comput. Mol. Sci. **2**, 73 (2012).

[63]F. Neese, Wiley Interdiscip. Rev.: Comput. Mol. Sci. **8**, e1327 (2018).

[64]A. Martorana, G. Bellapadrona, A. Feintuch, E. Di Gregorio, S. Aime, and D. Goldfarb, J. Am. Chem. Soc. **136**, 13458 (2014).

[65]S. Dunkel, L. Pulagam, H.-J. Steinhoff, and J. Klare, Phys. Chem. Chem. Phys. **17**, 4875 (2015).

[66]M. Oranges, R. Guzzi, D. Marsh, and R. Bartucci, Chem. Phys. Lipids **212**, 130 (2018).

[67]B. Endeward, A. Marko, V. Denysenkov, S. T. Sigurdsson, and T. Prisner, *Methods in Enzymology* (Elsevier, 2015), Vol. 564, pp. 403–425.

[68]J. Oddershede and J. Geertsen, J. Chem. Phys. **92**, 6036 (1990).

[69]J. Gauss, M. Kállay, and F. Neese, J. Phys. Chem. A **113**, 11541 (2009).

[70]J. L. Whitten, J. Chem. Phys. **58**, 4496 (1973).

[71]B. I. Dunlap, J. Connolly, and J. Sabin, J. Chem. Phys. **71**, 3396 (1979).

[72]O. Vahtras, J. Almlöf, and M. Feyereisen, Chem. Phys. Lett. **213**, 514 (1993).

[73]O. L. Malkina, S. Komorovský, L. Visscher, and V. G. Malkin, J. Chem. Phys. **134**, 086101 (2011).

Supplementary Information for: Gauge-origin dependence in electronic g-tensor calculations

Michael Glasbrenner, Sigurd Vogler, and Christian Ochsenfeld

Chair of Theoretical Chemistry and Center for Integrated Protein Science Munich (CIPSM), Department of Chemistry, University of Munich (LMU), Butenandtstr. 7, 81377 Munich, Germany

May 4, 2018

Contents

1 g-shifts

1.1 Sections "Significance of the gauge-origin dependence in g-tensor calculations" and "Influence of GIAOs on the basis set convergence"

Table 1: g-shifts in ppm for small molecules. Different gauges are employed : center of mass ("COM"), a gauge origin at (-10.0, -10.0 -10.0) ("m10") or a GIAO basis.

molecule	basis set	gauge	Δg_1 [ppm]	Δg_2 [ppm]	Δg_3 [ppm]
BeH	def2-qzvppd	COM	-40	-167	-167
CCH	def2-qzvppd	COM	277	277	-127
CH_2CH_3	def2-qzvppd	COM	686	536	-90
CH_2OH	def2-qzvppd	COM	2173	948	-144
CH_2	def2-qzvppd	COM	210	197	-67
CH_3	def2-qzvppd	COM	578	578	-90
$CHCH_2$	def2-qzvppd	COM	604	-98	-683
CHO	def2-qzvppd	COM	2340	-201	-7373
CH	def2-qzvppd	COM	1309	-158	-15408
$COCH_3$	def2-qzvppd	COM	2267	-254	-6544
NH_2	def2-qzvppd	COM	5037	1505	-153
NH	def2-qzvppd	COM	1287	1287	-107
NO	def2-qzvppd	COM	3302	-381	-110959
OCH_3	def2-qzvppd	COM	47607	6061	-195
OH	def2-qzvppd	COM	55799	5574	-217
ONO	def2-qzvppd	COM	3604	-624	-11072
OOH	def2-qzvppd	COM	27585	5529	-279
PH_2	def2-qzvppd	COM	15108	5063	-32
SiH_2	def2-qzvppd	COM	1230	1034	-498
SiH_3	def2-qzvppd	COM	2285	2285	-78
AlO	def2-qzvppd	COM	1076	1076	-125
AsO_3^{2-}	def2-qzvppd	COM	9325	9325	2783
BO	def2-qzvppd	COM	-69	-1742	-1742
BS	def2-qzvppd	COM	-84	-9099	-9099
C_3H_5	def2-qzvppd	COM	752	631	-73
CF_3Br^-	def2-qzvppd	COM	53176	53176	-465
CF_3Cl^-	def2-qzvppd	COM	11696	11696	-497
CH_4^+	def2-qzvppd	COM	23796	2599	-86
Continued on next page					

Table 1 – continued from previous page

molecule	basis set	gauge	Δg_1 [ppm]	Δg_2 [ppm]	Δg_3 [ppm]
ClO_2	def2-qzvppd	COM	15108	12793	-453
ClO_3	def2-qzvppd	COM	7316	7316	1105
CO_2^-	def2-qzvppd	COM	1041	-632	-4678
CO_3^-	def2-qzvppd	COM	11733	11733	2997
CO^+	def2-qzvppd	COM	-134	-2535	-2535
GeH_3	def2-qzvppd	COM	16762	16762	-107
H_2CO^+	def2-qzvppd	COM	5677	262	97
KrF	def2-qzvppd	COM	38364	38364	-307
MgF	def2-qzvppd	COM	-3	-1742	-1742
NF2	def2-qzvppd	COM	6586	3949	-644
NF_3^+	def2-qzvppd	COM	7109	7109	-563
NH_3^+	def2-qzvppd	COM	1657	1657	-156
NO_2	def2-qzvppd	COM	3836	-638	-11239
NO_3	def2-qzvppd	COM	15197	15197	343
O_3^-	def2-qzvppd	COM	17547	10452	-502
SO_2^-	def2-qzvppd	COM	6376	5072	-278
SO_3^-	def2-qzvppd	COM	2534	2534	169
BeH	def2-svp	GIAO	-42	-151	-151
CCH	def2-svp	GIAO	195	195	-128
CH_2CH_3	def2-svp	GIAO	599	434	-108
CH_2OH	def2-svp	GIAO	2053	846	-168
CH_2	def2-svp	GIAO	163	137	-73
CH_3	def2-svp	GIAO	481	481	-92
$CHCH_2$	def2-svp	GIAO	457	-117	-630
CHO	def2-svp	GIAO	2102	-224	-6943
CH	def2-svp	GIAO	1131	-175	-14448
$COCH_3$	def2-svp	GIAO	2075	-275	-6330
NH_2	def2-svp	GIAO	4651	1260	-153
NH	def2-svp	GIAO	1109	1109	-110
NO	def2-svp	GIAO	3117	-385	-124152
OCH_3	def2-svp	GIAO	48468	5437	-214
OH	def2-svp	GIAO	60473	4950	-210
ONO	def2-svp	GIAO	3422	-574	-10675
OOH	def2-svp	GIAO	26857	5077	-286
PH_2	def2-svp	GIAO	13370	4506	-23
SiH_2	def2-svp	GIAO	1099	941	-413
SiH_3	def2-svp	GIAO	2125	2125	-85
AlO	def2-svp	GIAO	1230	1230	-125
AsO_3^{2-}	def2-svp	GIAO	12798	12798	2361
BO	def2-svp	GIAO	-69	-1574	-1574
BS	def2-svp	GIAO	-83	-8358	-8358
C_3H_5	def2-svp	GIAO	511	510	-111
CF_3Br^-	def2-svp	GIAO	50377	50377	-474
CF_3Cl^-	def2-svp	GIAO	12016	12016	-488
CH_4^+	def2-svp	GIAO	27516	2379	-89
ClO_2	def2-svp	GIAO	17302	14141	-548
ClO_3	def2-svp	GIAO	9178	9178	630
CO_2^-	def2-svp	GIAO	1507	-629	-5194
CO_3^-	def2-svp	GIAO	11584	11584	2988
CO^+	def2-svp	GIAO	-136	-2255	-2255
GeH_3	def2-svp	GIAO	15005	15005	-211
H_2CO^+	def2-svp	GIAO	5168	164	72
KrF	def2-svp	GIAO	32609	32609	-301
MgF	def2-svp	GIAO	-5	-1661	-1661
NF2	def2-svp	GIAO	6617	3992	-653
NF_3^+	def2-svp	GIAO	6716	6716	-645
NH_3^+	def2-svp	GIAO	1475	1475	-156
NO_2	def2-svp	GIAO	3634	-587	-10828
NO_3	def2-svp	GIAO	14840	14840	547
O_3^-	def2-svp	GIAO	17659	10590	-519
SO_2^-	def2-svp	GIAO	9479	7461	-404
SO_3^-	def2-svp	GIAO	4036	4036	21
BeH	def2-svp	COM	-42	-150	-150
CCH	def2-svp	COM	195	195	-128
CH_2CH_3	def2-svp	COM	604	428	-95
CH_2OH	def2-svp	COM	2004	724	-146
CH_2	def2-svp	COM	152	132	-73
CH_3	def2-svp	COM	462	462	-98
$CHCH_2$	def2-svp	COM	496	-109	-606
CHO	def2-svp	COM	2056	-197	-6889
CH	def2-svp	COM	1094	-175	-14448
$COCH_3$	def2-svp	COM	2140	-268	-6278

Continued on next page

Table 1 – continued from previous page

molecule	basis set	gauge	Δg_1 [ppm]	Δg_2 [ppm]	Δg_3 [ppm]
NH_2	def2-svp	COM	4579	1157	-147
NH	def2-svp	COM	1062	1062	-110
NO	def2-svp	COM	3205	-357	-124152
OCH_3	def2-svp	COM	47937	5518	-200
OH	def2-svp	COM	60473	4805	-186
ONO	def2-svp	COM	3741	-558	-10660
OOH	def2-svp	COM	26796	5195	-271
PH_2	def2-svp	COM	12972	4104	-12
SiH_2	def2-svp	COM	1033	932	-442
SiH_3	def2-svp	COM	2062	2062	-89
AlO	def2-svp	COM	1070	1070	-125
AsO_3^{2-}	def2-svp	COM	13367	13367	2489
BO	def2-svp	COM	-69	-1607	-1607
BS	def2-svp	COM	-83	-8594	-8594
C_3H_5	def2-svp	COM	528	508	-74
CF_3Br^-	def2-svp	COM	49725	49725	-458
CF_3Cl^-	def2-svp	COM	11389	11389	-470
CH_4^+	def2-svp	COM	27180	2381	-85
ClO_2	def2-svp	COM	17661	15264	-489
ClO_3	def2-svp	COM	10056	10056	847
CO_2^-	def2-svp	COM	1555	-619	-5144
CO_3^-	def2-svp	COM	11914	11914	3502
CO^+	def2-svp	COM	-136	-2336	-2336
GeH_3	def2-svp	COM	14613	14613	-211
H_2CO^+	def2-svp	COM	5238	192	95
KrF	def2-svp	COM	33247	33247	-301
MgF	def2-svp	COM	-5	-1905	-1905
NF2	def2-svp	COM	6696	4033	-629
NF_3^+	def2-svp	COM	6818	6818	-609
NH_3^+	def2-svp	COM	1449	1449	-162
NO_2	def2-svp	COM	3966	-570	-10816
NO_3	def2-svp	COM	15243	15243	1046
O_3^-	def2-svp	COM	17780	10894	-498
SO_2^-	def2-svp	COM	9674	8063	-365
SO_3^-	def2-svp	COM	4379	4379	73
BeH	def2-svp	m10	-15	-88	-115
CCH	def2-svp	m10	640	555	-213
CH_2CH_3	def2-svp	m10	885	349	75
CH_2OH	def2-svp	m10	2787	-206	-446
CH_2	def2-svp	m10	176	-10	-117
CH_3	def2-svp	m10	462	462	-98
$CHCH_2$	def2-svp	m10	921	54	-578
CHO	def2-svp	m10	3430	-305	-4176
CH	def2-svp	m10	-97	-405	-14451
$COCH_3$	def2-svp	m10	3301	-381	-5488
NH_2	def2-svp	m10	2969	-142	-2059
NH	def2-svp	m10	732	-172	-1014
NO	def2-svp	m10	3679	-636	-124153
OCH_3	def2-svp	m10	47850	8793	-175
OH	def2-svp	m10	60557	666	248
ONO	def2-svp	m10	5578	-429	-8160
OOH	def2-svp	m10	25038	1984	-241
PH_2	def2-svp	m10	8545	-175	-524
SiH_2	def2-svp	m10	4123	2370	-100
SiH_3	def2-svp	m10	7360	6205	-1239
AlO	def2-svp	m10	2654	-5758	-8526
AsO_3^{2-}	def2-svp	m10	36236	31381	-2285
BO	def2-svp	m10	-15	-2080	-2134
BS	def2-svp	m10	176	-10993	-11250
C_3H_5	def2-svp	m10	1297	397	29
CF_3Br^-	def2-svp	m10	42320	41525	-1236
CF_3Cl^-	def2-svp	m10	18344	17384	-1421
CH_4^+	def2-svp	m10	26719	2453	-176
ClO_2	def2-svp	m10	17641	5836	-668
ClO_3	def2-svp	m10	16099	15221	-25
CO_2^-	def2-svp	m10	1117	-619	-6003
CO_3^-	def2-svp	m10	11914	11914	3502
CO^+	def2-svp	m10	374	-1107	-1617
GeH_3	def2-svp	m10	55387	46245	-9106
H_2CO^+	def2-svp	m10	7889	363	-39
KrF	def2-svp	m10	31177	31103	-373
MgF	def2-svp	m10	877	-5169	-6048
NF2	def2-svp	m10	9079	3809	-345

Table 1 – continued from previous page

molecule	basis set	gauge	Δg_1 [ppm]	Δg_2 [ppm]	Δg_3 [ppm]
NF_3^+	def2-svp	m10	8367	8252	-723
NH_3^+	def2-svp	m10	1449	1449	-162
NO_2	def2-svp	m10	2474	-686	-13728
NO_3	def2-svp	m10	15243	15243	1046
O_3^-	def2-svp	m10	15433	10472	-567
SO_2^-	def2-svp	m10	14809	7343	-468
SO_3^-	def2-svp	m10	8564	7852	-636
BeH	def2-tzvp	GIAO	-40	-154	-154
CCH	def2-tzvp	GIAO	167	167	-127
CH_2CH_3	def2-tzvp	GIAO	678	503	-106
CH_2OH	def2-tzvp	GIAO	2175	929	-168
CH_2	def2-tzvp	GIAO	198	173	-70
CH_3	def2-tzvp	GIAO	551	551	-89
$CHCH_2$	def2-tzvp	GIAO	517	-114	-682
CHO	def2-tzvp	GIAO	2256	-224	-7245
CH	def2-tzvp	GIAO	1256	-164	-15142
$COCH_3$	def2-tzvp	GIAO	2200	-275	-6536
NH_2	def2-tzvp	GIAO	4900	1453	-151
NH	def2-tzvp	GIAO	1240	1240	-107
NO	def2-tzvp	GIAO	3300	-402	-112426
OCH_3	def2-tzvp	GIAO	47755	5894	-213
OH	def2-tzvp	GIAO	56561	5406	-213
ONO	def2-tzvp	GIAO	3559	-638	-10980
OOH	def2-tzvp	GIAO	27277	5401	-300
PH_2	def2-tzvp	GIAO	15250	5044	-18
SiH_2	def2-tzvp	GIAO	1224	1031	-511
SiH_3	def2-tzvp	GIAO	2294	2294	-81
AlO	def2-tzvp	GIAO	1167	1167	-127
AsO_3^{2-}	def2-tzvp	GIAO	10627	10627	3371
BO	def2-tzvp	GIAO	-69	-1733	-1733
BS	def2-tzvp	GIAO	-84	-9322	-9322
C_3H_5	def2-tzvp	GIAO	590	559	-103
CF_3Br^-	def2-tzvp	GIAO	51125	51125	-479
CF_3Cl^-	def2-tzvp	GIAO	12155	12155	-508
CH_4^+	def2-tzvp	GIAO	24221	2487	-91
ClO_2	def2-tzvp	GIAO	16238	12888	-548
ClO_3	def2-tzvp	GIAO	7325	7325	953
CO_2^-	def2-tzvp	GIAO	1283	-668	-5144
CO_3^-	def2-tzvp	GIAO	11731	11731	2912
CO^+	def2-tzvp	GIAO	-135	-2514	-2514
GeH_3	def2-tzvp	GIAO	16501	16501	-116
H_2CO^+	def2-tzvp	GIAO	5517	249	75
KrF	def2-tzvp	GIAO	36640	36640	-304
MgF	def2-tzvp	GIAO	-3	-1651	-1651
NF2	def2-tzvp	GIAO	6548	3960	-679
NF_3^+	def2-tzvp	GIAO	6975	6975	-602
NH_3^+	def2-tzvp	GIAO	1600	1600	-153
NO_2	def2-tzvp	GIAO	3787	-652	-11142
NO_3	def2-tzvp	GIAO	14978	14978	257
O_3^-	def2-tzvp	GIAO	17795	10623	-557
SO_2^-	def2-tzvp	GIAO	8942	5544	-369
SO_3^-	def2-tzvp	GIAO	2612	2612	227
BeH	def2-tzvp	COM	-40	-151	-151
CCH	def2-tzvp	COM	155	155	-127
CH_2CH_3	def2-tzvp	COM	676	512	-93
CH_2OH	def2-tzvp	COM	2177	864	-149
CH_2	def2-tzvp	COM	193	171	-69
CH_3	def2-tzvp	COM	544	544	-92
$CHCH_2$	def2-tzvp	COM	545	-105	-668
CHO	def2-tzvp	COM	2248	-206	-7213
CH	def2-tzvp	COM	1237	-165	-15142
$COCH_3$	def2-tzvp	COM	2248	-261	-6506
NH_2	def2-tzvp	COM	4886	1409	-148
NH	def2-tzvp	COM	1221	1221	-107
NO	def2-tzvp	COM	3353	-379	-112426
OCH_3	def2-tzvp	COM	47485	5956	-198
OH	def2-tzvp	COM	56561	5357	-202
ONO	def2-tzvp	COM	3661	-614	-10973
OOH	def2-tzvp	COM	27234	5419	-283
PH_2	def2-tzvp	COM	15126	4889	-8
SiH_2	def2-tzvp	COM	1187	1033	-522
SiH_3	def2-tzvp	COM	2266	2266	-78

Continued on next page

Table 1 – continued from previous page

molecule	basis set	gauge	Δg_1 [ppm]	Δg_2 [ppm]	Δg_3 [ppm]
AlO	def2-tzvp	COM	1167	1167	-127
AsO_3^{2-}	def2-tzvp	COM	10348	10348	3892
BO	def2-tzvp	COM	-69	-1737	-1737
BS	def2-tzvp	COM	-84	-9338	-9338
C_3H_5	def2-tzvp	COM	556	459	-72
CF_3Br^-	def2-tzvp	COM	52769	52769	-477
CF_3Cl^-	def2-tzvp	COM	11907	11907	-506
CH_4^+	def2-tzvp	COM	24111	2508	-86
ClO_2	def2-tzvp	COM	16585	13118	-481
ClO_3	def2-tzvp	COM	7581	7581	1058
CO_2^-	def2-tzvp	COM	1358	-647	-5128
CO_3^-	def2-tzvp	COM	11881	11881	3186
CO^+	def2-tzvp	COM	-135	-2534	-2534
GeH_3	def2-tzvp	COM	16196	16196	-61
H_2CO^+	def2-tzvp	COM	5591	269	93
KrF	def2-tzvp	COM	37733	37733	-304
MgF	def2-tzvp	COM	-3	-1705	-1705
NF2	def2-tzvp	COM	6590	3953	-650
NF_3^+	def2-tzvp	COM	6967	6967	-576
NH_3^+	def2-tzvp	COM	1595	1595	-158
NO_2	def2-tzvp	COM	3890	-628	-11135
NO_3	def2-tzvp	COM	15144	15144	636
O_3^-	def2-tzvp	COM	17900	10781	-531
SO_2^-	def2-tzvp	COM	9530	5657	-319
SO_3^-	def2-tzvp	COM	2761	2761	366
BeH	def2-tzvp	m10	-38	-171	-173
CCH	def2-tzvp	m10	374	340	-161
CH_2CH_3	def2-tzvp	m10	737	472	52
CH_2OH	def2-tzvp	m10	2348	-204	-286
CH_2	def2-tzvp	m10	218	32	-146
CH_3	def2-tzvp	m10	544	544	-92
$CHCH_2$	def2-tzvp	m10	733	-42	-669
CHO	def2-tzvp	m10	2557	-316	-6106
CH	def2-tzvp	m10	646	-207	-15144
$COCH_3$	def2-tzvp	m10	2694	-298	-6266
NH_2	def2-tzvp	m10	3204	400	-113
NH	def2-tzvp	m10	797	545	-359
NO	def2-tzvp	m10	3252	-523	-112426
OCH_3	def2-tzvp	m10	47255	7053	-120
OH	def2-tzvp	m10	56586	2999	61
ONO	def2-tzvp	m10	4452	-630	-10283
OOH	def2-tzvp	m10	26154	3917	-309
PH_2	def2-tzvp	m10	11363	4083	40
SiH_2	def2-tzvp	m10	1431	1101	-823
SiH_3	def2-tzvp	m10	2619	2599	-99
AlO	def2-tzvp	m10	1070	1066	-131
AsO_3^{2-}	def2-tzvp	m10	23769	20934	1085
BO	def2-tzvp	m10	-57	-1946	-1958
BS	def2-tzvp	m10	-77	-9685	-9691
C_3H_5	def2-tzvp	m10	511	396	24
CF_3Br^-	def2-tzvp	m10	66382	65206	-1615
CF_3Cl^-	def2-tzvp	m10	13104	13055	-554
CH_4^+	def2-tzvp	m10	24218	2388	-157
ClO_2	def2-tzvp	m10	15408	12908	-393
ClO_3	def2-tzvp	m10	10185	9892	766
CO_2^-	def2-tzvp	m10	1178	-648	-5408
CO_3^-	def2-tzvp	m10	11881	11881	3186
CO^+	def2-tzvp	m10	-84	-2081	-2133
GeH_3	def2-tzvp	m10	29764	27576	-2217
H_2CO^+	def2-tzvp	m10	6679	279	159
KrF	def2-tzvp	m10	52699	51013	-1947
MgF	def2-tzvp	m10	119	-2504	-2625
NF2	def2-tzvp	m10	7891	3864	-402
NF_3^+	def2-tzvp	m10	7287	7281	-582
NH_3^+	def2-tzvp	m10	1595	1595	-158
NO_2	def2-tzvp	m10	3157	-657	-11851
NO_3	def2-tzvp	m10	15144	15144	636
O_3^-	def2-tzvp	m10	17786	10781	-609
SO_2^-	def2-tzvp	m10	7702	4725	-554
SO_3^-	def2-tzvp	m10	4709	4399	57

1.2 Section "Gauge-origin dependence in molecules with a single localized spin center"

Table 2: g-shifts in ppm for section "Gauge-origin dependence in molecules with a single localized spin center". Either a GIAO basis or a common gauge-origin (SDC or ECC) was used as indicated.

molecule	basis set	gauge	Δg_1 [ppm]	Δg_2 [ppm]	Δg_3 [ppm]
MTSL	def2-tzvp	ECC	7147	4295	-44
·NH-[CH$_2$-O]$_7$-CH$_2$-OH	def2-tzvp	ECC	5680	1505	207
·O-[CH$_2$]$_{18}$-CH$_3$	def2-tzvp	ECC	65816	7382	292
·Cys-Gly$_4$	def2-tzvp	ECC	219431	17158	83
·CH$_3$	def2-tzvp	ECC	543	543	-92
·CH$_2$-CH$_3$	def2-tzvp	ECC	683	514	-95
·CH$_2$-[CH$_2$]-CH$_3$	def2-tzvp	ECC	549	492	-106
·CH$_2$-[CH$_2$]$_2$-CH$_3$	def2-tzvp	ECC	580	497	-81
·CH$_2$-[CH$_2$]$_3$-CH$_3$	def2-tzvp	ECC	598	508	-72
·CH$_2$-[CH$_2$]$_5$-CH$_3$	def2-tzvp	ECC	639	536	-44
·CH$_2$-[CH$_2$]$_8$-CH$_3$	def2-tzvp	ECC	700	588	-4
·CH$_2$-[CH$_2$]$_{13}$-CH$_3$	def2-tzvp	ECC	802	695	48
·CH$_2$-[CH$_2$]$_{18}$-CH$_3$	def2-tzvp	ECC	905	819	84
LiH$^+$	def2-tzvp	ECC	-36	-36	-36
NaF$^+$	def2-tzvp	ECC	99018	69225	-265
MTSL	def2-tzvp	SDC	7139	3950	-163
·NH-[CH$_2$-O]$_7$-CH$_2$-OH	def2-tzvp	SDC	4137	1509	-132
·O-[CH$_2$]$_{18}$-CH$_3$	def2-tzvp	SDC	65935	6494	-143
·Cys-Gly$_4$	def2-tzvp	SDC	220569	17268	15
·CH$_3$	def2-tzvp	SDC	543	543	-92
·CH$_2$-CH$_3$	def2-tzvp	SDC	680	516	-100
·CH$_2$-[CH$_2$]-CH$_3$	def2-tzvp	SDC	544	486	-112
·CH$_2$-[CH$_2$]$_2$-CH$_3$	def2-tzvp	SDC	558	478	-106
·CH$_2$-[CH$_2$]$_3$-CH$_3$	def2-tzvp	SDC	561	475	-110
·CH$_2$-[CH$_2$]$_5$-CH$_3$	def2-tzvp	SDC	563	473	-111
·CH$_2$-[CH$_2$]$_8$-CH$_3$	def2-tzvp	SDC	566	473	-109
·CH$_2$-[CH$_2$]$_{13}$-CH$_3$	def2-tzvp	SDC	568	474	-108
·CH$_2$-[CH$_2$]$_{18}$-CH$_3$	def2-tzvp	SDC	568	474	-107
LiH$^+$	def2-tzvp	SDC	-36	-39	-39
NaF$^+$	def2-tzvp	SDC	99018	68179	-291
MTSL	def2-tzvp	GIAO	7084	3805	-226
·NH-[CH$_2$-O]$_7$-CH$_2$-OH	def2-tzvp	GIAO	4119	1540	-151
·O-[CH$_2$]$_{18}$-CH$_3$	def2-tzvp	GIAO	66021	6443	-175
·Cys-Gly$_4$	def2-tzvp	GIAO	220519	17231	17
·CH$_3$	def2-tzvp	GIAO	550	550	-89
·CH$_2$-CH$_3$	def2-tzvp	GIAO	682	507	-108
·CH$_2$-[CH$_2$]-CH$_3$	def2-tzvp	GIAO	526	469	-137
·CH$_2$-[CH$_2$]$_2$-CH$_3$	def2-tzvp	GIAO	537	465	-133
·CH$_2$-[CH$_2$]$_3$-CH$_3$	def2-tzvp	GIAO	534	460	-138
·CH$_2$-[CH$_2$]$_5$-CH$_3$	def2-tzvp	GIAO	533	460	-139
·CH$_2$-[CH$_2$]$_8$-CH$_3$	def2-tzvp	GIAO	533	460	-139
·CH$_2$-[CH$_2$]$_{13}$-CH$_3$	def2-tzvp	GIAO	533	460	-139
·CH$_2$-[CH$_2$]$_{18}$-CH$_3$	def2-tzvp	GIAO	534	460	-139
LiH$^+$	def2-tzvp	GIAO	-36	-39	-39
NaF$^+$	def2-tzvp	GIAO	99018	68231	-288
	def2-tzvp	ECC	786	593	45
	def2-tzvp	SDC	619	520	-52
	def2-tzvp	GIAO	566	504	-122
	def2-tzvp	ECC	774	491	274
	def2-tzvp	SDC	557	553	-28
	def2-tzvp	GIAO	552	549	-81
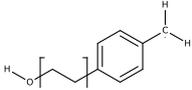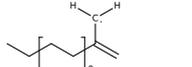
1.3 Section "Gauge origin dependence in molecules with multiple spin centers"

Table 3: g-shifts in ppm for section "Gauge-origin dependence in molecules with multiple spin centers". Either a GIAO basis or a common gauge-origin (SDC or ECC) was used as indicated.

molecule	basis set	gauge	Δg_1 [ppm]	Δg_2 [ppm]	Δg_3 [ppm]
O=$\dot{\text{C}}$-[CH$_2$]$_{17}$-C≡C·	def2-tzvp	ECC	1550	-92	-3268
·NF-CH$_2$-[CH=CH-CH$_2$]$_5$-CH$_2$-CH=$\dot{\text{C}}$H	def2-tzvp	ECC	4167	1788	-35
·Ala-Gly$_2$-Lys·	def2-tzvp	ECC	3252	1031	113
·NH-$\dot{\text{C}}$H$_2$	def2-tzvp	ECC	2296	887	-6
·NH-CH$_2$-$\dot{\text{C}}$H$_2$	def2-tzvp	ECC	2763	979	-28
·NH-[CH$_2$]$_2$-$\dot{\text{C}}$H$_2$	def2-tzvp	ECC	2647	944	179
·NH-[CH$_2$]$_3$-$\dot{\text{C}}$H$_2$	def2-tzvp	ECC	2812	901	67
·NH-[CH$_2$]$_4$-$\dot{\text{C}}$H$_2$	def2-tzvp	ECC	2688	913	256
·NH-[CH$_2$]$_9$-$\dot{\text{C}}$H$_2$	def2-tzvp	ECC	3155	954	129
·NH-[CH$_2$]$_{14}$-$\dot{\text{C}}$H$_2$	def2-tzvp	ECC	3251	917	457
·O-[CH$_2$]$_{18}$-$\dot{\text{N}}$H	def2-tzvp	ECC	35012	4516	1108
O=$\dot{\text{C}}$-[CH$_2$]$_{17}$-C≡C·	def2-tzvp	SDC	1487	-96	-3248
·NF-CH$_2$-[CH=CH-CH$_2$]$_5$-CH$_2$-CH=$\dot{\text{C}}$H	def2-tzvp	SDC	4197	1786	-28
·Ala-Gly$_2$-Lys·	def2-tzvp	SDC	3149	1032	111
·NH-$\dot{\text{C}}$H$_2$	def2-tzvp	SDC	2288	890	-7
·NH-CH$_2$-$\dot{\text{C}}$H$_2$	def2-tzvp	SDC	2764	978	-28
·NH-[CH$_2$]$_2$-$\dot{\text{C}}$H$_2$	def2-tzvp	SDC	2646	943	178
·NH-[CH$_2$]$_3$-$\dot{\text{C}}$H$_2$	def2-tzvp	SDC	2818	899	67
·NH-[CH$_2$]$_4$-$\dot{\text{C}}$H$_2$	def2-tzvp	SDC	2692	915	255
·NH-[CH$_2$]$_9$-$\dot{\text{C}}$H$_2$	def2-tzvp	SDC	3178	949	129
·NH-[CH$_2$]$_{14}$-$\dot{\text{C}}$H$_2$	def2-tzvp	SDC	3289	913	456
·O-[CH$_2$]$_{18}$-$\dot{\text{N}}$H	def2-tzvp	SDC	35002	4513	1110
·NH-[CH$_2$-O]$_7$-CH$_2$-S·	def2-tzvp	ECC	108751	9710	-100
·NH-[CH$_2$-O]$_7$-CH$_2$-S·	def2-tzvp	SDC	108742	9710	-100
O=$\dot{\text{C}}$-[CH$_2$]$_{17}$-C≡C·	def2-tzvp	GIAO	1353	-118	-3253
·NF-CH$_2$-[CH=CH-CH$_2$]$_5$-CH$_2$-CH=$\dot{\text{C}}$H	def2-tzvp	GIAO	3704	1518	-155
·Ala-Gly$_2$-Lys·	def2-tzvp	GIAO	2632	887	-32
·NH-$\dot{\text{C}}$H$_2$	def2-tzvp	GIAO	2196	901	-39
·NH-CH$_2$-$\dot{\text{C}}$H$_2$	def2-tzvp	GIAO	2623	996	-78
·NH-[CH$_2$]$_2$-$\dot{\text{C}}$H$_2$	def2-tzvp	GIAO	2468	948	112
·NH-[CH$_2$]$_3$-$\dot{\text{C}}$H$_2$	def2-tzvp	GIAO	2564	897	-15
·NH-[CH$_2$]$_4$-$\dot{\text{C}}$H$_2$	def2-tzvp	GIAO	2376	931	143
·NH-[CH$_2$]$_9$-$\dot{\text{C}}$H$_2$	def2-tzvp	GIAO	2579	927	-57
·NH-[CH$_2$]$_{14}$-$\dot{\text{C}}$H$_2$	def2-tzvp	GIAO	2378	928	143
·NH-[CH$_2$-O]$_7$-CH$_2$-S·	def2-tzvp	GIAO	108747	9710	-100
·O-[CH$_2$]$_{18}$-$\dot{\text{N}}$H	def2-tzvp	GIAO	34749	3795	561
	def2-tzvp	ECC	138717	20958	2980
	def2-tzvp	SDC	138718	20968	2975
	def2-tzvp	GIAO	138658	20575	2805
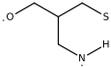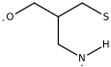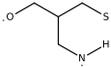
7.4 Manuscript IV: Accurate hyperfine coupling constants including electron correlation and dynamic effects

"Accurate Hyperfine Coupling Constants including Electron
Correlation and Dynamic Effects",
S. Vogler, J. C. B. Dietschreit, L. D. M. Peters, C. Ochsenfeld,
in preparation

The following manuscript analyzes the influence of both electron correlation and dynamic effects on the hyperfine coupling constants of organic radicals. Electron correlation is incorporated using the methods described in Paper **I** and **II**, whereas an *ab initio* molecular dynamics simulation is chosen to represent vibrational averaging. Our results indicate that both effects are important and that results close to experimental results can be obtained. Especially neglecting dynamic effects can in some cases deteriorate the results.

Accurate Hyperfine Coupling Constants including Electron Correlation and Dynamic Contributions

Sigurd Vogler, Johannes C. B. Dietschreit, Laurens D. M. Peters, and Christian Ochsenfeld*

Chair of Theoretical Chemistry and Center for Integrated Protein Science Munich (CIPSM), Department of Chemistry, University of Munich (LMU), Butenandtstr. 7, 81377 Munich, Germany

E-mail: christian.ochsenfeld@uni-muenchen.de

Abstract

The calculation of hyperfine coupling constants is a challenging task in balancing accuracy and computational effort. While previous work has shown both the importance of electron correlation and of molecular dynamic contributions, we present a thorough study simultaneously analyzing the influence of both on hyperfine coupling constants. To this end, we systematically study two organic radicals, namely dimethylamine and ethanal, proving the necessity for a high-level description of dynamic contributions as well as the large influence of electron correlation. Based on this study, we analyze the effect of both electron correlation and dynamic simulations on a set of 12 organic radicals, illustrating that both are vital for an accurate description and influence the *in silico* results on the same scale.

1

1   Introduction

Electron paramagnetic resonance (EPR) spectroscopy is an important tool for studying radicals. As a non-invasive method to characterize radicals it is indispensable in research tackling many biological systems.[1,2] However, the *ab initio* computation of the EPR parameters, namely the hyperfine coupling constants (HFCCs) and the g-tensors, remains challenging. These open-shell properties can be calculated using the unrestricted approach which incorporates spin polarization and delocalization, but this can lead to erratic results due to spin contamination. This can be improved by using a spin restriction within the restricted-unrestricted ansatz by Rinkevicius *et al.*,[3] where spin contamination is overcome while still including spin polarization. Nonetheless, the computationally less demanding unrestricted approach often leads to reliable results, especially for well localized radicals and when using density functional theory (DFT).[4–8]

Based on an unrestricted framework, further aspects need to be considered: EPR-specific basis sets have been demonstrated to be important,[9] and while DFT is often highly accurate, higher-order levels of theory taking electron correlation into account systematically are often beneficial. This includes methods such as second-order Møller-Plesset perturbation theory (MP2), double hybrid (DH)-DFT,[6,10] or coupled cluster approaches,[5,11–14] as well as different multi-reference ansätze.[15–17]

Nonetheless, even with elaborate methods, extensive basis sets, and in absence of spin contamination, the results can deviate from experimental values due to the neglect of dynamic contributions, i.e., vibrational averaging. This was shown in recent work by Massolle *et al.*[18] on verdazyl radicals where computational results at the DFT-level are improved by averaging over frames from a molecular dynamics simulation based on a quantum mechanically derived force field.[19] Similarly, studies on nitroxide radicals show that considering both vibrational averaging and solvent effects leads to more accurate results within the DFT framework.[20,21] The influence of molecular and intermolecular motion of explicitly solvated benzosemiquinone was studied in detail by Asher and Kaupp.[22] Nonetheless, solvent effects

are often small, and work by Rinkevicius *et al.*[23] shows that a description of the environment by means of molecular mechanics theory seems sufficient. A thorough investigation of the effect of the bending angle of the methyl radical, its incorporation within an *ab initio* molecular dynamics simulation,[24] and its solvation[25] also motivates the correct description of dynamic contributions. Similarly, significant ro-vibrational contributions were shown in the analysis of out-of-plane bending in H_2NO,[26] of dimethyl nitroxide,[27] and of other organic radicals.[28–30] Here, we also want to mention corresponding recent work in the computation of nuclear magnetic resonance shielding tensors by Grimme *et al.*,[31] that considers a set of conformers or rotamers to accurately describe flexible molecules in solution.

While both the effect of electron correlation and of dynamic contributions significantly improve the *in silico* results, their combined description is computationally cumbersome. A straightforward approach is to perform a molecular dynamics simulation and compute the EPR parameters for a set of frames. This requires the speed for computing the EPR parameters per frame to be small as to allow a sufficient number of frames to be computed for accurately incorporating the vibrational and rotational motion. The description of the correlation contribution by the cheapest wavefunction-based ansatz, MP2, is still expensive due to its conventionally large scaling behavior of $\mathcal{O}(N^5)$ as well as its large prefactor. This also applies to double-hybrid DFT that contains a second-order perturbation theory term analogous to MP2. The prefactor can be reduced by the resolution-of-the-identity (RI) approximation,[32–37] whereas linear scaling behavior can be achieved by a reformulation in local orbitals, e.g., in atomic orbitals (AO),[38] using distance-including integral estimates.[39,40] Analytic energy gradients at the MP2-level have been developed in the AO-basis.[41] By introducing the RI approximation and a Cholesky decomposition[42,43] in the computation of AO-MP2 energy gradients, we recently presented a low-scaling, low-prefactor implementation to compute HFCCs.[44] By computing only selected nuclei, the computational cost can be reduced further.[45] Using these methods, large-scale computations of HFCCs based on multiple frames from a molecular dynamics (MD) simulation are possible.

This work simultaneously analyzes the effect of both dynamic contributions and electron correlation on the HFCCs thus providing a computational protocol for calculating accurate HFCCs for large molecular systems using the efficient quantum chemical methods introduced above. We highlight the importance of accounting for dynamic contributions by investigating in depth the dependence of the HFCC on geometric parameters such as bond lengths, and bond and dihedral angles of two organic radicals, ethanale and dimethylamine. Subsequently, we analyze the contribution of electron correlation and dynamic contributions on isotropic HFCCs of a set of 12 organic radicals. Dynamic contributions are considered by computing a set of snapshots from an *ab initio* molecular dynamics (AIMD) simulation[46–48] of the radicals using the fast small basis set Hartree-Fock method (HF3c)[49] and the efficient three-fold corrected Perdew-Burke-Ernzerhoff generalized-gradient-approximation (PBEH3c),[50] thus providing their high-level and accurate description. Finally, we analyze the effect of electron correlation by comparing the isotropic HFCCs both at the Hartree-Fock (HF) and DFT-level, as well as using RI-MP2 and DH-DFT.

2 Theory

The isotropic HFCC can be calculated in the absence of spin-orbit coupling by:[51]

$$a_{\mathrm{iso}}(N) = \frac{\mu_0}{3} g_{\mathrm{e}} g_{\mathrm{N}} \beta_{\mathrm{e}} \beta_{\mathrm{N}} \langle S_z \rangle^{-1} \rho(N), \tag{1}$$

where μ_0 is the permeability of the vacuum, g_e and g_N are the electronic and nuclear g factor, β_{e} is the Bohr magneton, β_{N} is the nuclear magneton of the nucleus N, and $\langle S_z \rangle$ is the mean value of S_z in the current electronic state. $\rho(N)$ is the Fermi-contact integral, which is given by:

$$\rho(N) = \sum_{\mu\nu} P_{\mu\nu}^{\alpha-\beta} \langle \phi_\mu(r) | \delta(r - r_{\mathrm{N}}) | \phi_\nu(r) \rangle, \tag{2}$$

where $P_{\mu\nu}^{\alpha-\beta}$ is the difference between the α- and β-electron density matrices.

For methods beyond Hartree-Fock or DFT, the respective energy equation needs to be perturbed with respect to the nuclear magnetic moment M_k of nucleus k.[51] For an efficient low-scaling algorithm to obtain the analytic gradient of the MP2 expression, which analogously arises in DH-DFT, Schweizer *et al.*[41] proposed a reformulation in AOs by the Laplace-transform ansatz.[52,53] Thus, the locality of AOs can be exploited, ultimately leading to a linear scaling behavior. A further reduction of the computational cost can be achieved by the RI approximation[32–37] and a Cholesky decomposition[42,43] following the RI-CDD ansatz previously established for the computation of MP2 energies.[54] In combination with an extention of the QQR-type integral estimation technique to efficiently select significant two-electron contributions,[39,40] we recently introduced a low scaling efficient ansatz to compute first-order properties,[44] which can be improved upon by computing selected nuclei only, exploiting the locality of the perturbation.[45]

3   Computational Details

The isotropic HFCCs were obtained at the HF-, DFT-, DH-DFT- and MP2-level using the respective implementation in the program package FermiONs++.[55,56] The MP2-contributions are hereby computed using the aforementioned Laplace-transform RI-CDD ansatz[44] and a QQR-based integral screening.[39,40] The Laplace expansion coefficients are selected based on the minimax-approximation.[57] The extents of the QQR-type integral estimates are determined with the same thresholds as in Ref. 40. The QQR-screening threshold was set to 10^{-8} and seven Laplace expansion points were chosen based on the study of the accuracy in Ref. 44. The DL-UCPSCF was converged to a threshold of 10^{-4}. Deviations of less than 1 MHz can be expected with these thresholds.[44] The auxiliary basis set def2-TZVPP-RI/JK by Weigend[58] was chosen for the computations using the basis set EPR-III,[9] which was previously found suitable.[59] Reference coupled cluster computations with singles and doubles

excitations[60] were obtained with the program package Cfour.[61]

The *ab inito* molecular dynamics simulations at the PBEH3c-level[50] were performed as canonical (NVT-)ensembles with the Velocity Verlet propagator[62,63] at 298.15 K, using the Bussi-Donadio-Parrinello thermostat.[64] Each simulation included a 100 fs equilibration period and a 10 ps production run with a time step of 0.1 fs. Geometries were saved every 1 fs. Furthermore, the fully converged extended Lagrangian Born-Oppenheimer MD (XL-BOMD) method[65] was used to speed up SCF convergence. HFCCs were computed for every 100th geometry, i.e., every 100 fs, of the trajectory. To analyze whether a less computationally demanding AIMD simulation is sufficient, some simulations were performed analogously at the HF3c-level.[49]

4 Results

4.1 In-depth study of the ethanal and the dimethylamine radical

In this section, we analyze in detail the dependence of the HFCCs on bond lengths, bond angles, and dihedral angles of both ethanal and the dimethylamine radicals, supplementing previous work on other systems in Refs. 22,24,25,27,66.

We first analyze the distribution of the bond angles, bond lengths, and the dihedral angle in the ethanal radical as obtained from both a HF3c- and a PBEH3c-based AIMD. The results are shown in Figure 1. Anharmonicity can clearly be seen, especially in the distribution of the $\angle$(C-C=O) bond angle in the case of the PBEH3c-based simulation and in the C-H bond length in the HF3c simulation.

Anharmonicity is also present in the distribution of the δ(N-C) bond length and the $\angle$(C-N-C) angle in the dimethylamine radical in Figure 2. The distribution of the $\angle$(C-N-C) angle in the PBEH3c-based simulation is significantly broader than in the HF3c-based AIMD, which is reflected in the standard deviation of 4.3° vs. 2.7°.

When comparing the results obtained from a HF3c-based AIMD to the respective PBEH3c-
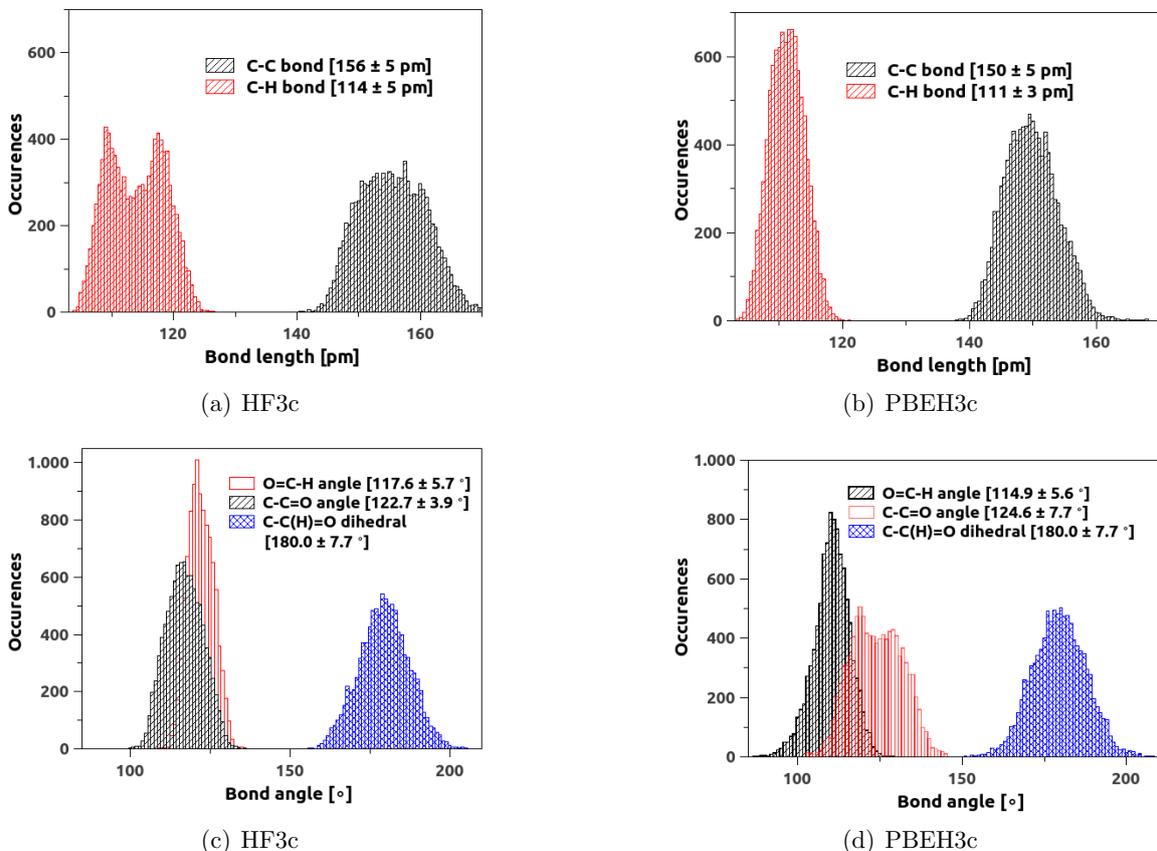
Figure 1: Anharmonicity in the distribution of bond lengths and bond and dihedral angles in the *ab initio* MD simulation of the ethanal radical at room temperature at the HF3c- and PBEH3c-level. Average bond lengths, angles, and dihedral angles and their standard deviations are shown in the legend.

based results, differences in the distribution are apparent (see, e.g., the distribution of ∠(C-C=O) in Figure 1). Thus, specific care has to be taken as to what level of theory is employed to describe the dynamic contributions, i.e., an accurate description of the potential energy surface of the system is paramount. This is shown in Table 1, where the isotropic HFCCs of both systems obtained from the HF3c- and PBEH3c-optimized structures as well as from averaging over the respective AIMD simulations are compared. As can be seen, considerable deviations larger than 20 MHz occur in the case of the hydrogen atom between the results obtained with structures obtained at the HF3c-level in comparison to HFCCs of PBEH3c-optimized structures. The deviation is larger than the influence of the dynamics,
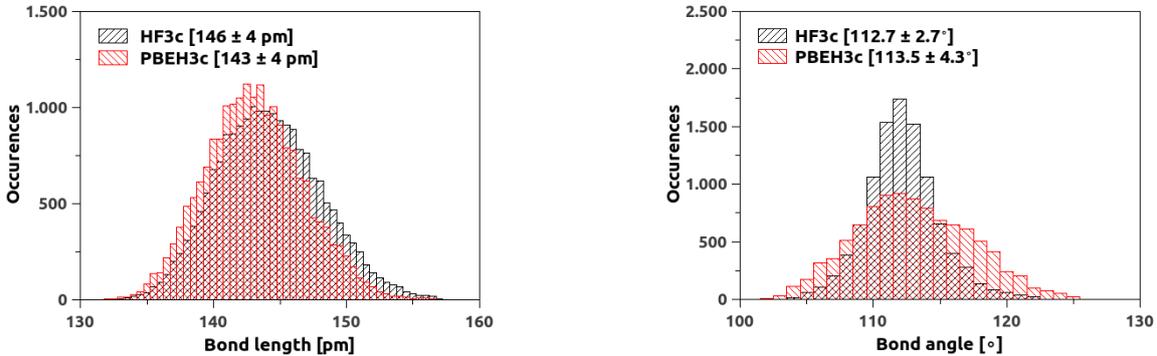
Figure 2: Anharmonicity in the distribution of the N-C bond length and and the ∠(C-N-C) bond angle in the *ab initio* MD simulation of the dimethylamine radical at room temperature at the HF3c- and PBEH3c-level. The average bond length and bond angle and the standard deviation is indicated.

as shown in Table 1, where the HFCCs obtained from the respective AIMD simulation is shown. It has to be noted, though, that (i) the deviation due to using the HF3c instead of PBEH3c is less severe than the effect of differing methods used for the HFCC computation shown in Table 2, and that (ii) the description of dynamic contributions between HF3c and PBEH3c is comparable, as can be seen from the Δ_{dyn} values in Table 1. The latter motivates the computation of the dynamic contributions based on a HF3c-based AIMD trajectory.

Though dynamic contributions results in changes in the HFCCs, these are on the same order as the differences between the results obtained from DFT, post-KS, and post-HF methods in Table 2. While this is not always the case, as shown in Sec. 4.2 where the neglect of dynamic contributions leads to results strongly deviating from experimental findings for a variety of radicals, the results in Table 2 motivate a careful choice of the method with which the isotropic HFCCs are computed. It furthermore has to be noted that the results obtained from the AIMD simulation are not always closer to the experimental results. A variety of reasons can be named for this: first and foremost the error of the approximation to the Schrödinger equation itself and, second, the exact experimental conditions were not sufficiently replicated, and including solution effects might be crucial.

The reason for the significant dynamic contributions, apart from the anharmonicity in

8

Table 1: Isotropic HFCCs in MHz of the dimethylamine and the ethanal radical using the HF3c-optimized and the PBEH3c-optimized structures at the B3LYP/def2-TZVPP level as well as based on the respective AIMD-simulations. Dynamic contributions Δ_{dyn}, i.e., the difference of the HFCCs obtained with the optimized structures to their respective AIMD simulations, are also shown.

Nucleus	Opt. structure			AIMD simulation			Δ_{dyn}	
	HF3c	PBEH3c	$\|\Delta_{\text{PBE-HF}}\|$	HF3c	PBEH3c	$\|\Delta_{\text{PBE-HF}}\|$	HF3c	PBEH3c
Ethanal radical								
^{19}O	-26.8	-28.6	1.8	-26.5	-28.6	2.1	0.3	0.0
^{13}C'	40.1	42.8	2.7	40.2	43.0	2.8	0.1	0.2
^{13}C"	-85.2	-75.2	10.0	-88.6	-77.3	11.3	3.4	2.1
^{1}H@C'	338.3	311.0	27.3	352.1	330.5	21.6	13.8	19.5
Dimethylamine radical								
^{14}N	30.6	29.6	1.0	31.0	30.6	0.4	0.6	1.0
^{13}C	-30.7	-32.8	2.1	-31.5	-33.6	2.1	-0.8	-0.8
^{1}H	67.2	73.2	6.0	69.5	76.4	6.9	2.3	3.2

Table 2: Isotropic HFCCs of the ethanal and the dimethylamine radicals obtained from the PBEH3c-optimized structure (1) and based on the respective AIMD simulation (2) using different levels of theory with the basis set EPR-III/def2-TZVPP-RI.

Nucleus	B3LYP [MHz]		RI-B2PLYP [MHz]		RI-CDD MP2 [MHz]		exptl.
	1	2	1	2	1	2	
Ethanal radical							
^{19}O	-43.6	-43.1	-57.2	-60.8	-48.8	-45.5	—
^{13}C'	42.9	43.2	48.4	47.6	51.3	52.7	—
^{13}C"	-73.4	-75.4	-84.6	-84.1	-89.3	-111.1	—
^{1}H@C'	335.5	356.3	329.3	355.0	301.0	334.2	381[66]
Dimethylamine radical							
^{14}N	34.7	35.3	40.8	41.6	30.8	37.6	41.4[67]
^{13}C	-32.0	-32.7	-36.4	-37.3	-35.3	-34.3	—
^{1}H	80.0	82.6	77.8	80.3	63.9	71.8	76.7[67]

the distribution of the structural parameters in the AIMD simulation, is the strong and non-linear dependency of the HFCCs on the bond lengths, bond angles, and dihedral angles. In order to investigate this relationship systematically, we start from a PBEH3c-optimized structure of both the ethanal and the dimethylamine radical. Subsequently, we modify one

structural parameter, i.e., the bond length, bond angle, or dihedral angle, at a time, and compute the isotropic HFCCs of all nuclei for each structure using the hybrid functional B3LYP, the double-hybrid DFT method B2PLYP, HF, RI-MP2, and CCSD with the basis set def2-TZVPP.

The results for the ethanal radical are shown in Figure 3. Noteworthy is the strong dependence of the isotropic HFCCs on the geometry, especially for the hydrogen nuclei. While for most nuclei the dependency of the HFCC on the structural parameters is consistent throughout the different methods studied here, the importance of correlation can be deduced from the differing results obtained with HF. This is most apparent in the dependency of the ^{13}C'-HFCC on the angle $\angle$(C-C=O) and the angle $\angle$(H-C=O), as well as on both bond lengths. In general, the HF results deviate substantially from the results obtained with the other methods. Thus, irrespective of dynamic contributions, HF computations are incapable of correctly describing the spin density in ethanal and methods incorporating electron correlation are required. Furthermore, strong non-linear dependencies indicate that dynamic contributions will change the *in silico* results considerably.

Similar results are obtained with the dimethylamine radical in Figure 4, except for in the description of the hydrogen atoms, where HF lies in between the other methods taking electron correlation into account. Both the carbon and the nitrogen nucleus, however, strongly depend on the electron correlation.

4.2 Study of a set of organic radicals

In order to determine both the influence of electron correlation and dynamic contributions on the accuracy of HFCCs, we computed the HFCCs based on AIMD simulations at the PBEH3c-level of theory for a set of organic radicals shown in Figure 5.
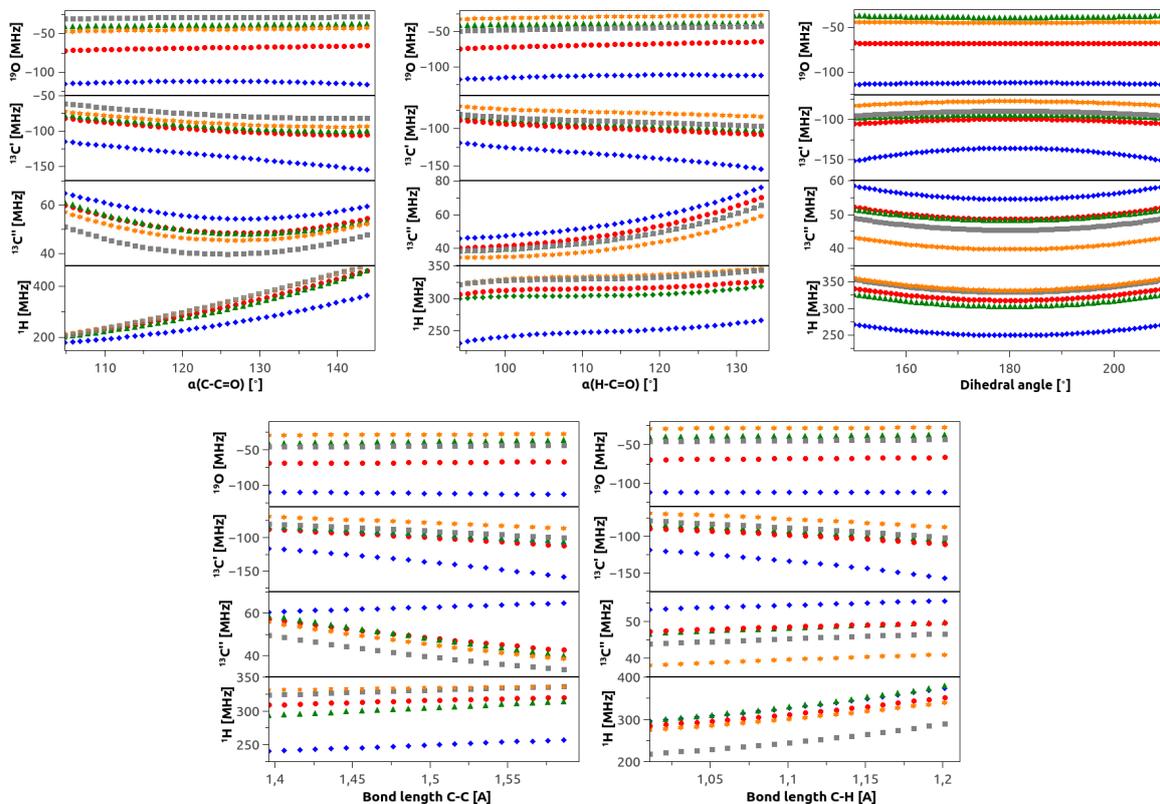
Figure 3: Dependency of the isotropic HFCCs of the different nuclei in the ethanal radical on the bond angles $\angle(C-C=O)$ and $\angle(O=C-H)$, the bond lengths δ(C-C) and δ(C-H) and the dihedral angle. The PBEH3c-based geometry optimization leads to $\angle(C-C=O) = 124.8°$, $\angle(O=C-H) = 114.2°$, $\delta(C-C) = 1.496$ Å, $\delta(C-H) = 1.111$ Å and $\angle$(dihedral) $= 180.0°$. All computations were performed with the def2-TZVPP basis set. HF is represented by blue diamonds, RI-MP2 by green triangles, CCSD by red circles, B3LYP by grey rectangulars, and RI-B2PLYP by orange stars.

4.2.1 Convergence with the number of frames

Prior to comparing the results of all radicals, we analyzed in detail the convergence with respect to the number of MD frames for which the HFCCs were computed. This is shown in Figure 6 for a selection of three radicals (**2**, **3**, and **11**). The change of the average HFCCs with the number of frames is considerably smaller than the standard deviation. Nonetheless, converged results require more than 50 frames (5 ps). Noteworthy is radical **3**, where a significant standard deviation is apparent, especially in the HFCCs of the hydrogen atoms.

11
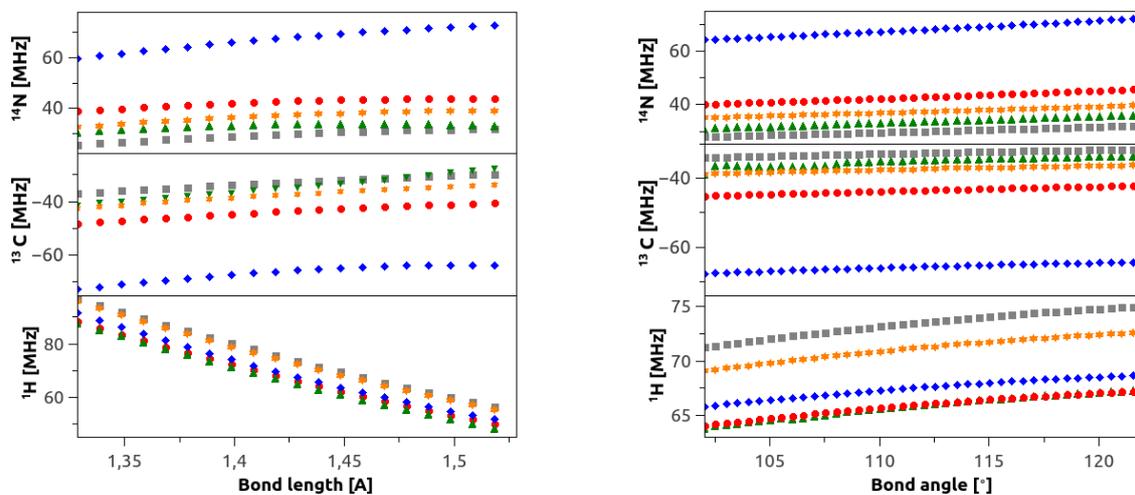
Figure 4: Dependency of the isotropic HFCCs of the different nuclei (carbon, nitrogen, and hydrogen) in dimethylamine on the N-C bond length and the ∠(C-N-C) bond angle. The PBEH3c-based geometry optimization leads to a N-C bond length of 1.428 Å and a ∠(C-N-C) bond angle of 112.04°. All computations were performed with the def2-TZVPP basis set. HF is represented by blue diamonds, RI-MP2 by green triangles, CCSD by red circles, B3LYP by grey rectangulars, and RI-B2PLYP by orange stars.
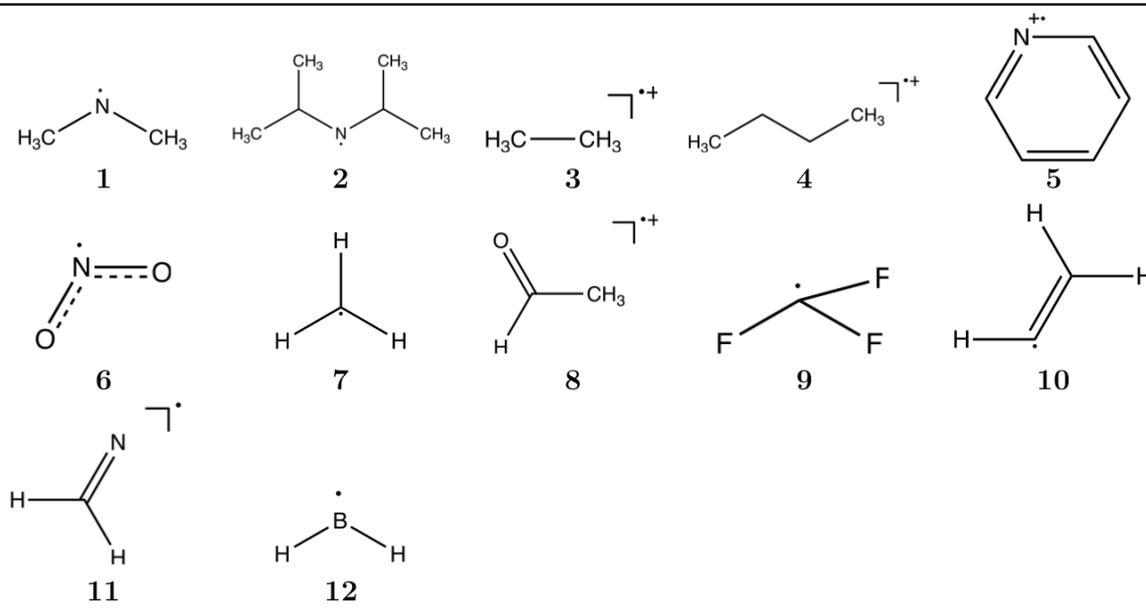
Figure 5: Organic radicals investigated in this work.

This can be explained by a Jahn-Teller distortion of the CH-bonds, which will be discussed in detail in Sec. 4.2.2.

(a) Radical **2**   (b) Radical **3**   (c) Radical **11**
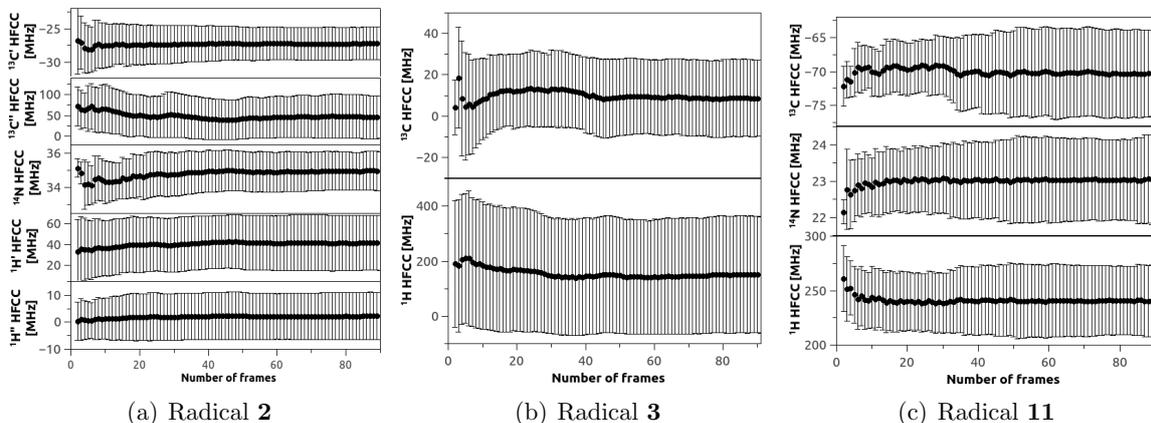
Figure 6: Convergence of the isotropic HFCCs at the B3LYP/EPR-III level with the number of frames taken from the PBEH3c-AIMD simulation. The frames are sorted chronologically.

4.2.2 Molecular dynamic and electron correlation contributions

The isotropic HFCCs of all organic radicals shown in Figure 5 were obtained with the PBEH3c-optimized structure and from the respective AIMD simulation with B3LYP, RI-B2PLYP, and RI-CDD MP2, and can be found in Table 3. We also state experimental results where available. While a comparison to the experimental results is instructive, caution is warranted as the AIMD simulations do not necessarily replicate the experimental conditions. The experimental conditions include a variety of temperatures and solvents/matrices, whereas all AIMD simulations were performed at room temperature in the gas phase.

Both dynamic contributions and electron correlation constitute a significant contribution to the overall obtained HFCCs. Taking into account electron correlation is often important and significantly changes the computed HFCCs. This is most apparent by comparing the Hartree-Fock and RI-MP2 results with deviations of 26.9 MHz on average. In most cases, including electron correlation significantly improves the agreement between the *in silico* HFCCs and the experimental results (see ^{14}N-HFCC in radical **1** and **2**, ^{19}F-HFCC in

13

radical **9**, ^{11}B-HFCC in radical **12**, and the ^{1}H-HFCCs in radicals **3**, **8**, and **12**). This is consistent with the results in Sec. 4.1. In this analysis, we disregarded the radicals with extensive spin contamination leading to erratic results, i.e., radicals **5**, **10**, and **11**. These systems are better described at the double hybrid DFT-level, where similarly the correlation contributions prove to be considerable, changing the computed HFCCs on average by 9.4 MHz. In most cases taking electron correlation into account via double-hybrid DFT improves the obtained HFCCs with respect to the experimental results (see ^{14}N-HFCC in radicals **1**, **2**, **5**, **6**, and **11**, ^{19}F-HFCC in radical **9**, ^{13}C-HFCC in radicals **7**, **10**, and **11**, and ^{1}H-HFCCs in radicals **7**, **8**, **10**, **11**, and **12**). The larger magnitude of the correlation effect on RI-MP2 than on RI-B2PLYP can be explained by the scaling factor of 0.27 present in the B2PLYP-functional and by the fact that electron correlation is also partially included in the hybrid-DFT ansatz.

The comparison to the hybrid DFT results obtained with B3LYP are somewhat more complicated, as B2PLYP and B3LYP vary significantly in their HF-exchange contribution. While the choice of optimal double hybrid functionals is challenging, recent work showcases that a correct determination of the optimal HF exchange contribution as well as spin component scaling in the second-order correlation contribution can significantly improve the computational results.[68] In this work, we chose the two well-established functionals B3LYP and B2PLYP. While the comparison can not directly be related to the additional treatment of electron correlation, B2PLYP significantly improves the *in silico* results in most cases with respect to the B3LYP results (see, e.g., for the ^{14}N-HFCC in radicals **1**, **2**, **5**, and **11**, the ^{19}F-HFCC in radical **9**, the ^{13}C-HFCCs in radicals **10** and **11**, and the ^{1}H-HFCCs in radicals **1**, **2**, **3**, **4**, **7**, and **10**).

Overall, molecular dynamic contributions change the computed HFCCs by on average 10.7 MHz for B3LYP, 11.3 MHz for B2PLYP, and by 31.9 MHz for RI-MP2. For the latter method we disregarded the systems with high spin contamination. While the dynamic contributions are on the same order as the correlation contribution, a detailed analysis discloses

that for a variety of systems the inclusion of dynamical effects is indispensable for the accurate description of the system. The results of these systems are shown in Table 3 and will be discussed in detail in the following. At this point we want to mention that the signs indicated in the experimental results are not directly obtained from the experiment but are assigned afterwards using theoretical results. As the absolute values match in the cases where the sign obtained with B3LYP, RI-B2PLYP, and RI-MP2 deviates from the corresponding values in the experimental column and since the three different methods lead to similar results, one can assume the assignment of the sign to the experimental value to be erroneous.

The hyperfine splitting of alkane radical cations, namely of the ethane radical **3** and the butane radical **4**, represent special cases, where a Jahn-Teller distortion breaks the symmetry of the six CH-bonds at the terminal methyl groups.[69–71] This can be verified experimentally by looking at the low temperature EPR spectrum of the ethane radical, which exhibits a triplett splitting due to a localization of the spin density at two equivalent hydrogen atoms.[69,70] Moving to higher temperatures, dynamic contributions lead to a septett splitting, as on average all six hydrogen atoms have become equivalent.[69,70] This is not due to a disappearance of the Jahn-Teller distortion, but due to this effect becoming dynamic. Our results confirm this behavior, as two C-H bonds are shorter (1.07 ± 0.03 Å) than the remaining four (1.13 ± 0.03 Å) in the AIMD simulation. We thus expect a significant change in the obtained HFCCs when dynamics are considered via an AIMD simulation. We can confirm this behavior in our computations of radical **3**, where we obtain HFCCs close to the experimental results. When moving towards higher temperatures, our computation confirm the septett splitting.

Table 3: Influence of dynamic contributions and electron correlation on the HFCCs in MHz of radicals **1-12** in Figure 5. HFCCs are obtained (1) with the PBEH3c-optimized structure and (2) from averaging 100 frames from PBEH3c-AIMD simulations at 298.15 K. Computations employ the basis set EPR-III. Where available, experimental results are shown arising from varying experimental conditions. Values in parentheses correspond to the contribution of the electron correlation as described by second-order perturbation theory. RI-MP2 that are erratic due to spin contamination are omitted and the respective (average) $< \hat{S}^2 >$ is shown.

Nucleus	B3LYP		RI-B2PLYP		RI-CDD MP2		exptl.
	1	2	1	2	1	2	
Radical **1**							
^{14}N	34.7	35.3	40.8 (-9.4)	41.6 (-9.6)	30.8 (-41.1)	37.6 (-36.3)	41.4[67]
^{13}C	-32.0	-32.7	-36.4 (6.4)	-37.3 (6.6)	-34.1 (30.8)	-34.3 (32.4)	—
^{1}H	119.9	} 82.6	116.7 (3.7)	} 80.3 (1.9)	100.8 (-5.1)	} 71.8 (-3.4)	} 76.7[67]
^{1}H	0.1		0.0 (-1.8)		-8.0 (-13.8)		

The non-dynamic computations exhibit four and two equivalent hydrogen atoms, leading to two separate hydrogen HFCCs. Averaging over these HFCCs results in 80.0 MHz (B3LYP), 77.8 MHz (B2PLYP), and 63.9 MHz (RI-CDD MP2) for the hydrogen atom.

Nucleus	B3LYP		RI-B2PLYP		RI-CDD MP2		exptl.
Radical **2**							
^{14}N	34.5	34.9	40.8 (-9.1)	41.2 (-9.4)	63.7 (-8.5)	34.3 (-39.0)	40.1[67]
^{13}C'	-26.8	-27.1	-30.0 (7.4)	-30.3 (7.8)	-16.7 (42.1)	-25.8 (35.0)	—
^{13}C"	70.0	} 41.1	69.9 (1.5)	} 40.7 (0.3)	91.9 (24.5)	} 37.3 (-4.2)	—
	14.0		13.9 (-0.6)		50.0 (-9.8)		—
^{1}H'	34.2	46.9	33.1 (0.0)	46.0 (1.2)	0.3 (-33.7)	37.7 (-5.6)	40.1[67]
^{1}H"	22.2	} 2.2	21.1 (2.2)	} 1.8 (0.9)	8.5 (-5.4)	} 2.2 (3.8)	} 1.8[67]
	-2.8		-3.0 (0.5)		-18.3 (-13.1)		

In the non-dynamic computations, the C" and H" exhibit separate HFCCs. Averaging over these HFCCs results in 42.0 MHz (B3LYP), 41.9 MHz (B2PLYP), and 71.0 MHz (RI-CDD MP2) for the carbon atoms, and 2.7 MHz (B3LYP), 2.2 MHz (B2PLYP), and -9.8 MHz (RI-CDD MP2) for the hydrogen atoms.

Table 3 – *Continued from previous page*

Nucleus	B3LYP		RI-B2PLYP		RI-CDD MP2		exptl.
	1	2	1	2	1	2	
Radical **3**							
^{13}C	20.5	8.6	17.7 (-2.6)	4.7 (-0.2)	6.3 (-14.1)	-5.7 (-7.1)	—
^{1}H'	471.0	} 149.5	461.6 (4.8)	} 145.7 (1.0)	445.4 (23.2)	} 136.8 (12.7)	427.4/
^{1}H'	-19.7		-23.3 (5.0)		-26.2 (17.6)		141.0 [69,70]

The first experimental value is obtained from a 4K measurement, where a triplett splitting is observed. At higher temperatures, the dynamics result in a septett splitting, i.e., an averaged HFCC over six equivalent hydrogen atoms, corresponding to the second experimental value from a 77K measurement. [69,70]

Radical **4**							
^{13}C'	-9.9	-10.9	-12.1 (0.5)	-13.0 (9.4)	-10.2 (8.2)	-13.8 (6.7)	—
^{13}C"	8.9	13.1	7.3 (2.5)	11.6 (1.5)	6.4 (0.1)	0.0 (-12.3)	—
^{1}H'	9.4	13.6	7.7 (3.1)	11.8 (4.0)	10.3 (20.1)	8.5 (15.6)	–
^{1}H"	214.2	154.3	197.5 (14.0)	142.3 (9.5)	179.7 (34.5)	125.9 (19.6)	171.8 [69,70]
	20.9	39.8	18.6 (1.2)	36.1 (3.0)	14.0 (0.4)	30.5 (6.9)	22.4 [69,70]

Similar to radical 10, a Jahn-Teller distortion can be observed at low temperatures. From the PBEH3c-optimized structure, two distinct hydrogen HFCCs of the methyl groups can be observed.

Radical **5**							
^{14}N	110.3	110.6	117.2 (-20.6)	116.4 (-20.9)			114.9 [72]
o-^{13}C	35.1	33.8	38.5 (-11.5)	37.5 (-14.2)			—
m-^{13}C	-14.0	-11.9	-15.9 (10.2)	-14.5 (12.1)			—
p-^{13}C	-10.5	-12.1	-10.7 (8.3)	-12.1 (9.2)	$< \hat{S}^2 > =$ 1.1997	$< \hat{S}^2 > =$ 1.2345	—
o-^{1}H	93.9	91.4	89.9 (4.6)	88.0 (3.5)			82.1 [72]
m-^{1}H	35.5	31.2	32.2 (9.5)	28.4 (10.9)			31.3 [72]
p-^{1}H	22.5	23.7	16.4 (1.8)	16.7 (-2.1)			24.1 [72]
Radical **6**							
^{14}N	143.5	142.5	146.8 (0.5)	146.4 (2.5)	145.9 (-7.7)	134.1 (-15.8)	152.0 [73]
^{17}O	-58.2	-57.0	-68.1 (-7.0)	-66.9 (-7.7)	-70.9 (3.5)	-69.5 (-3.5)	62.2 [73]
Radical **7**							
^{13}C	79.8	114.6	85.1 (-26.0)	120.0 (-26.3)	59.3 (-101.5)	94.8 (-102.2)	107.4 [74]
^{1}H	-64.4	-58.4	-70.1 (13.0)	-64.0 (12.5)	-71.9 (49.8)	-65.9 (48.6)	64.6 [74]

Table 3 – *Continued from previous page*

Nucleus	B3LYP		RI-B2PLYP		RI-CDD MP2		exptl.
	1	2	1	2	1	2	
Radical **8**							
^{17}O	-43.6	-43.1	-57.2 (20.5)	-60.8 (16.7)	-48.8 (79.4)	-45.5 (83.2)	—
^{13}C'	42.9	43.2	48.4 (-1.8)	47.6 (-2.3)	51.3 (-5.5)	52.7 (-4.2)	—
^{13}C''	-73.4	-75.4	-84.6 (13.3)	-84.1 (18.1)	-89.3 (41.1)	-111.1 (26.9)	—
^{1}H'	335.5	356.3	329.3 (33.4)	355.0 (37.3)	301.0 (51.9)	334.2 (65.7)	381[70]
^{1}H''	-3.0	-0.6	-5.3 (3.4)	0.5 (7.9)	-8.9 (-5.7)	-23.8 (-9.5)	-8[70]
Radical **9**							
^{13}C	728.0	737.1	745.6 (-22.5)	754.9 (21.8)	750.1 (-67.6)	762.2 (-64.9)	—
^{19}F	393.3	390.4	408.7 (-3.4)	406.5 (-4.5)	370.1 (-47.9)	398.0 (-16.9)	405.0[75]
Radical **10**							
^{13}C'	314.5	296.5	321.6 (-38.4)	304.1 (-39.8)			301.5[76,77]
^{13}C''	-11.9	-17.0	-17.3 (16.6)	-22.8 (18.0)	$<\hat{S}^2> =$ 0.9399	$<\hat{S}^2> =$ 0.9459	-24.1[76,77]
^{1}H'	53.2	44.6	46.0 (12.8)	37.4 (13.5)			38.7[76,77]
^{1}H''	178.7	182.5	177.9 (2.8)	181.9 (2.2)			184.8[76,77]
^{1}H'''	113.8	120.7	112.7 (-0.4)	119.6 (0.1)			111.0[76,77]
Radical **11**							
^{13}C	-68.0	-70.2	-77.7 (19.3)	-80.0 (21.8)			81.1[78]
^{14}N	23.1	23.1	27.2 (-9.1)	27.2 (-9.4)	$<\hat{S}^2> =$ 0.9356	$<\hat{S}^2> =$ 0.9410	28.6[78]
^{1}H	235.7	240.0	235.4 (8.3)	239.7 (7.3)			244.8[78]
Radical **12**							
^{11}B	363.7	342.0	366.5 (-14.9)	345.0 (-15.5)	357.7 (-43.1)	336.0 (-45.8)	358[79]
^{1}H	42.5	35.6	39.5 (5.2)	32.4 (5.8)	26.1 (17.9)	19.4 (19.9)	38[79]
Averages $\overline{\Delta}$ (correlation contribution $\overline{\Delta}_\text{corr.}$, dynamic contribution $\overline{\Delta}_\text{dyn}$)							
$\overline{\Delta}_\text{corr.}$	—		**9.4**		**26.9**		—
$\overline{\Delta}_\text{dyn}$	**10.7**		**11.3**		**31.9**		—

In the case of the butane radical cation, a triplett splitting is also observed experimentally at 77 K. We can confirm this splitting in both our PBEH3c-optimized structure and in the results based on our AIMD simulation, where we see two discrete hydrogen-HFCCs with

the larger component arising from two hydrogen atoms leading to the observed triplett splitting. While the inclusion of dynamic contributions leads to a reduction of the HFCCs in the direction of the experimental values, they overshoot, which is in line with our room temperature simulation resulting in a larger reduction.

Similarly, the HFCCs obtained from the PBEH3c-optimized structures of radicals **1** and **2** show that the spin density is primarily located at four of the six hydrogen atoms. An analysis of the CH bond lengths shows in analogy to the ethane radical that two CH-bonds are shorter than the remaining four (1.089 Å vs. 1.098 Å). When molecular dynamic contributions are considered, this distortion is averaged resulting in HFCCs that match the experimental results closely. It has to be noted, though, that similar results can be obtained from averaging the six hydrogen-HFCCs obtained with the PBEH3c-optimized structure. We therefore considered dynamic contributions in our overall analysis only by comparing to the averaged results. The effect is analogous in radical **2**.

Another example of extensive dynamic contributions is the methane radical cation **7**. Optimizing the structure leads to a planar configuration. When vibrational averaging is considered, non-planar configurations also contribute to the overall HFCCs, which leads to considerable changes of up to 40 % and to a better agreement with the experimental findings. It is apparent, that the spin density at the carbon atom increases, whereas the spin density at the hydrogen nuclei decreases considerably and consistently both for B3LYP and B2PLYP.

The results in Table 3 include dynamic contributions using the PBEH3c-AIMD simulations. Especially when moving towards larger molecular systems, the computational cost for such high-level MD simulations will constitute an impediment. While the results in Table 1 indicate that the computationally cheaper HF3c method leads to deviating results, these can mostly be contributed to a differing optimized ground state structure (see Figures 1 and 2). The dynamic contributions, however, are approximated decently using the HF3c-based AIMD simulations. We thus investigated whether adding the Δ_{dyn} contribution at the HF3c-AIMD level to the HFCCs obtained with a PBEH3c-optimized structure can be

considered a viable pragmatic approach. This is shown in the following Table 4, where we compare the results for radicals **1**, **8**, and **9** using this combined HF3c/PBEH3c-approach to the respective HFCCs obtained from a HF3c- and a PBEH3c-AIMD simulation.

Table 4: Comparison of the HFCCs in MHz obtained from the PBEH3c-optimized structure (stat$_{\text{PBEH3c}}$) and from HF3c- (dyn$_{\text{HF3c}}$) and PBEH3-simulations (dyn$_{\text{PBEH3c}}$) to the combined HF3c/PBEH3c-approach (dyn$_{\text{comb}}$). The HF3c/PBEH3c approach consists of adding the dynamical correction as obtained from an HF3c-AIMD simulation to the HFCCs obtained with the PBEH3c-optimized structure.

Nucleus	B3LYP				RI-B2PLYP			
	stat$_{\text{PBEH3c}}$	dyn$_{\text{HF3c}}$	dyn$_{\text{PBEH3c}}$	dyn$_{\text{comb}}$	stat$_{\text{PBEH3c}}$	dyn$_{\text{HF3c}}$	dyn$_{\text{PBEH3c}}$	dyn$_{\text{comb}}$
				Radical **1**				
^{14}N	34.7	35.7	35.3	35.0	40.8	41.9	41.6	41.2
^{13}C	-32.0	-30.6	-32.7	-32.8	-36.4	-34.8	-37.3	-37.3
^{1}H	80.0	75.1	82.6	82.4	77.8	73.1	80.3	80.2
				Radical **8**				
^{17}O	-43.6	-40.1	-43.1	-42.9	-57.2	-56.2	-60.8	-60.2
^{13}C'	42.9	40.9	43.2	43.1	48.4	45.5	47.6	50.8
^{13}C"	-73.4	-86.3	-75.4	-76.6	-84.6	-98.4	-84.1	-83.9
^{1}H'	335.5	379.9	356.3	350.5	329.3	383.8	355.0	333.4
^{1}H"	-3.0	-0.1	-0.6	-2.7	-5.3	-0.3	0.5	-2.2
				Radical **9**				
^{13}C	728.0	535.8	737.1	732.9	745.6	554.3	754.9	751.0
^{19}F	393.3	489.1	390.4	389.6	408.7	503.4	406.5	403.6

The results obtained directly from the HF3c-AIMD simulation deviate significantly from the respective PBEH3c-based results. This confirms the findings in Table 1. However, by adding the dynamic correction as obtained from the HF3c-AIMD simulation to the HFCCs of the PBEH3c-optimized structure, this error can be removed, and results decently incorporating the dynamical effect are achieved. A drastic example where this combined approach leads to improved results is is radical **9**, where the HF3c geometries exhibit significantly smaller δ(C-F) and thus leads to strongly deviating results. The dynamic contribution, however, is correctly incorporated, leading to good results using the combined HF3c/PBEH3c ansatz. The results of radical **9** reinforce the necessity to perform high-level geometry optimizations in order to obtain comparable results.

5 Conclusion

In this work, we studied a variety of organic radicals analyzing both the effect of electron correlation and dynamics simultaneously on the accuracy of the *in silico* HFCCs. In our test set, electron correlation was shown to be a significant contributor, strongly improving the obtained results. The importance of electron correlation can especially be seen when comparing the HF results to the respective RI-MP2 HFCCs. Despite the functional B3LYP leading to results agreeing reasonably well with experimental results, our findings show that further inclusion of electron correlation as within double-hybrid DFT is beneficial and can not be neglected.

While electron correlation has to be considered for accurate results, neglecting molecular dynamic contributions can in some cases lead to incomparable results. In our test set, this especially applies to alkane radicals where the Jahn-Teller distortion turns dynamic. Therefore, we conclude that for reliable in silico HFCCs both effects must be considered. When moving towards larger molecular systems using our established methodology, the cost to compute the HFCCs taking electron correlation into account can be reduced with our recently introduced efficient AO-based approach, whereas our results indicate that obtaining the dynamic correction at a computationally cheaper AIMD-level using our HF3c/PBEH3c approach can be sufficient in many cases to capture most of the dynamic contribution, providing a good compromise between accuracy and computational cost.

References

(1) Jeschke, G. EPR techniques for studying radical enzymes. *Biochim. Biophys. Acta* **2005**, *1707*, 91–102.

(2) Sahu, I. D.; McCarrick, R. M.; Lorigan, G. A. Use of Electron Paramagnetic Resonance to Solve Biochemical Problems. *Biochemistry* **2013**, *52*, 5967–5984.

(3) Rinkevicius, Z.; Telyatnyk, L.; Vahtras, O.; Ågren, H. Density functional theory for hyperfine coupling constants with the restricted-unrestricted approach. *J. Chem. Phys.* **2004**, *121*, 7614.

(4) Rega, N.; Cossi, M.; Barone, V. Development and validation of reliable quantum mechanical approaches for the study of free radicals in solution. *J. Chem. Phys.* **1996**, *105*, 11060.

(5) Munzarová, M.; Kaupp, M. A Critical Validation of Density Functional and Coupled-Cluster Approaches for the Calculation of EPR Hyperfine Coupling constants in Transition Metal Complexes. *J. Phys. Chem. A* **1999**, *103*, 9966–9983.

(6) Kossmann, S.; Kirchner, B.; Neese, F. Performance of modern density functional theory for the prediction of hyperfine structure: meta-GGA and double hybrid functionals. *Mol. Phys.* **2007**, *105*, 2049–2071.

(7) Barone, V.; Cimino, P. Accurate and feasible computations of structural and magnetic properties of large free radicals: The PBE0/N07D model. *Chem. Phys. Lett.* **2008**, *454*, 139–143.

(8) Kirste, B. DFT calculations of hyperfine coupling constants of organic π radicals and comparison with empirical equations and experiment. *Magn. Reson. Chem.* **2016**, *54*, 835–841.

(9) Barone, V. *Recent Advances in Density Functional Methods*; World Scientific Singapore, 1995; p 287.

(10) Grimme, S. Semiempirical hybrid density functional with perturbative second-order correlation. *J. Chem. Phys.* **2006**, *124*, 034108.

(11) Čížek, J. On the Correlation Problem in Atomic and Molecular Systems. Calculation of Wavefunction Components in Ursell-Type Expansion Using Quantum-Field Theoretical Methods. *J. Chem. Phys.* **1966**, *45*, 4256–4266.

(12) Čížek, J.; Paldus, J.; Šroubková, L. Cluster expansion analysis for delocalized systems. *Int. J. Quantum Chem.* **2007**, *3*, 149–167.

(13) Čížek, J.; Paldus, J. Correlation Problems in Atomic and Molecular Systems III. Rederivation of the Coupled-Pair Many-Electron Theory Using the Traditional Quantum Chemical Methods. *Int. J. Quantum Chem.* **2007**, *5*, 359–379.

(14) Verma, P.; Perera, A.; Morales, J. A. Massively parallel implementations of coupled-cluster methods for electron spin resonance spectra. I. Isotropic hyperfine coupling tensors in large radicals. *J. Chem. Phys.* **2013**, *139*, 174103.

(15) Suter, H. U.; Huang, M.-B.; Engels, B. A multireference configuration interaction study of the hyperfine structure of the molecules CCO, CNN, and NCN in their triplet ground states. *J. Chem. Phys.* **1994**, *101*, 7686.

(16) Neese, F. Sum-over-states based multireference *ab initio* calculation of EPR spin Hamiltonian parameters for transition metal complexes. A case study. *Magn. Reson. Chem.* **2004**, *42*, S187–S198.

(17) Shiozaki, T.; Yanai, T. Hyperfine Coupling Constants from Internally Contracted Multireference Perturbation Theory. *J. Chem. Theory Comput.* **2016**, *12*, 4347–4351.

(18) Massolle, A.; Dresselhaus, T.; Eusterwiemann, S.; Doerenkamp, C.; Eckert, H.; Studer, A.; Neugebauer, J. Towards reliable references for electron paramagnetic resonance parameters based on quantum chemistry: the case of verdazyl radicals. *Phys. Chem. Chem. Phys.* **2018**, *20*, 7661–7675.

(19) Grimme, S. A General Quantum Mechanically Derived Force Field (QMDFF) for Molecules and Condensed Phase Simulations. *J. Chem. Theory Comput.* **2014**, *10*, 4497–4514.

(20) Barone, V.; Cimino, P.; Pedone, A. An integrated computational protocol for the accurate prediction of EPR and PNMR parameters of aminoxyl radicals in solution. *Magn. Reson. Chem.* **2010**, *48*, S11–S22.

(21) Stendardo, E.; Pedone, A.; Cimino, P.; Menziani, M. C.; Crescenzi, O.; Barone, V. Extension of the AMBER force-field for the study of large nitroxides in condensed phases: an *ab initio* parametrization. *Phys. Chem. Chem. Phys.* **2010**, *12*, 11697–11709.

(22) Asher, J. R.; Kaupp, M. Hyperfine Coupling Tensors of the Benzosemiquinone Radical Anion from Car-Parrinello Molecular Dynamics. *Chem. Phys. Chem.* **2007**, *8*, 69–79.

(23) Rinkevicius, Z.; Murugan, N. A.; Kongsted, J.; Frecus, B.; Steindal, A. H.; Agren, H. Density Functional Restricted-Unrestricted/Molecular Mechanics Theory for Hyperfine Coupling Constants of Molecules in Solution. *J. Chem. Theory Comput.* **2011**, *7*, 3261–3271.

(24) Tachikawa, H.; Igarashi, M.; Ishibashi, T. Ab initio molecular dynamics (MD) calculations of hyperfine coupling constants of methyl radical. *Chem. Phys. Lett.* **2002**, *352*, 113–119.

(25) Igarashi, M.; Ishibashi, T.; Tachikawa, H. A direct ab initio molecular dynamics study of

the finite temperature effects on the hyperfine coupling constant of methyl radicalwater complexes. *J. Mol. Struct.* **2002**, *594*, 61–69.

(26) Neugebauer, J.; Louwerse, M. J.; Belanzoni, P.; Wesolowski, T. A.; Baerends, E. J. Modeling solvent effects on electron-spin-resonance hyperfine couplings by frozen-density embedding. *J. Chem. Phys.* **2005**, *123*, 114101.

(27) Pavone, M.; Benzi, C.; De Angelis, F.; Barone, V. Hyperfine coupling constants of dimethyl nitroxide in aqueous solution: CarParrinello molecular dynamics and discrete-continuum approaches. *Chem. Phys. Lett.* **2004**, *395*, 120–126.

(28) Chen, X.; Rinkevicius, Z.; Cao, Z.; Ruud, K.; Ågren, H. Zero-point vibrational corrections to isotropic hyperfine coupling constants in polyatomic molecules. *Phys. Chem. Chem. Phys.* **2011**, *13*, 696–707.

(29) Chen, X.; Rinkevicius, Z.; ; Ruud, K.; Ågren, H. Role of zero-point vibrational corrections to carbon hyperfine coupling constants in organic π radicals. *J. Chem. Phys.* **2013**, *138*, 054310.

(30) Adam, A. Y.; Yachmenev, A.; Yurchenko, S. N.; Jensen, P. Ro-vibrational averaging of the isotropic hyperfine coupling constant for the methyl radical. *J. Chem. Phys.* **2015**, *143*, 244306.

(31) Grimme, S.; Bannwarth, C.; Dohm, S.; Hansen, A.; Pisarek, J.; Pracht, P.; Seiber, J.; Neese, F. Fully Automated Quantum-Chemistry-Based Computation of Spin-Spin-Coupled Nuclear Magnetic Resonance Spectra. *Angew. Chem. Int. Ed.* **2017**, *56*, 14763–14769.

(32) Whitten, J. L. Coulombic potential energy integrals and approximations. *J. Chem. Phys.* **1973**, *58*, 4496–4501.

(33) Dunlap, B. I.; Connolly, J. W. D.; Sabin, J. R. On some approximations in applications of Xα theory. *J. Chem. Phys.* **1979**, *71*, 3396–3402.

(34) Feyereisen, M.; Fitzgerald, G.; Komornicki, A. Use of approximate integrals in ab initio theory: An application in MP2 energy calculations. *Chem. Phys. Lett.* **1993**, *208*, 359–363.

(35) Vahtras, O.; Almlöf, J.; Feyereisen, M. Integral approximations for LCAO-SCF calculations. *Chem. Phys. Lett.* **1993**, *213*, 514–518.

(36) Weigend, F.; Häser, M. RI-MP2: first derivatives and global consistency. *Theor. Chem. Acc.* **1997**, *97*, 331–340.

(37) Weigend, F.; Häser, M.; Patzelt, H.; Ahlrichs, R. RI-MP2: optimized auxiliary basis sets and demonstration of efficiency. *Chem. Phys. Lett.* **1998**, *294*, 143–152.

(38) Doser, B.; Zienau, J.; Clin, L.; Lambrecht, D. S.; Ochsenfeld, C. A Linear-Scaling MP2 Method for Large Molecules by Rigorous Integral-Screening Criteria. *Z. Phys. Chem.* **2010**, *224*, 397.

(39) Maurer, S. A.; Lambrecht, D. S.; Flaig, D.; Ochsenfeld, C. Distance-dependent Schwarz-based integral estimates for two-electron integrals: Reliable tightness vs. rigorous upper bounds. *J. Chem. Phys.* **2012**, *136*, 144107.

(40) Maurer, S. A.; Lambrecht, D. S.; Kussmann, J.; Ochsenfeld, C. Efficient distance-including integral screening in linear-scaling Møller-Plesset perturbation theory. *J. Chem. Phys.* **2013**, *138*, 014101.

(41) Schweizer, S.; Doser, B.; Ochsenfeld, C. An atomic orbital-based reformulation of energy gradients in second-order Møller-Plesset perturbation theory. *J. Chem. Phys.* **2008**, *128*, 154101.

26

(42) Koch, H.; Sánchez de Meras, A.; Pedersen, T. B. Reduced scaling in electronic structure calculations using Cholesky decompositions. *J. Chem. Phys.* **2003**, *118*, 9481–9484.

(43) Boström, J.; Pitoňák, M.; Aquilante, F.; Neogrády, P.; Pedersen, T. B.; Lindh, R. Coupled Cluster and Møller-Plesset Perturbation Theory Calculations of Noncovalent Intermolecular Interactions using Density Fitting with Auxiliary Basis Sets from Cholesky Decompositions. *J. Chem. Theory Comput.* **2012**, *8*, 1921–1928.

(44) Vogler, S.; Ludwig, M.; Maurer, M.; Ochsenfeld, C. Low-scaling first-order properties within second-order Møller-Plesset perturbation theory using Cholesky decomposed density matrices. *J. Chem. Phys.* **2017**, *147*, 024101.

(45) Vogler, S.; Savasci, G.; Ludwig, M.; Ochsenfeld, C. Selected-Nuclei Method for the Computation of Hyperfine Coupling Constants within Second-Order Møller-Plesset Perturbation Theory. *J. Chem. Theory Comput.* **2018**, *14*, 3014–3024.

(46) Warshel, A.; Karplus, M. Semiclassical trajectory approach to photoisomerization. *Chem. Phys. Lett.* **1975**, *32*, 11–17.

(47) Leforestier, C. Classical trajectories using the full ab iniot potential energy surface H^- + $CH_4 \rightarrow CH_4 + H^-$. *J. Chem. Phys.* **1978**, *68*, 4406.

(48) Peters, L. D. M.; Kussmann, J.; Ochsenfeld, C. Efficient and Accurate BornOppenheimer Molecular Dynamics for Large Molecular Systems. *J. Chem. Theory Comput.* **2017**, *13*, 5479–5485.

(49) Sure, R.; Grimme, S. Corrected small basis set HartreeFock method for large systems. *J. Comput. Chem.* **2013**, *34*, 1672–1685.

(50) Grimme, S.; Brandenburg, J. G.; Bannwarth, C.; Hansen, A. Consistent structures and interactions by density functional theory with small atomic orbital basis sets. *J. Chem. Phys.* **2015**, *143*, 054107.

(51) Gauss, J. *Modern Methods and Algorithms of Quantum Chemistry*; John von Neumann Institute for Computing, Jülich, 2000; pp 541–592.

(52) Almlöf, J. Elimination of energy denominators in Møller-Plesset perturbation theory by a Laplace transform approach. *Chem. Phys. Lett.* **1991**, *181*, 319–320.

(53) Häser, M.; Almlöf, J. Laplace transform techniques in Møller-Plesset perturbation theory. *J. Chem. Phys.* **1992**, *96*, 489–494.

(54) Maurer, S. A.; Clin, L.; Ochsenfeld, C. Cholesky-decomposed density MP2 with density fitting: accurate MP2 and double-hybrid DFT energies for large systems. *J. Chem. Phys.* **2014**, *140*, 224112.

(55) Kussmann, J.; Ochsenfeld, C. Pre-selective screening for matrix elements in linear-scaling exact exchange solutions. *J. Chem. Phys.* **2013**, *138*, 134114.

(56) Kussmann, J.; Ochsenfeld, C. Preselective Screening for Linear-Scaling Exact Exchange-Gradient Calculations for Graphics Processing Units and General Strong-Scaling Massively Parallel Calculations. *J. Chem. Theory Comput.* **2015**, *11*, 918–922.

(57) Takatsuka, A.; Ten-no, S.; Hackbusch, W. Minimax approximation for the decomposition of energy denominators in Laplace-transformed Møller-Plesset perturbation theories. *J. Chem. Phys.* **2008**, *129*, 044112.

(58) Weigend, F. A fully direct RI-HF algorithm: Implementation, optimised auxiliary basis sets, demonstration of accuracy and efficiency. *Phys. Chem. Chem. Phys.* **2002**, *4*, 4285–4291.

(59) Kossmann, S.; Neese, F. Correlated ab Initio Spin Densities for Larger Molecules: Orbital-Optimized Spin-Component-Scaled MP2 Method. *J. Phys. Chem. A* **2010**, *114*, 11768–11781.

(60) Purvis III, G. D.; Bartlett, R. J. A full coupled-cluster singles and doubles model: The inclusion of disconnected triples. *J. Chem. Phys.* **1982**, *76*, 1910.

(61) Stanton, J. F.; Gauss, J.; Cheng, L.; Harding, M. E.; Matthews, D. A.; Szalay, P. G. CFOUR, Coupled-Cluster techniques for Componentutational Chemistry, a quantum-chemical program package. With contributions from A.A. Auer, R.J. Bartlett, U. Benedikt, C. Berger, D.E. Bernholdt, Y.J. Bomble, O. Christiansen, F. Engel, R. Faber, M. Heckert, O. Heun, M. Hilgenberg, C. Huber, T.-C. Jagau, D. Jonsson, J. Jusélius, T. Kirsch, K. Klein, W.J. Lauderdale, F. Lipparini, T. Metzroth, L.A. Mück, D.P. O'Neill, D.R. Price, E. Prochnow, C. Puzzarini, K. Ruud, F. Schiffmann, W. Schwalbach, C. Simmons, S. Stopkowicz, A. Tajti, J. Vázquez, F. Wang, J.D. Watts and the integral packages MOLECULE (J. Almlöf and P.R. Taylor), PROPS (P.R. Taylor), ABACUS (T. Helgaker, H.J. Aa. Jensen, P. Jørgensen, and J. Olsen), and ECP routines by A. V. Mitin and C. van Wüllen. For the current version, see http://www.cfour.de.

(62) Verlet, L. Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. *Phys. Rev.* **1967**, *159*, 98.

(63) Swope, W. C.; Andersen, H. C.; Berens, P. H.; Wilson, K. R. A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. *J. Chem. Phys.* **1982**, *76*, 637.

(64) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **2007**, *126*, 014101.

(65) Niklasson, A. M. N.; Steneteg, P.; Odell, A.; Bock, N.; Challacombe, M.; Tymczak, C. H.; Holmström, E.; Zheng, G.; Weber, V. Extended Lagrangian Born-Oppenheimer molecular dynamics with dissipation. *J. Chem. Phys.* **2009**, *130*, 214109.

(66) Boon, J., Philip; Symons, M. C. R.; Ushida, K.; Shida, T. Radical-cations of Aldehydes

and Ketones generated by Ionizing Radiation: an Electron Spin Resonance Study. *J. Chem. Soc. Perkin Trans. 2* **1984**, *0*, 1213.

(67) Danen, W. C.; Kensler, T. T. Electron Spin Resonance Study of Dialkylamino Free Radicals in Solution. *J. Am. Chem. Soc.* **1970**, *92*, 5235.

(68) Kozuch, S.; Gruzman, D.; Martin, J. M. L. DSD-BLYP: A General Purpose Double Hybrid Density Functional Including Spin Component Scaling and Dispersion Correction. *J. Phys. Chem. C* **2010**, *114*, 20801–20808.

(69) Toriyama, K.; Nunome, K.; Iwasaki, M. Structures and reactions of radical cations of some prototype alkanes in low temperature solids as studied by ESR spectroscopy. *J. Chem. Phys.* **1982**, *77*, 5891.

(70) Symons, M. C. R. Radical cations in condensed phases. *Chem. Soc. Rev.* **1984**, *13*, 393–439.

(71) Jacovella, U.; Stein, C. J.; Grütter, M.; Freitag, L.; Lauzin, C.; Reiher, M.; Merkt, F. Structure and dynamics of the radical cation of ethane arising from the Jahn-Teller and pseudo-Jahn-Teller effects. *Phys. Chem. Chem. Phys.* **2018**, *20*, 1072–1081.

(72) Shida, T.; Kato, T. ESR and optical studies on the cation-radical of pyridine in a γ-irradiated rigid matrix at low temperatures. *Chem. Phys. Lett.* **1979**, *68*, 106–110.

(73) Morton, J. R.; Preston, K. F.; Strach, S. J. Paramagnetic Oxides of Nitrogen Observed in a Sulfur Hexafluoride Matrix. *J. Phys. Chem.* **1979**, *83*, 533–536.

(74) Fessenden, R. W. Electron Spin Resonance Spectra of Isotopically Substituted Hydrocarbon Radicals. *J. Phys. Chem.* **1967**, *71*, 74–83.

(75) Fessenden, R. W.; Schuler, R. H. ESR Spectra and Structure of the Fluorinated Methyl Radicals. *J. Chem. Phys.* **1965**, *43*, 2704.

(76) Perera, S. A.; Salemi, L. M.; Bartlett, R. J. Hyperfine coupling constants of organic radicals. *J. Chem. Phys.* **1997**, *106*, 4061.

(77) Lan, T. N.; Kurashige, Y.; Yanai, T. Toward Reliable Prediction of Hyperfine Coupling Constants Using Ab Initio Density Matrix Renormalization Group Method: Diatomic $^2\Sigma$ and Vinyl Radicals as Test Cases. *J. Chem. Theory Comput.* **2014**, *10*, 1953–1967.

(78) McManus, H. J.; Fessenden, R. W.; Chipman, D. M. ^{13}C Hyperfine Constants of H$_2$CN$^{\cdot}$, H(HO)CN$^{\cdot}$, and $^{\cdot}$CONH$_2$. *J. Phys. Chem.* **1988**, *92*, 3781–3784.

(79) Knight, L. B.; Winiski, M.; Miller, P.; Arrington, C. A.; Feller, D. Neon matrix electron spin resonance and theoretical investigations of ^{10,11}BH$_2$, ^{12,13}CH$_2^-$,and ^{12,13}CH$_2^+$. *J. Chem. Phys.* **1989**, *91*, 4468.