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

Development of Efficient and Accurate Quantum Chemical Methods for Correlation Energies and Magnetic Resonance Parameters

Michael Anton Glasbrenner

aus

Starnberg

2021

Erklärung

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

Eidesstattliche Versicherung

Diese Dissertation wurde eigenständig und ohne unerlaubte Hilfe erarbeitet.

München, 03.03.2022 ______________________________
(Michael Glasbrenner)

Dissertation eingereicht am:	25.11.2021
1. Gutachterin/Gutachter:	Prof. Dr. Christian Ochsenfeld
2. Gutachterin/Gutachter:	Prof. Dr. Regina de Vivie-Riedle
Mündliche Prüfung am:	16.02.2022

Danksagung

Bei meinem Doktorvater, Prof. Dr. Christian Ochsenfeld, möchte ich mich für die Betreuung dieser Arbeit bedanken und dafür, dass er mich durch seine Vorlesungen für die Theoretische Chemie begeistert hat. Außerdem danke ich ihm für den fachlichen Austausch und die Unterstützung im Laufe der Jahre.

Ich bedanke mich auch bei Prof. Dr. Regina de Vivie–Riedle für das Anfertigen des Zweitgutachtens und bei den weiteren Mitgliedern der Promotionskommission.

Der Studienstiftung des deutschen Volkes danke ich für die ideelle und finanzielle Unterstützung im Rahmen eines Promotionsstipendiums.

Ich danke dem gesamten Arbeitskreis Ochsenfeld für die schöne Zeit und das nette Miteinander. Besonders bedanken möchte ich mich bei Dr. Sigurd Vogler für die Betreuung während der Anfangszeit meiner Promotion. Dr. Daniel Graf, Felix Bangerter und Dr. Sigurd Vogler danke ich für die hervorragende Zusammenarbeit bei den gemeinsamen Forschungsprojekten. Gökçen Savaşçi und den IT-Teams danke ich für den Einsatz für unseren Cluster.

Mein Dank gilt auch meiner Familie und Freunden für ihren steten Rückhalt und die Unterstützung, ohne welche diese Arbeit nicht möglich gewesen wäre.

Summary

Despite tremendous progress in electronic structure theory and computer technology over the last decades there still is great demand for more accurate and computationally efficient quantum chemical computation methods. In this thesis, contributions to this field of research are made by introducing several novel electronic structure methods for computing correlation energies, NMR shielding tensors, and electronic g-tensors.

An efficient method for computing correlation energies is derived from atomic orbital-based second-order Møller–Plesset perturbation theory (AO-MP2). Low-rank approximations are introduced for the pseudo-densities in the form of a Cholesky decomposition and for the two-electron integrals using the resolution-of-the-identity (RI) approximation with an attenuated Coulomb metric. Sparse linear algebra and integral screening based on Schwarz estimates allow to achieve linear scaling for large systems. This method is also extended to allow for computing NMR shieldings using an analytical second derivative. A suitable Z-vector approach is described, which ensures high efficiency by minimizing the number of coupled-perturbed self-consistent field equations that have to be solved. In addition, the first method for computing NMR shieldings with the post-Kohn–Sham random phase approximation (RPA) is presented. Benchmark calculations indicate a strong dependence of the accuracy on the employed orbitals. RPA shieldings with Hartree–Fock orbitals are shown to be significantly more accurate than MP2 shieldings, which renders RPA a promising method for NMR shielding calculations.

A second method for computing MP2 energies is presented, which employs the tensor hypercontraction (THC) ansatz. A low-scaling approach for decomposing two-electron integrals in THC format is described that relies on Cholesky-decomposed density matrices and RI with an attenuated Coulomb metric. Furthermore, it is demonstrated how the correlation energy can be efficiently computed from the decomposed integrals by exploiting their sparsity.

Finally, also methods for efficiently computing electronic g-tensors at the density functional theory (DFT) level are presented. Integral screening and sparse linear algebra approaches are shown to enable linear scaling for large molecules. Sublinear scaling is achieved for molecules with a local spin density by exploiting the locality of the contributions to the g-tensor. A benchmark study on the gauge-origin dependence of the electronic g-tensor is presented, which demonstrates that methods relying on a common gauge-origin can suffer from large errors, especially for extended molecules. Distributed gauge-origin approaches and a new way of choosing a common gauge-origin are suggested as potential alternatives.

List of Publications

This is a cumulative dissertation, comprising six articles in peer-reviewed journals. In the following, all articles are stated together with the author's contribution to each of them.

I **Michael Glasbrenner**, Daniel Graf, Christian Ochsenfeld
"Efficient Reduced-Scaling Second-Order Møller–Plesset Perturbation Theory with Cholesky-Decomposed Densities and an Attenuated Coulomb Metric",
J. Chem. Theory Comput. **16**, 6856 (2020).
Contributions by the author: Conjoint development of the theory, major parts of the implementation, all calculations, and writing of the manuscript

II Felix H. Bangerter, **Michael Glasbrenner**, Christian Ochsenfeld
"Low-Scaling Tensor Hypercontraction in the Cholesky Molecular Orbital Basis Applied to Second-Order Møller–Plesset Perturbation Theory",
J. Chem. Theory Comput. **17**, 211 (2021).
Contributions by the author: Assisted with the derivations, with the implementation, and with the writing of the manuscript

III **Michael Glasbrenner**, Sigurd Vogler, Christian Ochsenfeld
"Efficient low-scaling computation of NMR shieldings at the second-order Møller–Plesset perturbation theory level with Cholesky-decomposed densities and an attenuated Coulomb metric",
J. Chem. Phys. **155**, 224107 (2021).
Contributions by the author: Most of the derivations, major parts of the implementation, all calculations, and writing of the manuscript

IV **Michael Glasbrenner**, Daniel Graf, Christian Ochsenfeld
"Benchmarking the Accuracy of the Direct Random Phase Approximation and σ-Functionals for NMR Shieldings",
J. Chem. Theory Comput. **18**, 192 (2022).
Contributions by the author: Conjoint development of the theory, major parts of the implementation, all calculations, and writing of the manuscript

V **Michael Glasbrenner**, Sigurd Vogler, Christian Ochsenfeld
“Gauge-origin dependence in electronic g-tensor calculations”,
J. Chem. Phys. **148**, 214101 (2018).
Contributions by the author: All of the derivations, the implementation, all calculations, and writing of the manuscript

VI **Michael Glasbrenner**, Sigurd Vogler, Christian Ochsenfeld
“Linear and sublinear scaling computation of the electronic g-tensor at the density functional theory level”,
J. Chem. Phys. **150**, 024104 (2019).
Contributions by the author: All of the derivations, the implementation, all calculations, and writing of the manuscript

Contents

Chapter 1

Introduction

Quantum chemistry is concerned with the application of the fundamental laws of quantum mechanics to chemical systems. In its early years, quantum chemistry could provide an improved *qualitative* understanding of many chemical phenomena and, e.g., refine concepts of chemical bonding.[1] With the enormous progress in modern computer technology and the development of efficient numerical methods and approximations, *quantitatively* accurate quantum chemical calculations on increasingly large molecules became possible.

The majority of quantum chemical methods available today can be grouped into two categories: wave function based methods and density functional methods. Wave function methods employ a certain ansatz for the many-electron wave function and are systematically improvable towards the exact solution of the Schrödinger equation, but are often expensive in terms of computational resources if high accuracy is desired. Some wave function methods are based on the variational principle[2] and minimize the energy expectation value assuming a specific form of the wave function. The Hartree–Fock (HF) method[3–5] belongs to this class and uses a single Slater determinant as an ansatz for the wave function. HF theory describes the electron-electron interaction in a mean field manner; the corresponding error of this approximation compared to the exact solution of the electronic Schrödinger equation is called the correlation energy.[6] One distinguishes static and dynamic electron correlation. In systems with static correlation, a single determinant does not provide a qualitatively correct description of the electronic structure, and multi-reference methods like complete active space self-consistent field (CASSCF)[7] or the density matrix renormalization group (DMRG)[8–10] need to be used. Dynamic correlation accounts for the small, but numerous contributions from excited determinants, which are created by removing electrons from occupied molecular orbitals (MOs) in the reference determinant and placing them in virtual MOs.

Configuration interaction (CI)[11] methods are one class of variational methods for describing dynamic correlation and expand the wave function as a linear combination of Slater determinants. The CI wave function contains contributions both from the HF determinant and from excited determinants. CI methods are characterized by the maximum number of allowed excitations. In the CI singles and doubles method (CISD), e.g., determinants with one and two excited electrons are included, but no higher excitations. Unless all possible Slater determinants are included, which is called full CI[12] and is equivalent to the exact solution of the Schrödinger equation in the given basis, truncated CI methods are not size-extensive. This means that chemical systems with different sizes and numbers of electrons are not treated at the same level of accuracy.[13] In general, CI recovers a decreasing fraction of the total correlation energy for increasingly large molecules, which is a serious drawback.

Coupled cluster (CC) methods[14,15] are a size-extensive alternative and employ an exponential ansatz for the wave function. Most formulations of coupled cluster are not variational.[16] However, as the variational principle only holds for absolute energies, but not for energy differences, the non-variational nature of CC is of little concern in practice. It is outweighed by the benefits of size-extensivity and faster convergence of CC to the full CI limit compared to truncated CI methods.[13]

Perturbation theory[17] is another theoretical framework that allows to approximately solve the many-electron Schrödinger equation. Møller–Plesset perturbation theory[18] (MPPT) provides a hierarchy of methods for computing correlation energies. In the most popular MPPT method, the perturbation series is truncated after second order, which results in the MP2 model. MP2 is one of the computationally cheapest methods for describing electron correlation and allows to account for dispersion effects, which is highly important, e.g., in the study of biomolecules. Perturbation theory can also be applied in the context of CC theories. In the CCSD(T) method, e.g., a perturbatively computed triples correction is added to the CC singles and doubles (CCSD) energy.[19] The highly accurate CCSD(T) method is widely regarded as the "gold standard of quantum chemistry", but can only be applied to rather small molecules due to an unfavorable $\mathcal{O}(N^7)$ scaling.

Density functional theory (DFT) methods aim to circumvent the use of the wave function in favor of the electron density, which is a much simpler function as it only depends on three coordinates irrespective of the number of electrons. A sound theoretical basis for DFT was provided by Hohenberg and Kohn[20] in 1964, who proved that the exact energy can be expressed as a functional of the electron density. However, the exact functional remains elusive up to this day and therefore approximate functionals need to be used in practice. Most popular density functional methods are based on the Kohn–Sham formulation of DFT[21] and have similar computational cost as HF theory, albeit at significantly higher accuracy. Some methods display properties of both density functional and wave function methods and cannot be unambiguously assigned to one category. These include methods based on the adiabatic-connection fluctuation-dissipation theorem (ACFDT)[22–26] such as the random phase approximation (RPA),[27] which allow to systematically improve density functional methods. Another example for methods at the interface of density functional and wave function methods are double-hybrid density functionals, which contain an MP2-like contribution and have been shown to provide highly accurate energies.[28]

Even though many of the mentioned quantum chemistry methods allow for much more efficient calculations than the exact "brute-force" full CI approach, the high requirements of computational resources still limit their applicability, especially for larger molecules. Of high importance in this context is the scaling of a method with respect to the molecule size. HF theory displays a comparatively low $\mathcal{O}(N^4)$ scaling, which means that doubling the system size results in an increase of the computational cost by a factor of $2^4 = 16$. Other wave function methods show an even higher scaling; MP2, e.g., scales as $\mathcal{O}(N^5)$ and CC methods scale at least as $\mathcal{O}(N^6)$. Applying these methods to larger molecules quickly becomes infeasible. However, the steep scaling is often unphysical. The electronic structure of many molecules is local,[29,30] which explains the success of local bonding models or the concept of functional groups. Therefore, it should, in principle, be possible to reduce the scaling of quantum chemical methods, ideally to linear or $\mathcal{O}(N^1)$, for large enough systems with a significant electronic gap. Much work has been done on reduced-scaling methods that are able to exploit the locality of the electronic structure both for energies and molecular properties.[31]

This thesis is also mainly concerned with the development of novel quantum chemical methods with reduced scaling, especially at the MP2 level of theory. The presented MP2 methods are based on atomic orbital-based MP2 (AO-MP2),[32–36] which expresses the MP2 equation solely in terms of atomic orbitals (AOs) avoiding the delocalized canonical MOs. While the scaling of AO-MP2 can be reduced to linear using integral screening,[36] the high computational cost, especially for extended basis sets, still limits its applicability to larger molecules. For this reason, additional approximations such as the resolution-of-the-identity (RI)[37,38] and Cholesky decomposition[39–42] of pseudo-density matrices have been introduced into AO-MP2 theory, resulting in the RI-CDD-MP2 method from Maurer *et al.*[43] Together with distance-dependent integral screening[44] cubic scaling with a low prefactor was shown to be possible. Also, a linear-scaling version of RI-CDD-MP2 with local density fitting[45] was presented by Maurer *et al.*[43] However, for the latter approach, the overhead due to local density fitting was substantial and computational savings occurred only for very large systems. In **Publication I**, RI-CDD-MP2 is improved further by introducing an attenuated Coulomb metric[46–49] for the RI. Sparse linear algebra approaches are described, which allow to exploit the resulting sparsity in the three-center integrals and reduce computational cost and memory demands. In addition, it is shown that a separate treatment of Coulomb and exchange contributions with specialized algorithms is beneficial, if an attenuated Coulomb metric is used. Efficient integral screening based on Schwarz estimates[50] is used for the time-dominating exchange contributions. Calculations on DNA strands with up to 1052 atoms and 11230 basis functions demonstrate the potential of the method for applications on large biomolecules.

An alternative path to efficient reduced-scaling MP2 is explored in **Publication II** by combining the RI-CDD-MP2 method with the tensor hypercontraction[51,52] (THC) ansatz, which provides a decomposition of four-center integral tensors into matrices. An approach for obtaining the THC decomposition with reduced scaling is described, which employs an attenuated Coulomb metric and Cholesky-decomposed density matrices. Furthermore, it is shown how the MP2 energy can be computed efficiently from the THC-decomposed integrals by exploiting the inherent sparsity of the obtained factors.

In addition to the computation of energies, quantum chemical methods can also be used for calculating various molecular properties. Especially the computation of spectroscopically measurable properties provides important links between theory and experiments. In this thesis, quantum chemical methods are presented for computing nuclear shielding tensors measurable by nuclear magnetic resonance (NMR) spectroscopy and electronic g-tensors of electron paramagnetic resonance (EPR) spectroscopy. Both properties can be obtained using a second derivative of the energy.

A method for computing NMR shieldings based on AO-MP2 has previously been introduced by Maurer *et al.*[53] In **Publication III**, the method for computing MP2 energies from **Publication I** is extended for NMR shieldings using an analytical second derivative. The employed approximations significantly extend the applicability to larger systems compared to the previous method from Maurer *et al.*[53] A Z-vector approach[54,55] is used, which reduces the number of coupled self-consistent field (CPSCF) equations[56] that need to be solved and allows to efficiently compute the entire set of shieldings for a given molecule.

In **Publication IV**, the first method for computing NMR shieldings at the post-Kohn–Sham RPA level of theory is introduced using a non-trivial numerical derivative approach. While better efficiency could likely be obtained with an analytical derivative combined with a Z-vector approach, the described implementation allowed us to perform benchmark calculations on a set of small molecules and compare the computed shieldings to high-level CCSD(T) shieldings. The results show that RPA based on a HF reference gives highly accurate NMR shieldings. In particular, the obtained accuracy is higher than for MP2, which makes the method attractive considering also the lower $\mathcal{O}(N^4)$ scaling. In addition, the basis set convergence of the method and the accuracy of the closely related σ-functionals[57] was analyzed in this study.

The electronic g-tensor is a related molecular property and describes the interaction between the spin of an unpaired electron and an external magnetic field. In **Publication VI**, an efficient method for computing g-tensors at the DFT level of theory is described. Linear scaling for large systems is achieved with integral screening approaches. It is shown that for molecules with a localized spin density, also asymptotically constant, $\mathcal{O}(N^0)$ or $\mathcal{O}(1)$, scaling is possible by exploiting the locality of the contributions to the g-tensor. As for NMR shieldings, the problem of gauge-origin dependence[58] of the computed values can occur in g-tensor calculations. **Publication V** provides an extensive benchmark study on the gauge-origin dependence of the g-tensor. It is demonstrated that methods relying on a common gauge-origin do not reliably give accurate g-tensors. The findings suggest that, if possible, distributed gauge-origin methods should be used for g-tensors, which has not been done in most of the previous implementations described in the literature.[59–76] Furthermore, it is shown that for molecules with a single spin-center and well-localized spin density, a common gauge-origin can be sufficient, but should be chosen according to the spin density distribution. Previous common gauge-origin approaches do not fulfill this requirement and we suggest a novel way of choosing an appropriate gauge-origin in **Publication V**, whose suitability is illustrated using numerical data.

The following chapters of this cumulative thesis are structured as follows: In Chapter 2 the theoretical basis of the quantum chemical methods and many of the approximations used in this thesis is presented. Chapter 3 contains the main body of work in the form of the published articles and manuscripts, which are reproduced in their entirety. Finally, a conclusion with an outlook on potential future lines of further research is provided in Chapter 4.

Chapter 2

Theoretical Background

2.1 Molecular Schrödinger Equation

The central goal in non-relativistic quantum chemistry is to solve the molecular Schrödinger equation[77] or at least to accurately approximate its solution. In its time-independent formulation, the Schrödinger equation is given by

$$\hat{H}\Psi(\{\mathbf{r}\},\{\mathbf{R}\}) = E\Psi(\{\mathbf{r}\},\{\mathbf{R}\}). \tag{2.1}$$

In this equation, E stands for the energy of the molecular system. $\{\mathbf{r}\}$ and $\{\mathbf{R}\}$ denote the sets of coordinates of all electrons and nuclei, respectively. Ψ is the wave function, which describes the spatial distributions of electrons and nuclei. $|\Psi|^2$ can be interpreted as a probability density for finding these particles at certain positions in space. Note that Eq. 2.1 has in general many different solutions with different energies E and wave functions Ψ, which correspond to the ground state and to excited states of the molecule. In this thesis, the focus lies on ground state energies and properties. $\hat{H}$ denotes the Hamilton operator, which in atomic units has the following form for molecules:

$$\begin{aligned}\hat{H} &= \hat{T}_\mathrm{e} + \hat{T}_\mathrm{N} + \hat{V}_\mathrm{eN} + \hat{V}_\mathrm{ee} + \hat{V}_\mathrm{NN} \\ &= -\sum_i \frac{1}{2}\nabla_i^2 - \sum_A \frac{1}{2M_A}\nabla_A^2 - \sum_i\sum_A \frac{Z_A}{r_{iA}} + \frac{1}{2}\sum_i\sum_j \frac{1}{r_{ij}} + \frac{1}{2}\sum_A\sum_B \frac{Z_A Z_B}{R_{AB}}.\end{aligned} \tag{2.2}$$

The individual terms in Eq. 2.2 correspond to the kinetic energy of the electrons $\hat{T}_\mathrm{e}$, the kinetic energy of the nuclei $\hat{T}_\mathrm{N}$, the electrostatic attraction of nuclei and electrons $\hat{V}_\mathrm{eN}$, the electron-electron repulsion $\hat{V}_\mathrm{ee}$, and the nuclear-nuclear repulsion $\hat{V}_\mathrm{NN}$, respectively. The indices i and j denote electrons, while A and B denote nuclei. M_A and Z_A are mass and charge of nucleus A. r and R denote distances between particles.

The problem of solving Eq. 2.1 can be simplified considerably by applying the Born–Oppenheimer (BO) approximation,[78] which allows to separate the motion of nuclei and electrons. It is motivated by the fact that the nuclei are much heavier than the electrons and their motion usually happens on different time scales. The molecular wave function can in general be expanded in the following form:

$$\Psi(\{\mathbf{r}\},\{\mathbf{R}\}) = \sum_k \Psi_{\mathrm{e},k}(\{\mathbf{r}\};\{\mathbf{R}\})\Psi_{\mathrm{N},k}(\{\mathbf{R}\}), \tag{2.3}$$

where k labels the electronic states. Applying the BO approximation allows to replace Eq. 2.1 with one electronic and one nuclear Schrödinger equation:

$$\underbrace{\left(\hat{T}_{\mathrm{e}} + \hat{V}_{\mathrm{eN}} + \hat{V}_{\mathrm{ee}} + \hat{V}_{\mathrm{NN}}\right)}_{\hat{H}_{\mathrm{e}}} \Psi_{\mathrm{e},k}(\{\mathbf{r}\};\{\mathbf{R}\}) = E_{\mathrm{e},k}\Psi_{\mathrm{e},k}(\{\mathbf{r}\};\{\mathbf{R}\}), \tag{2.4}$$

$$\left(\hat{T}_{\mathrm{N}} + E_{\mathrm{e},k}(\mathbf{R})\right)\Psi_{\mathrm{N},k}(\{\mathbf{R}\}) = E\Psi_{\mathrm{N},k}(\{\mathbf{R}\}). \tag{2.5}$$

The nuclear Schrödinger equation (Eq. 2.5) describes the quantum-mechanical motion of the nuclei in the potential created by the electrons. Solving the (time-dependent) nuclear Schrödinger equation allows to rigorously treat the nuclear dynamics during chemical reactions and light-induced processes. In the rest of this thesis, the focus lies on the electronic Schrödinger equation given in Eq. 2.4, which parametrically depends on the nuclear coordinates $\mathbf{R}$. For the sake of simplicity, the subscript "e" is dropped in the following; $\hat{H}$, E, and Ψ shall denote the *electronic* Hamiltonian, energy, and wave function.

For all but the simplest systems, an analytical solution of the electronic Schrödinger equation is not possible, and one has to resort to approximations. Many of the commonly used approximation methods are based on the variational principle,[2] which states that an approximate wave function $\tilde{\Psi}$ always gives an energy expectation value $E_{\tilde{\Psi}}$ greater than or equal to the exact ground state energy E_0 associated with the ground state wave function Ψ_0:

$$E_{\tilde{\Psi}} = \frac{\left\langle \tilde{\Psi} \middle| \hat{H} \middle| \tilde{\Psi} \right\rangle}{\langle \tilde{\Psi} | \tilde{\Psi} \rangle} \geq \frac{\left\langle \Psi_0 \middle| \hat{H} \middle| \Psi_0 \right\rangle}{\langle \Psi_0 | \Psi_0 \rangle} = E_0. \tag{2.6}$$

This suggests the following strategy for determining approximate wave functions and energies: First, an ansatz for the wave function that depends on some parameters needs to be chosen. Then, the energy as a function of these parameters needs to be minimized. One suitable ansatz for an electronic wave function is a single Slater determinant[79] or a linear combination of such determinants. A Slater determinant is a determinant containing one-electron functions $\varphi(\boldsymbol{\tau})$ called molecular orbitals (MOs):

$$\Psi(\boldsymbol{\tau}_1,\boldsymbol{\tau}_2,...,\boldsymbol{\tau}_{N_{\mathrm{el}}}) = \frac{1}{\sqrt{N_{\mathrm{el}}!}} \begin{vmatrix} \varphi_1(\boldsymbol{\tau}_1) & \varphi_2(\boldsymbol{\tau}_1) & \dots & \varphi_{N_{\mathrm{el}}}(\boldsymbol{\tau}_1) \\ \varphi_1(\boldsymbol{\tau}_2) & \varphi_2(\boldsymbol{\tau}_2) & \dots & \varphi_{N_{\mathrm{el}}}(\boldsymbol{\tau}_2) \\ \vdots & \vdots & & \vdots \\ \varphi_1(\boldsymbol{\tau}_{N_{\mathrm{el}}}) & \varphi_2(\boldsymbol{\tau}_{N_{\mathrm{el}}}) & \dots & \varphi_{N_{\mathrm{el}}}(\boldsymbol{\tau}_{N_{\mathrm{el}}}) \end{vmatrix}, \tag{2.7}$$

where N_{el} is the number of electrons. The MOs depend on the combined coordinates $\boldsymbol{\tau}$, which encompass three spatial coordinates $\mathbf{r} = (x, y, z)$ and a spin coordinate ω. An MO

φ_p can be written as the product of a spatial orbital ϕ_p and a spin function, which may be either an α or β function:

$$\varphi_p(\boldsymbol{\tau}_1) = \begin{cases} \phi_p(\mathbf{r}_1)\alpha(\omega) \\ \phi_p(\mathbf{r}_1)\beta(\omega) \end{cases} . \tag{2.8}$$

Slater determinants fulfill two necessary requirements for many-electron wave functions. Firstly, they treat all electrons on an equal footing and thus account for the fact that electrons are indistinguishable particles. Secondly, Slater determinants are anti-symmetric under the exchange of two electrons; this requirement, which is also called the Pauli principle,[80] has to hold for any wave function of fermions.

2.2 Electronic Structure Methods

2.2.1 Hartree–Fock Theory

In Hartree–Fock (HF) theory,[3–5] a single Slater determinant is used as ansatz for the wave function. If the Slater determinant is inserted into the energy expectation value, the energy becomes a functional of the molecular orbitals. Minimization of this functional with respect to the MOs, under the constraint that the MOs are orthonormal, leads to a set of coupled one-electron equations called canonical Hartree–Fock equations:

$$\hat{F}(\boldsymbol{\tau}_1)\,\varphi_i(\boldsymbol{\tau}_1) = \epsilon_i\,\varphi_i(\boldsymbol{\tau}_1), \tag{2.9}$$

where $\hat{F}$ is the so-called Fock operator and ϵ_i is the orbital energy of orbital φ_i. The Fock operator is an effective one-particle operator, which describes the electron-electron interaction using a mean-field potential:

$$\hat{F}(\boldsymbol{\tau}_1)\,\varphi_i(\boldsymbol{\tau}_1) = \hat{h}(\boldsymbol{\tau}_1)\,\varphi_i(\boldsymbol{\tau}_1) + \sum_j \Big[\hat{J}_j\varphi_i(\boldsymbol{\tau}_1) - \hat{K}_j\varphi_i(\boldsymbol{\tau}_1)\Big], \tag{2.10}$$

where

$$\hat{J}_j\varphi_i(\boldsymbol{\tau}_1) = \int d\boldsymbol{\tau}_2\; \varphi_j^*(\boldsymbol{\tau}_2)\; \varphi_j(\boldsymbol{\tau}_2)\frac{1}{r_{12}}\varphi_i(\boldsymbol{\tau}_1), \tag{2.11}$$

$$\hat{K}_j\varphi_i(\boldsymbol{\tau}_1) = \int d\boldsymbol{\tau}_2\; \varphi_j^*(\boldsymbol{\tau}_2)\; \varphi_i(\boldsymbol{\tau}_2)\frac{1}{r_{12}}\varphi_j(\boldsymbol{\tau}_1). \tag{2.12}$$

The one-electron part $\hat{h}$ includes the kinetic energy of the electrons and the electron-nuclear attraction. The Coulomb operator $\hat{J}$ describes the electrostatic repulsion of electrons, whereas the exchange operator $\hat{K}$ does not have a classical interpretation and results from antisymmetry of the wave function.

In practice, the spatial part of the MOs is usually expanded as a linear combination of basis functions χ, for which most commonly Gaussian basis functions are used:[81]

$$\phi_i(\mathbf{r}_1) = \sum_\nu C_{\nu i}\chi_\nu(\mathbf{r}_1), \tag{2.13}$$

If this ansatz is inserted into Eq. 2.9, and one multiplies from the left with the ket $\langle \chi_\mu |$, the Roothaan–Hall equation[82,83] is obtained:

$$\sum_\nu F_{\mu\nu} C_{\nu i} = \sum_\nu S_{\mu\nu} C_{\nu i} \epsilon_i, \tag{2.14}$$

where $\mathbf{S}$ is the so-called overlap matrix:

$$S_{\mu\nu} = \int d\mathbf{r_1}\, \chi_\mu^*(\mathbf{r_1}) \chi_\nu(\mathbf{r_1}), \tag{2.15}$$

and $\mathbf{F}$ is the Fock matrix, whose elements can be computed as follows:

$$F_{\mu\nu} = \int d\mathbf{r_1}\, \chi_\mu^*(\mathbf{r_1}) \hat{F} \chi_\nu(\mathbf{r_1}). \tag{2.16}$$

The Roothaan–Hall equation is a generalized eigenvalue problem and can be solved using the Löwdin orthogonalization.[84] As the Fock operator depends on the MOs and thus on its own eigenfunctions, the canonical HF equation needs to be solved iteratively in the self-consistent field (SCF) procedure. For a system with N basis functions, N MOs can be obtained from the solution of the Roothaan–Hall equation. These are filled with electrons according to the *Aufbau principle*, which means that the N_el energetically lowest MOs become occupied MOs. In the following, the indices $i, j, k, l, ...$ are used for denoting occupied MOs. The remaining $(N - N_\mathrm{el})$ MOs are unoccupied or virtual and will be denoted with indices $a, b, c, d, ...$

Different variants of HF theory exist, which impose different conditions on the spatial orbitals. For closed-shell systems, restricted Hartree–Fock (RHF) theory is appropriate, which requires that each spatial orbital is occupied with one α and one β electron. For open-shell systems, the most popular variants are unrestricted Hartree–Fock (UHF)[85] and restricted open-shell Hartree–Fock (ROHF).[86] In ROHF theory, all spatial orbitals are either doubly occupied with electrons of opposite spin or singly occupied. This ensures that the ROHF wave function is an eigenfunction of the $\hat{S}^2$ operator. This is not guaranteed for a UHF wave function, in which α and β MOs are not required to have identical spatial parts. In UHF, an unphysical mixing of states with different total spin can occur for open-shell systems, which is called spin contamination.[87] On the other hand, the UHF wave function has more flexibility and can account for spin polarization. In **Publications V** and **VI** methods for computing electronic g-tensors are studied; as the g-tensor is a property of molecules with unpaired electrons, UHF and the related unrestricted DFT (UDFT) is used in these works.

2.2.2 Møller–Plesset Perturbation Theory

Perturbation theory[17] is a method for solving the Schrödinger equation with a Hamiltonian $\hat{H}$ and requires that the eigenvalues and eigenstates of a similar Hamiltonian $\hat{H}_0$ are known. The difference between $\hat{H}$ and $\hat{H}_0$ is called the perturbation $\hat{V}$. Furthermore, a parameter λ

is introduced, which controls the strength of the perturbation and renders the Hamiltonian λ-dependent:

$$\hat{H}(\lambda) = \hat{H}_0 + \lambda\hat{V}. \tag{2.17}$$

If the Hamiltonian from Eq. 2.17 is inserted into the Schrödinger equation, also the energy and the wave function become functions of λ. In perturbation theory, these functions $E(\lambda)$ and $\Psi(\lambda)$ are expanded in a Taylor series around $\lambda = 0$. From this expansion, a hierarchy of equations can be derived, which allow to obtain wave function and energy corrections to a certain order in λ. At the end of the derivation, λ is set to 1, which corresponds to the (fully perturbed) system of interest.

In Møller–Plesset perturbation theory (MPPT),[18] electron correlation effects are described perturbatively. The perturbation operator in MPPT is given by the difference between the exact electron-electron interaction and the mean-field interaction in HF theory:

$$\hat{V}^{\mathrm{MPPT}} = \hat{H} - \sum_i \hat{F}(\boldsymbol{\tau}_i), \tag{2.18}$$

$$\hat{H}_0^{\mathrm{MPPT}} = \sum_i \hat{F}(\boldsymbol{\tau}_i). \tag{2.19}$$

The sum of the zeroth and first order energy in MPPT is equal to the HF energy. In second-order MPPT (MP2), which is one of the computationally cheapest methods for approximately computing electron correlation effects, the second-order energy correction is included. For a closed-shell molecule, the following expression holds for the MP2 energy contribution:

$$E_{\mathrm{MP2}} = \sum_{ijab} \frac{(ia|jb)\,[2(ai|bj) - (bi|aj)]}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b}. \tag{2.20}$$

This formula contains electron repulsion integrals (ERIs) in the MO basis, which are defined as follows:

$$(ia|jb) = \iint d\mathbf{r}_1 d\mathbf{r}_2\, \phi_i^*(\mathbf{r}_1)\phi_a(\mathbf{r}_1)\frac{1}{r_{12}}\phi_j^*(\mathbf{r}_2)\phi_b(\mathbf{r}_2). \tag{2.21}$$

Computing the MP2 energy formally scales as $\mathcal{O}(N^5)$. Several reduced-scaling MP2 methods have been developed in order to enable calculations of the correlation energy of large molecules. Local correlation methods[88–93] employ localized molecular orbitals (LMOs)[94] and partition the correlation energy into contributions from LMOs and associated orbital domains. Disadvantages of local correlation methods are that often large orbital domains are required for accurate results and discontinuities in the potential energy surfaces can occur.[95–97] Other reduced scaling methods partition the molecule into fragments, like the fragment molecular orbital (FMO)[98,99] or the divide-and-conquer approach.[100] A different class of approaches employ a Laplace transform for decomposing the orbital energy denominator in Eq. 2.20:[101–103]

$$\frac{1}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b} = \int_0^\infty dt\, e^{-t(\epsilon_i+\epsilon_j-\epsilon_a-\epsilon_b)} \approx \sum_\alpha w_\alpha\, e^{-t_\alpha(\epsilon_i+\epsilon_j-\epsilon_a-\epsilon_b)}. \tag{2.22}$$

The integral in Eq. 2.22 can be accurately approximated numerically; in most cases 5-8 integration points are sufficient.[102] With the help of a Laplace transform, a formulation of MP2 purely based on AOs (AO-MP2) can be devised.[32–36] In combination with integral screening approaches, the scaling of AO-MP2 can be reduced to linear. The RI-CDD-MP2 method[43] further increases the efficiency of AO-MP2 by using the resolution-of-the-identity approximation (see Sec. 2.4.1) and a Cholesky decomposition[39–42] of pseudo-density matrices (see Sec. 2.4.3). Also distance-dependent QQR[44] integral screening is used. In **Publication I**, RI-CDD-MP2 is improved further by introducing a local attenuated Coulomb metric[46–48] for the RI and by efficiently exploiting the resulting additional sparsity in the three-center integrals. In **Publication II**, an alternative path to efficient MP2 energy computations on large molecules is explored by using the tensor hypercontraction approach (see also Sec. 2.4.2).

In recent years, also several variations of conventional MP2 theory have been proposed. In spin-component scaled MP2 (SCS-MP2) from Grimme,[104] the Coulomb and exchange type contributions to the MP2 energy are scaled using different constant factors, which has been shown to further improve the accuracy. In the related scaled opposite-spin MP2 (SOS-MP2) from Jung *et al.*,[105] the computationally demanding exchange contributions can be neglected entirely without degrading the accuracy of MP2. In **Publications I-III**, calculations with both conventional MP2 and SOS-MP2 were carried out showing that the SOS-approximation leads to significant computational savings for the developed methods. The recently introduced class of double-hybrid density functionals[28] contains an MP2-like energy contribution. Double-hybrid functionals have been shown to give excellent accuracy for energies[106] and different molecular properties.[107,108] All of the presented developments for efficient MP2 calculations in this thesis are also applicable in the context of double-hybrid density functionals.

2.2.3 Density Functional Theory

Density functional theory (DFT) aims to compute the energy and properties of molecules without resorting to the electronic wave function, but rather using the electron density $\rho(\mathbf{r})$ as the central quantity. This would be desirable, because the electron density is a much simpler quantity and depends only one three spatial coordinates, while the wave function depends on $3\,N_{\mathrm{el}}$ spatial coordinates. In 1964, Hohenberg and Kohn[20] proved that the exact energy can be expressed as a functional of the electron density. However, the exact density functional remains unknown to date. For this reason, approximate functionals are used in practice.[109]

Contrary to orbital-free DFT approaches,[110] the popular Kohn–Sham DFT (KS-DFT)[21] re-introduces MOs. This is advantageous because it enables an accurate computation of the kinetic energy (exact only for *non-interacting* systems), which is challenging for orbital-free DFT. The electron density $\rho(\mathbf{r})$ can be obtained from the MOs as shown in Eq. 2.23:

$$\rho(\mathbf{r}) = \sum_i \phi_i^*(\mathbf{r})\phi_i(\mathbf{r}). \tag{2.23}$$

Like in HF theory, a Slater determinant composed of orthonormal MOs is used in KS-DFT. These MOs fulfill an equation that has the same structure as the canonical HF equation shown in Eq. 2.9:

$$\hat{h}^{\mathrm{KS}}\varphi_i(\boldsymbol{\tau}_1) = \epsilon_i\varphi_i(\boldsymbol{\tau}_1). \tag{2.24}$$

Like the Fock operator from HF theory, $\hat{h}^{\mathrm{KS}}$ is a one-particle operator. It contains the following contributions:

$$\hat{h}^{\mathrm{KS}}(\mathbf{r}_i) = -\frac{\nabla_i^2}{2} \underbrace{-\sum_A \frac{Z_A}{r_{iA}} + \int d\mathbf{r}_2 \frac{\rho(\mathbf{r}_2)}{r_{i2}} + V_{\mathrm{xc}}(\mathbf{r}_i)}_{V_{\mathrm{KS}}}, \tag{2.25}$$

where V_{KS} is the Kohn–Sham potential. The first three terms on the right-hand side of Eq. 2.25 account for the kinetic energy of the electrons, the nuclear-electron attraction, and the Coulomb interaction; all these interactions are also included in the Fock operator from HF theory. The exchange operator $\hat{K}$ from Eq. 2.10, however, is replaced by the exchange-correlation potential V_{xc}. As its name suggests, V_{xc} should ideally account for all exchange and electron correlation effects (and include a small correction for the kinetic energy). However, only approximate expressions for V_{xc} are known. LDA functionals depend on $\rho(\mathbf{r})$ only, while GGA functionals also depend on the gradient of $\rho(\mathbf{r})$.[109] Hybrid functionals also include a fraction of exact Hartree–Fock exchange and often provide better accuracy.[111]

After iteratively solving Eq. 2.24, the DFT energy can be computed from the obtained MOs and the electron density as follows:

$$E_{\mathrm{DFT}} = -\sum_i \int d\mathbf{r}\, \phi_i^*(\mathbf{r}) \frac{\nabla_i^2}{2} \phi_i(\mathbf{r}) - \sum_A \int d\mathbf{r} \frac{Z_A\, \rho(\mathbf{r})}{|\mathbf{r} - \mathbf{R}_A|} + \iint d\mathbf{r}_1 d\mathbf{r}_2 \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{r_{12}} + E_{\mathrm{xc}}[\rho], \tag{2.26}$$

where E_{xc} is the exchange-correlation energy, which is a functional of ρ and can be related to V_{xc} by differentiating with respect to the electron density:

$$V_{\mathrm{xc}}[\rho](\mathbf{r}) = \frac{\partial E_{\mathrm{xc}}[\rho]}{\partial \rho(\mathbf{r})}. \tag{2.27}$$

For energetics and many molecular properties, DFT can provide significantly higher accuracy than Hartree–Fock at similar or even lower computational cost (lower in the case of pure density functionals without HF exchange). DFT is also more accurate than HF for electronic g-tensors;[112] for this reason, we use unrestricted DFT (UDFT) in addition to UHF in our studies on g-tensors in **Publications V** and **VI**. The excellent cost-to-accuracy ratio makes DFT a highly popular method in modern computational chemistry. Furthermore, DFT calculations can also serve as a starting point for the random phase approximation method, which is discussed in Sec. 2.2.4.

2.2.4 Random Phase Approximation

The adiabatic-connection formalism first studied by Langreth and Perdew[22,23] and by Gunnarsson and Lundqvist[113] allows to derive an exact expression for the correlation energy. This formalism is based on an interpolation between the so-called interacting system with the Schrödinger equation given by Eq. 2.4 and the non-interacting Kohn–Sham system. The corresponding Hamiltonian depending on the coupling strength λ reads as follows:

$$\hat{H}^{\lambda} = \hat{T}_{\mathrm{e}} + \lambda \hat{V}_{ee} + \hat{V}^{\lambda}. \tag{2.28}$$

$\hat{V}^{\lambda}$ is defined such that the ground state density is independent of λ. For $\lambda = 0$, $\hat{V}^{\lambda}$ reduces to the Kohn–Sham potential $\hat{V}^{\mathrm{KS}}$, which was defined in Eq. 2.25. In the fully interacting limit with $\lambda = 1$, $\hat{H}^{\lambda}$ is equal to the exact Hamiltonian. It can be shown that the correlation energy in the adiabatic-connection formalism is given by:[22,113]

$$E_{\mathrm{corr}} = \int\limits_0^1 d\lambda \left\langle \Psi_0^{\lambda} \middle| \hat{V}_{\mathrm{ee}} \middle| \Psi_0^{\lambda} \right\rangle - \left\langle \Psi_{\mathrm{KS}} \middle| \hat{V}_{\mathrm{ee}} \middle| \Psi_{\mathrm{KS}} \right\rangle, \tag{2.29}$$

where Ψ_0^{λ} is the ground state eigenfunction of $\hat{H}^{\lambda}$ and Ψ_{KS} is the Kohn–Sham determinant.

Using the adiabatic-connection fluctuation-dissipation theorem (ACFDT),[22–26] the expression for the correlation energy from Eq. 2.29 can be reformulated as:

$$E_{\mathrm{corr}} = -\frac{1}{2\pi} \int\limits_0^1 d\lambda \int\limits_0^{\infty} d\omega \iint d\boldsymbol{\tau}_1 d\boldsymbol{\tau}_2 \frac{\chi_{\lambda}(\boldsymbol{\tau}_1, \boldsymbol{\tau}_2, i\omega) - \chi_{\mathrm{KS}}(\boldsymbol{\tau}_1, \boldsymbol{\tau}_2, i\omega)}{r_{12}}, \tag{2.30}$$

where χ_{KS} is the Kohn–Sham frequency-dependent response function and χ_{λ} is the corresponding interacting response function.[114] These response functions describe the relationship between an external, frequency-dependent perturbation acting at position $\mathbf{r}_2$ and with frequency ω and the induced change of the density at position $\mathbf{r}_1$. For the Kohn–Sham response function, an exact expression is known:[115]

$$\chi_{\mathrm{KS}}(\boldsymbol{\tau}_1, \boldsymbol{\tau}_2, \omega) = \sum_{ia} \frac{2\,(\epsilon_a - \epsilon_i)\,\varphi_i(\boldsymbol{\tau}_1)\varphi_a(\boldsymbol{\tau}_1)\varphi_i(\boldsymbol{\tau}_2)\varphi_a(\boldsymbol{\tau}_2)}{\omega^2 - (\epsilon_a - \epsilon_i)^2}. \tag{2.31}$$

The interacting response can, in principle, be computed exactly from a Dyson-type equation:[116]

$$\chi(\boldsymbol{\tau}_1, \boldsymbol{\tau}_2, \omega) = \chi_{\mathrm{KS}}(\boldsymbol{\tau}_1, \boldsymbol{\tau}_2, \omega) + \iint d\boldsymbol{\tau}_3 d\boldsymbol{\tau}_4\, \chi_{\mathrm{KS}}(\boldsymbol{\tau}_1, \boldsymbol{\tau}_3, \omega) \left[\frac{1}{r_{34}} + f_{\mathrm{xc}}(\boldsymbol{\tau}_3, \boldsymbol{\tau}_4, \omega)\right] \chi(\boldsymbol{\tau}_4, \boldsymbol{\tau}_2, \omega), \tag{2.32}$$

where $f_{\mathrm{xc}}(\boldsymbol{\tau}_3, \boldsymbol{\tau}_4, \omega)$ is the exchange-correlation kernel, for which no exact expression is known. In its time-domain representation, f_{xc} describes the change of the exchange-correlation potential at time t_3 and position $\boldsymbol{\tau}_3$ that is caused by a change of the electron density at time t_4 and position $\boldsymbol{\tau}_4$:

$$f_{\mathrm{xc}}(\boldsymbol{\tau}_3, \boldsymbol{\tau}_4, t_3, t_4) = \frac{\partial V_{\mathrm{xc}}(\boldsymbol{\tau}_3, t_3)}{\partial \rho(\boldsymbol{\tau}_4, t_4)}. \tag{2.33}$$

$f_{\mathrm{xc}}(\boldsymbol{\tau}_3, \boldsymbol{\tau}_4, \omega)$ is obtained from the function in Eq. 2.33 by a Fourier transform.

When working in a basis of MOs, Eq. 2.32 can also be written in matrix form:

$$\boldsymbol{\Pi}(\omega) = \boldsymbol{\Pi}^{\mathrm{KS}}(\omega) + \boldsymbol{\Pi}^{\mathrm{KS}}(\omega)\left[\mathbf{V} + \mathbf{F}^{\mathrm{xc}}(\omega)\right]\boldsymbol{\Pi}(\omega), \tag{2.34}$$

where all involved matrices have dimensions $(n_{\mathrm{occ}} n_{\mathrm{virt}}) \times (n_{\mathrm{occ}} n_{\mathrm{virt}})$

$$\Pi^{\mathrm{KS}}_{ia,jb}(\omega) = \frac{2(\epsilon_a - \epsilon_i)\delta_{ij}\delta_{ab}}{\omega^2 - (\epsilon_a - \epsilon_i)^2}, \tag{2.35}$$

$$F^{\mathrm{xc}}_{ia,jb}(\omega) = \iint d\boldsymbol{\tau}_1 d\boldsymbol{\tau}_2\, \varphi_i(\boldsymbol{\tau}_1)\varphi_a(\boldsymbol{\tau}_1) f_{\mathrm{xc}}(\boldsymbol{\tau}_1, \boldsymbol{\tau}_2, \omega)\varphi_j(\boldsymbol{\tau}_2)\varphi_b(\boldsymbol{\tau}_2), \tag{2.36}$$

$$V_{ia,jb} = (ia|jb)\,. \tag{2.37}$$

Different approximations can be employed for the exchange-correlation kernel; the arguably simplest one is the direct random phase approximation (RPA),[27] in which the exchange-correlation kernel is neglected entirely:

$$\boldsymbol{\Pi}^{\mathrm{RPA}} = \boldsymbol{\Pi}^{\mathrm{KS}} + \boldsymbol{\Pi}^{\mathrm{KS}}\,\mathbf{V}\,\boldsymbol{\Pi}^{\mathrm{RPA}}. \tag{2.38}$$

Different formulations of RPA exist, which differ in the final working equations, like the plasmon formulation[26] or the dielectric matrix formulation.[117] In **Publication IV**, we employ the $\mathcal{O}(N^4)$ scaling RI-RPA formulation from Eshuis *et al.*:[118]

$$E^{\mathrm{RPA}}_{\mathrm{corr}} = \frac{1}{2\pi}\int\limits_0^\infty d\omega\, \mathrm{Tr}\{\ln(\mathbf{I} - \mathbf{X}^{\mathrm{KS}}(i\omega)\mathbf{V}) + \mathbf{X}^{\mathrm{KS}}(i\omega)\mathbf{V}\}, \tag{2.39}$$

$$X^{\mathrm{KS}}_{PQ}(i\omega) = \sum_{iajb} (P|ia)\,\Pi^{\mathrm{KS}}_{ia,jb}(i\omega)\,(jb|Q)\,, \tag{2.40}$$

$$V_{PQ} = (P|Q)^{-1}\,. \tag{2.41}$$

Eq. 2.39 holds only for real MOs; a more general expression that is also valid for complex MOs is derived in **Publication IV**.

RPA systematically improves energies obtained with DFT and gives particularly accurate results for dispersion-dominated systems and molecules with a small HOMO-LUMO gap.[119–121] It has also been shown that RPA provides accurate geometries and first-order molecular properties.[122] In **Publication IV**, we present the first method for computing NMR shieldings with RPA using a numerical derivative approach. With HF orbitals, RPA NMR shieldings are shown to be more accurate than MP2 shieldings and almost as accurate as CCSD shieldings.

The accuracy of direct RPA can be further increased by including exchange contributions. To this end, several approaches were suggested, like the second-order screened exchange (SOSEX)[117,123,124] or the approximate exchange kernel (AXK).[125] Very recently, Trushin *et al.*[57] introduced a novel class of functionals called σ-functionals, which significantly increase the accuracy of direct RPA while having virtually the same computational cost. These novel functionals are also included in our benchmark study on NMR shieldings from **Publication IV**.

2.3 Reduced-Scaling Quantum Chemistry Approaches

The scaling is one of the most important properties of a quantum chemical algorithm. The computational effort W can in general be approximately described by

$$W(M) = \alpha \times M^x, \tag{2.42}$$

where α is a constant called prefactor, M is some measure for the system size like the number of basis functions and the exponent x characterizes different quantum chemical methods. Low scaling, ideally linear scaling ($x = 1$), is desirable especially for studying larger molecules, as beyond a certain system size (called the crossover point) the method with lower scaling is always faster. The location of the crossover point depends both on the prefactors and the scaling of the compared methods.

In some cases, tensor decompositions like tensor hypercontraction can reduce the scaling of a quantum chemical method. Such decompositions will be discussed in Sec. 2.4.2. In the current section, approaches are discussed that reduce the scaling for large molecules by exploiting sparsity in the integrals or other intermediates.

2.3.1 Sparse Linear Algebra

Matrices and matrix operations are ubiquitous in quantum chemical methods and their implementations. Some of these matrices are sparse, which means that a large number of elements has a very small or zero value. Prime examples of sparse matrices are the AO density matrix in HF or DFT or the pseudo-density matrices that occur in AO-MP2 and related methods. Multiplications and other operations involving sparse matrices can be accelerated significantly by avoiding to include negligible elements during the computation. While this is guaranteed to reduce the required number of floating point operations, it can in practical implementations lead to a significant computational overhead compared to highly optimized routines for densely populated matrices. Therefore it is important to use suitable data structures and algorithms that allow to exploit the available sparsity in an *efficient* manner.

The mentioned density and pseudo-density matrices as well as their Cholesky factors usually display a diagonally-dominant structure if the atoms and basis functions are ordered appropriately (e.g., with the reverse Cuthill–McKee algorithm[126]). The magnitude of the elements decreases with increasing distance from the main diagonal. This sparsity pattern can be exploited using block-sparse linear algebra, which is applied in **Publications VI** and **III**. To this end, the matrix is divided into submatrices or blocks. In multiplications, blocks are excluded if their norm is smaller than a chosen threshold, which speeds up the multiplication. At the same time, highly optimized matrix multiplication routines can be employed for multiplying blocks, which reduces the computational overhead compared to dense matrix multiplications. Furthermore, no memory needs to be allocated for insignificant blocks, which lowers the memory demands of the method.

In addition, a special data format for three-center integrals called natural blocking[47,127] is used in **Publications I-III**. As an example, natural blocking shall be discussed for the three-center integrals $\left(P|\frac{\mathrm{erfc}(\omega r_{12})}{r_{12}}|\mu\nu\right)$. The elements in the corresponding third order integral tensor can be arranged such that there is a $N \times N$ matrix for every auxiliary index P, where N is the number of basis functions. If natural blocking is applied, all columns and all rows in these matrices that do not contain any significant element above a chosen threshold are deleted. This can lead to significant reduction of the size of these matrices. In particular, if all three indices of a third order tensor only couple over short distances (as it is the case for the three-center integrals with an attenuated Coulomb metric), then asymptotically only $\mathcal{O}(1)$ columns and rows are expected to be significant. Due to the reduced dimensions of the matrices, multiplications with matrices or contractions with other integral tensors can be accelerated and in many cases, also the asymptotic scaling can be lowered as demonstrated in **Publications I-III**. Furthermore, also the memory and disk space requirements for the three-center integrals can be lowered with natural blocking.

2.3.2 Integral Screening

In integral screening approaches, upper bounds or estimates for integrals occurring in quantum chemical methods are used for avoiding the computation of negligible integrals or negligible contributions involving certain integrals. One example for integral screening with a rigorous upper bound are Schwarz estimates[50] for the ERIs. These estimates are based on the following inequality:

$$|\,(\mu\nu|\lambda\sigma)\,| \leq Q_{\mu\nu}Q_{\lambda\sigma}, \tag{2.43}$$

which is a special case of the more general Cauchy–Schwarz inequality.[128] The Schwarz factors Q are defined as follows:

$$Q_{\mu\nu} = \sqrt{|\,(\mu\nu|\mu\nu)\,|}. \tag{2.44}$$

Formally, there are $\mathcal{O}(N^4)$ ERIs. However, ERIs show an exponential decay with increasing distance between the basis functions χ_μ and χ_ν in the bra and between χ_λ and χ_σ in the ket. For large molecules, many of the ERIs thus have values close to zero and are negligible. Due to the exponential coupling, only $\mathcal{O}(N)$ $\mu\nu$ pairs and $\lambda\sigma$ pairs are significant. Schwarz screening can capture this decay behavior and allows to compute only the $\mathcal{O}(N^2)$ significant ERIs.

In **Publication I**, Schwarz estimates are applied to the exchange contributions to the ω-RI-CDD-MP2 energy:

$$\left|\left(\underline{i}\overline{a}|\underline{j}\overline{b}\right)\left(\underline{i}\overline{b}|\underline{j}\overline{a}\right)\right| \leq Q_{\underline{i}\overline{a}}Q_{\underline{j}\overline{b}}Q_{\underline{i}\overline{b}}Q_{\underline{j}\overline{a}}. \tag{2.45}$$

In the ERI $\left(\underline{i}\overline{a}|\underline{j}\overline{b}\right)$, the Cholesky pseudo-MOs $\phi_{\underline{i}}$ and $\phi_{\overline{a}}$ in the bra are coupled exponentially as well as $\phi_{\underline{j}}$ and $\phi_{\overline{b}}$ in the ket. In the integral $\left(\underline{i}\overline{b}|\underline{j}\overline{a}\right)$, $\phi_{\underline{i}}$ is exponentially coupled to $\phi_{\overline{b}}$ and $\phi_{\underline{j}}$ is coupled exponentially to $\phi_{\overline{a}}$. In consequence, all four orbitals $\phi_{\underline{i}}$, $\phi_{\overline{a}}$, $\phi_{\underline{j}}$, and $\phi_{\overline{b}}$ need to be spatially close to each other; otherwise, the integral product in Eq. 2.45

vanishes. The Schwarz estimates from Eq. 2.45 are able to capture this decay behavior and reduce the number of computed contributions to linear. In a similar way, Schwarz estimates have recently also been applied to the exchange contributions in RPA-X methods by Beuerle *et al.*[129]

In some integral screening approaches, also other quantities like the HF density matrix are taken into account. One example is the LinK screening,[130,131] which enables a linear-scaling computation of the elements of the HF exchange matrix:

$$K_{\mu\nu} = \sum_{\lambda\sigma} (\mu\sigma|\lambda\nu)\, P_{\lambda\sigma}, \tag{2.46}$$

$$|\,(\mu\sigma|\lambda\nu)\, P_{\lambda\sigma}| \leq Q_{\mu\sigma} Q_{\lambda\nu} |P_{\lambda\sigma}|. \tag{2.47}$$

For systems with a non-vanishing HOMO-LUMO gap, the density matrix elements $P_{\lambda\sigma}$ decay exponentially with distance between χ_λ and χ_σ.[132–134] Because of the exponential coupling of the basis functions in the bra and in the ket and the additional coupling of bra and ket through the density matrix the number of significant contributions to $\mathbf{K}$ is reduced to asymptotically $\mathcal{O}(N)$. A similar screening approach, which takes into account both the HF density and the spin density matrix, was applied to the spin-orbit mean field (SOMF) matrix[135] in **Publication VI**. This significantly reduces the number of two-electron spin-orbit integrals that need to be computed, particularly for systems with a localized spin density.

2.4 Tensor Decompositions

Matrices and higher-order tensors play an important role in quantum chemical methods. Of particular significance is the fourth-order tensor of electron repulsion integrals. Tensor decomposition approaches can be used to reduce the memory requirements for storing the tensor elements and for accelerating tensor contractions. Also in this thesis, several different tensor decomposition techniques are applied, which will be described in the following.

2.4.1 Resolution-of-the-Identity

An ERI $(\mu\nu|\lambda\sigma)$ describes the electrostatic interaction between two charge densities $\rho_{\mu\nu}(\mathbf{r}_1)$ and $\rho_{\lambda\sigma}(\mathbf{r}_2)$, which are both given by a product of basis functions:

$$\rho_{\mu\nu}(\mathbf{r}_1) = \chi^*_\mu(\mathbf{r}_1)\chi_\nu(\mathbf{r}_1) \qquad \rho_{\lambda\sigma}(\mathbf{r}_2) = \chi^*_\lambda(\mathbf{r}_2)\chi_\sigma(\mathbf{r}_2). \tag{2.48}$$

When using the resolution-of-the-identity approximation,[37,38] also called density fitting,[45] at least one of these two charge densities is approximated using a linear combination of auxiliary functions χ_P:

$$\rho_{\lambda\sigma} = \chi^*_\lambda(\mathbf{r}_2)\chi_\sigma(\mathbf{r}_2) \approx \sum_P \chi_P(\mathbf{r}_2) C^P_{\lambda\sigma} = \tilde{\rho}_{\lambda\sigma}. \tag{2.49}$$

Different methods can be used to obtain the density fitting coefficients $C^P_{\lambda\sigma}$. Minimizing the error in the charge density $\Delta_{\lambda\sigma} \equiv \tilde{\rho}_{\lambda\sigma} - \rho_{\lambda\sigma}$ in a least-squares sense results in the so-called overlap metric:[136]

$$C^P_{\lambda\sigma} = \sum_Q (\lambda\sigma Q)\,(QP)^{-1}. \tag{2.50}$$

This formula contains two types of overlap integrals:

$$(\lambda\sigma Q) = \int d\mathbf{r}\,\chi^*_\lambda(\mathbf{r})\chi_\sigma(\mathbf{r})\chi_Q(\mathbf{r}), \tag{2.51}$$

$$(QP) = \int d\mathbf{r}\,\chi_Q(\mathbf{r})\chi_P(\mathbf{r}). \tag{2.52}$$

If both of the charge densities in the bra and in the ket of an ERI $(\mu\nu|\lambda\sigma)$ are fitted using the overlap metric, the following approximation for the ERI is obtained:

$$(\mu\nu|\lambda\sigma) \approx \sum_{PQRS} (\mu\nu P)\,(PQ)^{-1}\,(Q|R)\,(RS)^{-1}\,(S\lambda\sigma)\,. \tag{2.53}$$

Alternatively, one might employ the so-called Coulomb metric,[136] which minimizes the error in the Coulomb potential of the fitted charge densities. With the Coulomb metric, a four-center ERI can be decomposed into three- and two-center ERIs:

$$(\mu\nu|\lambda\sigma) \approx \sum_{PQRS} (\mu\nu|P)\,(P|Q)^{-1}\,(Q|R)\,(R|S)^{-1}\,(S|\lambda\sigma) = \sum_{PQ} (\mu\nu|P)\,(P|Q)^{-1}\,(Q|\lambda\sigma)\,. \tag{2.54}$$

With a given auxiliary basis, the (four-center) ERIs are usually significantly more accurately approximated with the Coulomb metric than with the overlap metric.[136] Therefore, the overlap metric is used rarely in practice, even though it leads to more sparsity in the three-center integrals because of an exponential decay with increasing bra-ket separation (while the Coulomb metric leads to a very slow $\frac{1}{r_{12}}$-decay).

The attenuated Coulomb metric[46–48] is another important RI metric. The corresponding decomposition of an ERI reads as follows:

$$(\mu\nu|\lambda\sigma) \approx \sum_{PQRS} (\mu\nu\mathrel{\text{¦}}P)(P\mathrel{\text{¦}}Q)^{-1}(Q|R)(R\mathrel{\text{¦}}S)^{-1}(S\mathrel{\text{¦}}\lambda\sigma). \tag{2.55}$$

The dashed vertical lines in this equation represent an erfc-attenuated electron-electron interaction operator given by $\frac{\mathrm{erfc}(\omega r_{12})}{r_{12}}$, which depends on the attenuation parameter ω. If ω is set to zero, the standard Coulomb metric is obtained as $\mathrm{erfc}(0) = 1$. In the limit of $\omega \to \infty$, the overlap metric is obtained. Thus, the attenuated Coulomb metric allows to interpolate between Coulomb and overlap metric. It has been shown that a suitable choice of ω allows to combine the advantages of these two metrics.[49] With $\omega \approx 0.1$ the accuracy is comparable to the Coulomb metric; additionally, a high degree of sparsity in the three-center integrals is obtained. This sparsity can be exploited in order to accelerate calculations on large molecules and to reduce the scaling. Because of these favorable properties, the attenuated Coulomb metric is used for increasing the efficiency of MP2 energy and MP2 NMR shielding calculations in **Publications I** and **III** and for speeding up the decomposition of the ERIs in tensor hypercontraction format in **Publication II**.

For many methods like, e.g., MP2, the RI approximation does not lower the formal scaling, but often significantly reduces the prefactor.[37] Additionally, the RI approximation often reduces the memory demands. Irrespective of the employed metric, RI allows to replace the fourth-order ERI tensor by third- and second-order integral tensors. Storing the elements of these lower order tensors usually requires significantly less memory, which is especially beneficial for electron correlation methods. A local metric such as the attenuated Coulomb can enable a further reduction in the amount of needed memory, if the sparsity in the three-center integrals is exploited. For this purpose, sparse data formats such as natural blocking and block-sparse matrices are employed in **Publications I** and **III** in order to lower the memory and disk space requirements for storing three-center integrals.

2.4.2 Tensor Hypercontraction

Tensor hypercontraction (THC) is a rather novel tensor decomposition technique and was first introduced by Martínez and coworkers in 2012.[51] Later, applications to various electronic structure methods were published, including MP2,[137,138] CCSD,[139,140] and second-order approximate coupled cluster singles and doubles (CC2).[141,142] More recently, THC has also been applied in the context of second-order complete active space perturbation theory (CASPT2).[143] For many of these methods, THC allows to reduce the formal scaling. For both MP2 and CCSD, e.g., the scaling can be reduced to $\mathcal{O}(N^4)$,[51,139] whereas the scaling with undecomposed ERIs or RI-approximated ERIs amounts to $\mathcal{O}(N^5)$ and $\mathcal{O}(N^6)$, respectively.

THC can be applied to ERIs both in the AO or MO basis. Here, the THC decomposition format is shown for an ERI in the MO basis:

$$(ia|jb) \approx \sum_{PQ} X_i^P X_a^P Z^{PQ} X_j^Q X_b^Q. \tag{2.56}$$

In the first article on THC, Hohenstein *et al.*[51] used an iterative PARAFAC decomposition[144] for obtaining the factors $\mathbf{X}$ and $\mathbf{Z}$. Later, the grid-based least-squares THC approach (LS-THC)[52] was introduced, which provides a good compromise between accuracy and computational efficiency of the THC factorization. In LS-THC, the $\mathbf{X}$-matrices contain the values of AOs or MOs at the spatial positions of a grid:

$$X_i^P = \phi_i(\mathbf{r}_P)\sqrt[4]{w_P}, \tag{2.57}$$

where w_P is the weight of the grid point at position $\mathbf{r}_P$. It has been shown that a rather small number of grid points per atom (less than 100) can be sufficient for an accurate THC decomposition of the ERIs.[145] A particularly good ratio between accuracy and the number of grid points per atom is obtained with grids from Ref. 145, which were optimized for THC and specific basis sets. Using the **X**-matrices from Eq. 2.57, the corresponding **Z**-matrix that minimizes the error in the ERI tensor in a least-squares sense

$$\frac{\partial}{\partial Z^{PQ}}\left[\sum_{iajb}\left((ia|jb)-\sum_{PQ}X_i^PX_a^PZ^{PQ}X_j^QX_b^Q\right)^2\right]\overset{!}{=}0\ \ \forall P,Q \tag{2.58}$$

can be determined by evaluating the following analytical expression in LS-THC:[52]

$$\mathbf{Z}=\mathbf{S}^{-1}\mathbf{E}\mathbf{S}^{-1}, \tag{2.59}$$

where

$$S^{PQ}=\sum_{ia}X_i^PX_a^PX_i^QX_a^Q, \tag{2.60}$$

$$E^{PQ}=\sum_{iajb}X_i^PX_a^P\,(ia|jb)\,X_j^QX_b^Q. \tag{2.61}$$

In **Publication II**, local Cholesky-MOs are used instead of delocalized, canonical MOs in Eqs. 2.60 and 2.61. Additionally, the ERIs in Eq. 2.61 are approximated using RI with an attenuated Coulomb metric. Together with sparse linear algebra, this allows for a highly efficient computation of the **E** and **S** intermediates.

In Eq. 2.58, **Z** is fitted to ERIs in the MO basis of the form $(ia|jb)$; this is denoted as MO-THC. Alternatively, **Z** can also fitted to the ERIs in the AO basis (AO-THC); for MP2 energies this is significantly less accurate, because the number of ERIs in the AO basis is much larger than the number of $(ia|jb)$-type integrals.

The THC approximation has shown great potential in the development of efficient quantum chemistry methods due to several properties. THC can reduce the formal scaling of many methods. As an example, consider the computation of the Coulomb contribution $E^{\mathrm{J}}_{\mathrm{MP2}}$ in RI-CDD-MP2 from Maurer *et al.*:[146]

$$E^{\mathrm{J}}_{\mathrm{MP2}}\equiv-2\sum_{\tau}\sum_{\underline{i}\underline{j}\bar{a}\bar{b}}\left(\underline{i}\bar{a}|\underline{j}\bar{b}\right)^{(\tau)}\left(\underline{i}\bar{a}|\underline{j}\bar{b}\right)^{(\tau)}, \tag{2.62}$$

where the sum over τ runs over all Laplace quadrature points. With RI, evaluating Eq. 2.62 scales as $\mathcal{O}(N^4)$. As shown in **Publication II**, the scaling can be lowered to $\mathcal{O}(N^3)$ with THC:

$$E^{\mathrm{J}}_{\mathrm{MP2}}=-2\sum_{\tau}\mathrm{Tr}\left[\left(\mathbf{C}^{(\tau)}\mathbf{Z}\right)\left(\mathbf{C}^{(\tau)}\mathbf{Z}\right)\right], \tag{2.63}$$

where

$$C^{PQ,(\tau)}=\Big(\sum_{\underline{i}}X_{\underline{i}}^PX_{\underline{i}}^Q\Big)\Big(\sum_{\bar{a}}X_{\bar{a}}^PX_{\bar{a}}^Q\Big). \tag{2.64}$$

THC also allows to lower the formal scaling of the RI-CDD-MP2 exchange contribution from $\mathcal{O}(N^5)$ to $\mathcal{O}(N^4)$.

In **Publication II**, it is shown how the effective scaling of THC-MP2 can be reduced further by exploiting sparsity. One might wonder whether reduced-scaling methods derived from THC-MP2 have any advantages compared to RI-MP2 based methods with the same asymptotic scaling. One needs to consider, however, that reduced-scaling methods achieve lowered scaling only for systems beyond a certain size by exploiting sparsity. For small molecules, the scaling behavior always follows the formal scaling. The molecular size range, in which a transition from the formal to the asymptotic scaling can be observed, strongly depends on the system. In systems, in which this transition occurs rather late, the THC-based methods will have a natural advantage, because they benefit from the lower formal scaling. Molecular systems with a late onset of reduced scaling include dense three-dimensional molecules (as opposed to linear, extended molecules), molecules with delocalized electronic structure, and systems with a large basis set containing diffuse basis functions. One can thus expect the THC-based methods to perform particularly well for these kind of systems.

THC can also greatly reduce the memory requirements compared to methods based on RI. The formal scaling of the required memory is quadratic for THC and cubic for RI methods. The crossover occurs rather early and also for molecules that are tractable with current methods, significant memory savings are possible. Consider, e.g., a system with 1000 basis functions. Typically, 3000-4000 auxiliary basis functions and roughly 10000 THC grid points (assuming optimized grids as from Kokkila Schumacher *et al.*[145]) are needed for such a system. Not considering sparsity or symmetry, the three-center integrals in the RI formulation would require 22.4 - 29.8 GB of memory ($1000^2 \times 3000 \times 8\,\text{bytes} = 22.4\,\text{GB}$, as 8 bytes of memory are needed for a double-precision float). In contrast, only 0.75 GB ($10000^2 \times 8$ bytes) memory are required for storing the matrix $\mathbf{Z}$ from Eq. 2.59; the $\mathbf{X}$ matrices require at least an order of magnitude less memory. As calculations of molecular properties using RI are often more memory-intensive than energy calculations due to the necessity of storing several different types of transformed integrals (see, e.g., **Publication III**), THC might be particularly useful in property calculations.

2.4.3 Cholesky Decomposition

A Cholesky decomposition (CD)[39] allows to decompose a symmetric positive definite matrix $\mathbf{A}$ into the product of a lower triangular matrix $\mathbf{L}$ and its transpose:

$$\mathbf{A} = \mathbf{L}\mathbf{L}^{\top}. \tag{2.65}$$

Pivoted CD, which is a variant of CD, allows to decompose also positive semi-definite matrices.[40]

CD can be applied to different quantities occurring in quantum chemistry. CD of the fourth-order ERI tensor decomposes it into two third-order tensors – similar to the RI approximation.[42,147,148] CD has also been applied to the MP2 orbital energy denominator from Eq. 2.20.[149] In this thesis, CD is used for decomposing the occupied and virtual pseudo-density matrices, which occur in AO-MP2 theory:

$$\underline{P}_{\mu\nu} = \sum_{\underline{i}} \underline{L}_{\mu\underline{i}} \underline{L}_{\nu\underline{i}}, \tag{2.66}$$

$$\overline{P}_{\mu\nu} = \sum_{\overline{a}} \overline{L}_{\mu\overline{a}} \overline{L}_{\nu\overline{a}}. \tag{2.67}$$

Such a decomposition of pseudo-densities has previously been applied to MP2[41,146] as well as to RPA.[49,150] The Cholesky factors $\underline{L}_{\mu\underline{i}}$ and $\overline{L}_{\mu\overline{a}}$ can be interpreted as the expansion coefficients of Cholesky pseudo-MOs $\phi_{\underline{i}}$ and $\phi_{\overline{a}}$, respectively. The number of occupied (virtual) Cholesky pseudo-MOs is smaller than or equal to the number of occupied (virtual) orbitals due to the rank deficiency of the pseudo-density matrices. This can be used for reducing the elements of the integral tensors in electron correlation methods like MP2 and RPA, as especially the number of occupied MOs is much smaller than the number of basis-functions for high-quality basis sets. It also improves the scaling with respect to the size of the basis set for fixed molecule size, because the number of occupied orbitals is constant in this case.

It also has been shown that the orbitals obtained from a CD are local as opposed to the delocalized canonical MOs.[151] CD can thus be used as an alternative to iterative orbital localization approaches,[94,152–155] which generally are computationally more demanding and can suffer from convergence problems.

In order to benefit from the rank reduction and locality of the Cholesky pseudo-MOs, the decompositions from Eqs. 2.66 and 2.67 are applied to MP2 energies in **Publications I** and **II** and to MP2-NMR shieldings in **Publication III**. In **Publication I**, also the HF density matrix is subjected to a CD – similar to Ref. 150. The obtained Cholesky factors are used for transforming one index of the AO three-center integrals. For each Laplace point, the Cholesky MO index is transformed again to the basis of occupied Cholesky pseudo-MOs. This two-step transformation enables both computational savings and reduces the disk space requirements by reducing the size of the integral tensors.

2.5 Computation of Magnetic Resonance Parameters

Spectroscopic techniques are widely used in modern experimental chemistry and can reveal a wealth of information about molecular geometry and electronic structure. A quantum chemical calculation of many spectroscopic properties is possible via derivatives of the energy. In this thesis, the focus lies on NMR shieldings and electronic g-tensors, which both can be computed using a mixed second derivative of the energy. Both properties require the treatment of magnetic fields. In general, magnetic fields can be accounted for by replacing the kinetic momentum $\mathbf{p}$ with $\boldsymbol{\pi} = \mathbf{p} - q\mathbf{A}$ according to the principle of minimal coupling.[156] q is the charge (-1 for electrons) and $\mathbf{A}$ the vector potential, which is related to the magnetic field $\mathbf{B}$ via $\mathbf{B} = \nabla \times \mathbf{A}$. The electronic Hamiltonian from Eq. 2.4 then changes to

$$\hat{H}_{\mathrm{e}} = \frac{1}{2} \sum_i \left(\hat{\mathbf{p}}_i + \mathbf{A}(\mathbf{r}_i)\right)^2 + \hat{V}_{\mathrm{eN}} + \hat{V}_{\mathrm{ee}} + \hat{V}_{\mathrm{NN}}. \tag{2.68}$$

The vector potential contains a contribution $\mathbf{A}_{\mathrm{ext}}$ from the external magnetic field. For NMR shieldings, also a contribution $\mathbf{A}_{\mathbf{m}}$ from the magnetic moment $\mathbf{m}^A$ of the nuclei needs to be included:

$$\mathbf{A} = \underbrace{\frac{1}{2}\mathbf{B} \times (\mathbf{r} - \mathbf{R}_0)}_{\mathbf{A}_{\mathrm{ext}}} + \underbrace{\alpha^2 \sum_A \frac{\mathbf{m}^A \times (\mathbf{r} - \mathbf{R}_A)}{|\mathbf{r} - \mathbf{R}_A|^3}}_{\mathbf{A}_{\mathbf{m}}}, \tag{2.69}$$

where $\mathbf{R}_0$ denotes the gauge-origin. Similarly, the contribution of the magnetic moment of the electron to the vector potential needs to be accounted for when computing g-tensors.

2.5.1 Gauge-Origin Dependence

Gauge-origin dependence is an important phenomenon that can occur in the computation of magnetic properties.[58] In a gauge-origin dependent calculation, the computed values show an unphysical dependence on the choice of the coordinate system. This is an artifact caused by the use of an incomplete basis set.[58] In principle, it could be eliminated by using a complete basis set, but this is not feasible in practice. Alternatively, gauge-origin independence can also be ensured with distributed gauge-origin methods like gauge-including atomic orbitals (GIAOs),[157–160] or, today less commonly, the individual gauge for localized orbitals (IGLO)[161] or the localized orbital/local origin gauge (LORG) approach.[162] GIAOs are the most commonly used approach and standard for computing NMR shieldings. They are constructed by multiplying AOs with a magnetic field dependent phase factor:

$$\chi_\mu(\mathbf{r}, \mathbf{B}) = \chi_\mu(\mathbf{r}) \exp\left[-\frac{i}{2}\mathbf{B} \times [(\mathbf{R}_\mu - \mathbf{R}_0) \cdot \mathbf{r}]\right]. \tag{2.70}$$

For electronic g-tensors, most implementations described in the literature to date employ a common gauge-origin.[59–76] The most commonly used gauge-origin is the electronic charge centroid (ECC),[163] whose position can be computed according to

$$\mathbf{r}_{\mathrm{ECC}} = \frac{1}{N_{\mathrm{el}}} \sum_{\mu\nu} P_{\mu\nu} \langle \mu | \mathbf{r} | \nu \rangle . \tag{2.71}$$

The ECC can be thought of as the center of mass of the electron cloud. In **Publication V**, we provided a benchmark study for the gauge-origin dependence in g-tensor and showed that especially for large molecules, common gauge-origin approaches can lead to substantial errors. The findings suggest that GIAOs or related approaches should be used for such molecules in order to obtain reliable, gauge-origin independent g-tensors. However, if and only if the spin density is well-localized in a small region of space, a common gauge-origin, which is located in this region, can be sufficient. Therefore, as an alternative to ECC that takes the spin density distribution into account, we proposed the spin density center (SDC) given by

$$\mathbf{r}_{\mathrm{SDC}} = \frac{1}{\mathrm{Tr}\left(|\mathbf{P}^{\alpha-\beta}|\mathbf{S}\right)} \sum_{\mu\nu} |P^{\alpha-\beta}_{\mu\nu}| \langle \mu | \mathbf{r} | \nu \rangle . \tag{2.72}$$

Here, $\mathbf{P}^{\alpha-\beta}$ is the spin density matrix, which is obtained from the difference of α and β densities. By construction, the SDC lies within the region with significant spin density; numerical data in **Publication V** demonstrates that the obtained g-tensors are significantly more accurate than with the ECC.

2.5.2 NMR Shielding Tensors

Nuclei possess an intrinsic magnetic moment and can occupy energetically different orientations in a magnetic field. The splitting of these energy levels can be probed experimentally using NMR spectroscopy. In a molecule, the effective magnetic field interacting with the nuclear magnetic moment is usually different from the externally applied field due to secondary magnetic fields originating from induced currents. The shielding of a nucleus is thus a sensitive probe of the chemical environment and electronic structure of the molecule. NMR spectra contain a wealth of information about structure and geometry of a molecule.[164] However, especially for larger molecules, it can be difficult to unambiguously assign a spectrum to a particular structure. In such situations, a comparison between experimental and theoretically simulated spectra can be very helpful.

The NMR shielding tensor $\boldsymbol{\sigma}^A$ of a nucleus A can be computed quantum-mechanically by differentiating the energy with respect to an external magnetic field $\mathbf{B}$ and the nuclear magnetic moment $\mathbf{m}^A$:

$$\sigma^A_{rs} = \left.\frac{\partial E}{\partial m^A_r \partial B_s}\right|_{\mathbf{m}^A=0,\mathbf{B}=0}. \tag{2.73}$$

In order to treat the perturbation caused by the nuclear magnetic moment, both paramagnetic and diamagnetic spin-orbit coupling operators need to be included in the Hamiltonian.[165] For the magnetic field perturbation, the interaction of angular momentum of the electronic motion and the magnetic field needs to be considered.[165] Furthermore, as for other magnetic properties, the problem of gauge-origin dependence (see Sec. 2.5.1) needs to be addressed. Simple approaches with a common gauge-origin give unreliable NMR shieldings. In order to ensure gauge-origin independence distributed gauge-origin approaches like GIAOs are usually employed in NMR shielding calculations.

Many analytical derivative implementations for NMR shieldings also employ a Z-vector approach[54,55] in order to minimize the number of CPSCF equations that need to be solved. In general, a Z-vector approach can be used, when perturbed density matrices $\mathbf{P}^\xi$ ($\equiv \frac{\partial}{\partial \xi}\mathbf{P}$) for some general perturbation ξ appear in an expression of the following form:

$$\mathrm{Tr}[\mathbf{X}\mathbf{P}^\xi] = \mathrm{Tr}[\mathbf{X}(\mathbf{A}^{-1}\mathbf{b}^\xi)] = \mathrm{Tr}[\mathbf{Z}\mathbf{b}^\xi], \tag{2.74}$$

where $\mathbf{A}$ is the HF Hessian and $\mathbf{b}^\xi$ is the right-hand side of the CPSCF equation for $\mathbf{P}^\xi$. $\mathbf{Z} = \mathbf{A}^{-1}\mathbf{X}$ is the so-called Z-vector. If multiple perturbations ξ need to be considered, the Z-vector approach can lead to large computational savings, because $\mathbf{b}^\xi$ can typically be calculated with low effort and only a single CPSCF equation needs to be solved for $\mathbf{Z}$.

Over the last decades, many quantum chemical methods for computing NMR shieldings were developed. This includes methods with HF,[166–169] DFT,[170–172] MP2,[173,174] and coupled cluster theory.[175–177] Also several reduced-scaling methods at the SCF level of theory were introduced, which are applicable to large molecules with 1000 atoms or more.[178–180] As MP2 gives highly accurate shieldings and outperforms HF and DFT,[171,181–183] there is great demand for efficient methods based on MP2. Several local correlation approaches for MP2 shieldings were proposed.[184–186] An asymptotically linear-scaling method based on

Laplace-transformed AO-MP2 was presented by Maurer *et al.*[53] The method from Maurer *et al.*[53] employs a selected-nuclei approach,[180] which makes it particularly efficient for computing the shieldings for a few selected nuclei in the molecule. The computation of the perturbed density matrices $\mathbf{P}^{\mathbf{B}}$ and $\mathbf{P}^{\mathbf{Bm}}$ is avoided using a Z-vector approach. Only the matrices $\mathbf{P}^{\mathbf{m}}$ are calculated explicitly for the selected nuclei and their locality is exploited.

In **Publication III**, a similar method is introduced, which, however, employs an all-nuclei formulation that is more suitable for computing the entire NMR spectrum of a given molecule. This is achieved by formulating a Z-vector approach that allows to circumvent the explicit computation of $\mathbf{P}^{\mathbf{m}}$ and $\mathbf{P}^{\mathbf{Bm}}$, which would require the solution of $9 \times N_{\text{atoms}}$ and $3 \times N_{\text{atoms}}$ CPSCF equations, respectively (where N_{atoms} denotes the number of atoms). Only three CPSCF equations need to be solved for $\mathbf{P}^{\mathbf{B}}$ irrespective of the number of atoms. Furthermore, several additional approximations such as the RI approximation with an attenuated Coulomb metric and Cholesky decomposition of pseudo-densities are introduced in order to improve the computational efficiency and reduce the memory requirements.

Largely unexplored to date has been the computation of NMR shieldings with methods based on the ACFDT such as RPA. The first method for NMR shieldings with post-Kohn–Sham RPA and the related σ-functionals is presented by us in **Publication IV**. In this study, the derivative from Eq. 2.73 is evaluated numerically using a finite-difference scheme.

2.5.3 Electronic g-Tensors

The electronic g-tensor, measured using EPR spectroscopy, can be viewed as the analogue of NMR shieldings for unpaired electrons. It describes the interaction between the magnetic moment of an unpaired electron and an external magnetic field. A quantum chemical computation of g-tensors is possible by computing the following second derivative of the energy:

$$g_{rs} = \frac{2}{\alpha} \frac{\partial E}{\partial B_r \partial s_s}\bigg|_{\mathbf{B}=0,\mathbf{s}=0}, \tag{2.75}$$

where $\mathbf{s}$ denotes the electronic spin vector. In many aspects, the calculation of g-tensors is similar to the calculation of NMR shielding tensors. Both properties require the treatment of external magnetic fields and the problem of gauge-origin dependence, which for g-tensors is analyzed in detail in **Publication V**. In contrast to NMR shieldings, which involve only electron-nuclear spin-orbit coupling, also two-electron spin-orbit coupling effects need be considered for g-tensors. Several approximations have been developed for the two-electron spin-orbit coupling, like effective one-electron approaches with scaled nuclear charges[187] or the spin-orbit mean field (SOMF)[135] method. The latter describes the two-electron spin-orbit interaction with a mean field approach that resembles the description of the electron-electron interaction in HF theory and is used in **Publications V** and **VI**. The implementation that is presented in these publications is the first one to combine the SOMF operator and GIAOs and therefore enables accurate and gauge-origin independent calculations. It also enables a linear- or sublinear-scaling computation of g-tensors using integral screening techniques.

Chapter 3

Publications

3.1 Efficient Reduced-Scaling Second-Order Møller–Plesset Perturbation Theory with Cholesky-Decomposed Densities and an Attenuated Coulomb Metric

M. Glasbrenner, D. Graf, C. Ochsenfeld.
J. Chem. Theory Comput. **16**, 6856 (2020).

Abstract: We present a novel, highly efficient method for the computation of second-order Møller–Plesset perturbation theory (MP2) correlation energies, which uses the resolution of the identity (RI) approximation and local molecular orbitals obtained from a Cholesky decomposition of pseudodensity matrices (CDD), as in the RI-CDD-MP2 method developed previously in our group [Maurer, S. A.; Clin, L.; Ochsenfeld, C. *J. Chem. Phys.* **2014**, *140*, 224112]. In addition, we introduce an attenuated Coulomb metric and subsequently redesign the RI-CDD-MP2 method in order to exploit the resulting sparsity in the three-center integrals. Coulomb and exchange energy contributions are computed separately using specialized algorithms. A simple, yet effective integral screening protocol based on Schwarz estimates is used for the MP2 exchange energy. The Coulomb energy computation and the preceding transformations of the three-center integrals are accelerated using a modified version of the natural blocking approach [Jung, Y.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2006**, *8*, 2831-2840]. Effective subquadratic scaling for a wide range of molecule sizes is demonstrated in test calculations in conjunction with a low prefactor. The method is shown to enable cost-efficient MP2 calculations on large molecular systems with several thousand basis functions.

Reprinted with permission from:

M. Glasbrenner, D. Graf, C. Ochsenfeld.
“Efficient Reduced-Scaling Second-Order Møller–Plesset Perturbation Theory with Cholesky-Decomposed Densities and an Attenuated Coulomb Metric.”
J. Chem. Theory Comput. **16**, 6856 (2020).

`https://pubs.acs.org/doi/abs/10.1021/acs.jctc.0c00600`

JCTC Journal of Chemical Theory and Computation

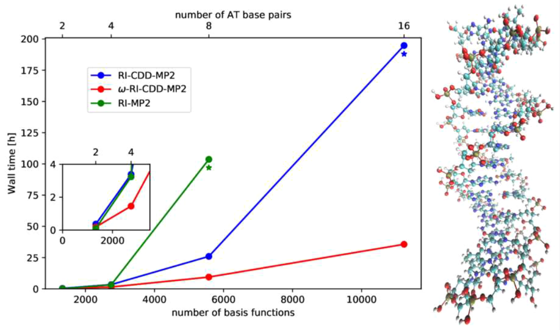
pubs.acs.org/JCTC Article

Efficient Reduced-Scaling Second-Order Møller−Plesset Perturbation Theory with Cholesky-Decomposed Densities and an Attenuated Coulomb Metric

Michael Glasbrenner, Daniel Graf, and Christian Ochsenfeld*

Cite This: *J. Chem. Theory Comput.* 2020, 16, 6856−6868 Read Online

ABSTRACT: We present a novel, highly efficient method for the computation of second-order Møller−Plesset perturbation theory (MP2) correlation energies, which uses the resolution of the identity (RI) approximation and local molecular orbitals obtained from a Cholesky decomposition of pseudodensity matrices (CDD), as in the RI-CDD-MP2 method developed previously in our group [Maurer, S. A.; Clin, L.; Ochsenfeld, C. *J. Chem. Phys.* **2014**, *140*, 224112]. In addition, we introduce an attenuated Coulomb metric and subsequently redesign the RI-CDD-MP2 method in order to exploit the resulting sparsity in the three-center integrals. Coulomb and exchange energy contributions are computed separately using specialized algorithms. A simple, yet effective integral screening protocol based on Schwarz estimates is used for the MP2 exchange energy. The Coulomb energy computation and the preceding transformations of the three-center integrals are accelerated using a modified version of the natural blocking approach [Jung, Y.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2006**, *8*, 2831−2840]. Effective subquadratic scaling for a wide range of molecule sizes is demonstrated in test calculations in conjunction with a low prefactor. The method is shown to enable cost-efficient MP2 calculations on large molecular systems with several thousand basis functions.



1. INTRODUCTION

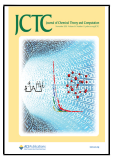
The accurate and efficient computation of electron correlation energies is one of the central challenges in *ab initio* electronic structure theory. Proper treatment of electron correlation is essential for a quantitative description of many chemical phenomena including dispersion, which is highly important in inter- and intramolecular interactions of many molecular systems. One of the most cost-efficient and commonly used wave function-based methods for computing correlation energies is second-order Møller−Plesset perturbation theory (MP2).[1] MP2 is significantly more accurate than Hartree−Fock, and because of its comparatively low N^5 scaling, it is computationally cheaper than other wave function methods such as Coupled Cluster (see, e.g., ref 2), which scale conventionally at least as N^6. Due to this good compromise between accuracy and computational cost, MP2 is one of the most widely used quantum chemistry methods.

In recent years, several variants of MP2 theory have been proposed. Empirical scaling of the same-spin and opposite-spin contributions to the MP2 energy in the spin-component-scaled MP2 (SCS-MP2)[3] has been shown to further increase the accuracy for energetics and molecular properties. The related scaled-opposite spin MP2 model (SOS-MP2)[4] leads to very efficient methods, since it allows to avoid the computationally challenging exchange contributions to the MP2 energy. Furthermore, MP2 has also been combined with density functional theory (DFT)[5] in the context of double-hybrid functionals.[6] These functionals contain an MP2-like term and have been shown to provide excellent accuracy for many applications.[6−8]

Due to the widespread use of MP2 and its related methods, the development of efficient MP2 algorithms, which allow one to treat large molecules, is an important endeavor. The unfavorable fifth-order scaling of canonical MP2 severely limits its applicability to larger systems. Over the last decades, much work has been devoted to the development of more efficient MP2 methods with reduced scaling and prefactor (see, e.g., refs 9−39).

Many efficient MP2 methods employ decompositions of the electron repulsion integral (ERI) tensor. The most commonly used decomposition is the resolution of the identity (RI) approximation,[40] which decomposes the fourth-order ERI tensor into third- and second-order tensors using a set of

Received: June 10, 2020
Published: October 19, 2020



auxiliary basis functions. Although RI-MP2 has the same scaling as canonical MP2 with four-center integrals, the prefactor and the storage requirements are significantly reduced. Apart from RI, also Cholesky decomposition of the ERI tensor[21,30] and the pseudospectral approximation[12] have been applied to MP2. More recently, tensor hypercontraction has been introduced by Martínez and co-workers,[31,41] which allows one to reduce the formal scaling of MP2 to quartic.

In order to tackle the steep scaling of MP2 algorithms for larger systems, one can also exploit the locality of electron correlation. Pulay and Saebø pioneered the field of local correlation methods and introduced localized molecular orbitals (MOs) and the concept of correlation domains into MP2 theory.[9,10] Many local MP2 methods have been proposed since then,[13,14,16−20,22,35] including the divide-expand-consolidate (DEC) ansatz,[29,36] which employs iterative optimization of the orbital spaces. Recently, Neese and co-workers have employed local pair natural orbitals for MP2 and other correlation methods using the domain-based local pair natural orbital approach.[34,42,43]

Another path to reduced-scaling MP2 algorithms starts from a Laplace transformation of the orbital energy denominator appearing in canonical MP2, which was introduced by Almlöf and Häser.[11,44,45] Later, an atomic-orbital-based reformulation of Laplace-transformed MP2 (AO-MP2) was developed.[11,15] In combination with integral screening approaches, the scaling of AO-MP2 can be reduced to be asymptotically linear.[23,26,27,32] However, the prefactor of AO-MP2 is high, especially for large basis sets, leading to a late crossover with canonical MP2.

The efficiency of the AO-MP2 method is improved in the RI-CDD-MP2 method[28,33] by introducing the RI approximation and using local orbitals obtained from a Cholesky decomposition of the pseudodensity matrices (CDD). Both modifications lead to a significantly reduced prefactor compared to AO-MP2 and make RI-CDD-MP2 applicable to large molecules also with high-quality basis sets. RI-CDD-MP2 in its standard formulation scales asymptotically cubic.[33] Maurer et al.[33] also presented a second formulation, which employs the local density fitting approach from Werner et al.[22] and scales asymptotically linear. However, the prefactor of linear-scaling RI-CDD-MP2 is high, and speed-ups compared to the cubic scaling RI-CDD-MP2 formulation are only observed for very large systems.[33]

In this work, we aim to improve the RI-CDD-MP2 method by replacing the Coulomb metric used for the RI with an attenuated Coulomb metric.[46] Furthermore, we redesigned the RI-CDD-MP2 algorithm in order to optimally exploit the additional sparsity in the three-center integrals. In particular, we compute the Coulomb and exchange contributions to the MP2 energy separately using specialized algorithms for each contribution. For the exchange term, we present an efficient screening approach based on Schwarz estimates,[47] which is able to capture the exponential decay behavior of this contribution. The Coulomb contribution and the transformations of the three-center integrals from the AO basis to the basis of Cholesky-pseudo-MOs are accelerated with the natural blocking approach from Jung et al.[24,25] We also employ several upper bounds for the three-center integrals in order to reduce the number of elements in the naturally blocked integral tensors at an early stage during the transformations. In addition, an efficient transformation sequence for the three-center integrals is proposed, which involves an initial transformation of the AO-three-center integrals with the Cholesky factor of the ground state density followed by a transformation from occupied Cholesky-MOs to occupied Cholesky-pseudo-MOs for every Laplace point. For the efficient computation of three-center integrals in the atomic orbital basis, a distance including screening based on the integral partition bounds from Thompson and Ochsenfeld[48] is used.

The new MP2 method is shown to scale subquadratically for sufficiently large and sparse systems and displays a small prefactor. Significant speed-ups compared to the RI-CDD-MP2 method are obtained for molecular systems of various sizes. The asymptotic scaling of the method is cubic; however, the cubic scaling steps have a small prefactor and are irrelevant for all molecules used in the present study, of which the largest system is a DNA strand with 16 AT base pairs, 1052 atoms, 11 230 basis functions, and 37 248 auxiliary functions. The new method therefore enables efficient MP2 and double-hybrid DFT calculations on large molecular systems.

2. THEORY

2.1. Review of AO-MP2 and RI-CDD-MP2.

The MP2 energy of a closed-shell molecule can be obtained from the following expression:

$$E_{\mathrm{MP2}} = -\sum_{ijab} \frac{(ia|jb)[2(ia|jb) - (ib|ja)]}{\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j} \tag{1}$$

The indices *ijk*... denote occupied MOs, and *abc*... denote virtual MOs. ϵ values are the orbital energies. The denominator from eq 1 can be decoupled using a Laplace transformation.[44] In general, the Laplace transformation allows one to replace the fraction $\frac{1}{x}$ by an integral of an exponential function:

$$\frac{1}{x} = \int_0^\infty \exp(-xt)\,\mathrm{d}t \approx \sum_\alpha \omega_\alpha \exp(-xt_\alpha) \tag{2}$$

The integral from eq 2 can be approximated by numerical integration using quadrature points t_α and quadrature weights ω_α. In the case of the MP2 orbital energy denominator, only 5 to 8 Laplace points are usually sufficient to obtain errors on the order of a few μHartree.[44] The major benefit of applying the Laplace transformation to the MP2 energy denominator is the possibility to factorize the resulting exponential into product form:

$$E_{\mathrm{MP2}} = -\sum_\alpha \omega_\alpha \sum_{ijab} e^{-t_\alpha\epsilon_a} e^{-t_\alpha\epsilon_b} e^{+t_\alpha\epsilon_i} e^{+t_\alpha\epsilon_j} (ia|jb) \times [2(ia|jb) - (ib|ja)] \tag{3}$$

Inserting the expansion of the MOs as a linear combination of atomic orbitals (LCAO) and subsequently summing over MO indices gives the energy expression of AO-MP2:[11]

$$E_{\mathrm{MP2}}^{\mathrm{AO}} = -\sum_\alpha \sum_{\mu\nu\lambda\sigma} \sum_{\mu'\nu'\lambda'\sigma'} \underline{P}^{(\alpha)}_{\mu\mu'} \overline{P}^{(\alpha)}_{\nu\nu'} \underline{P}^{(\alpha)}_{\lambda\lambda'} \overline{P}^{(\alpha)}_{\sigma\sigma'} (\mu\nu|\lambda\sigma) \times [2(\mu'\nu'|\lambda'\sigma') - (\mu'\sigma'|\lambda'\nu')] \tag{4}$$

where $\underline{\mathbf{P}}^{(\alpha)}$ and $\overline{\mathbf{P}}^{(\alpha)}$ denote the occupied and virtual pseudodensities, respectively. The pseudodensity matrices are defined as follows:

$$\underline{P}^{(\alpha)}_{\mu\mu'} = \sqrt[4]{\omega_\alpha} \sum_{i} C_{\mu i} e^{+\epsilon_i t_\alpha} C_{\mu' i} \tag{5}$$

$$\overline{P}^{(\alpha)}_{\nu\nu'} = \sqrt[4]{\omega_\alpha} \sum_{a} C_{\nu a} e^{-\epsilon_a t_\alpha} C_{\nu' a} \tag{6}$$

In the CDD-MP2 method,[28] the pseudodensity matrices are subjected to a Cholesky decomposition with complete pivoting:[49]

$$\underline{P}^{(\alpha)}_{\mu\mu'} = \sum_{\underline{i}} \underline{L}^{(\alpha)}_{\mu\underline{i}} \underline{L}^{(\alpha)}_{\mu'\underline{i}} \tag{7}$$

$$\overline{P}^{(\alpha)}_{\nu\nu'} = \sum_{\overline{a}} \overline{L}^{(\alpha)}_{\nu\overline{a}} \overline{L}^{(\alpha)}_{\nu'\overline{a}} \tag{8}$$

The obtained Cholesky factors are the expansion coefficients of occupied and virtual Cholesky-pseudo-MOs $\phi^{(\alpha)}_{\underline{i}}$ and $\phi^{(\alpha)}_{\overline{a}}$, respectively:

$$\phi^{(\alpha)}_{\underline{i}}(\mathbf{r}) = \sum_{\mu} \underline{L}^{(\alpha)}_{\mu\underline{i}} \chi_\mu(\mathbf{r}) \tag{9}$$

$$\phi^{(\alpha)}_{\overline{a}}(\mathbf{r}) = \sum_{\nu} \overline{L}^{(\alpha)}_{\nu\overline{a}} \chi_\nu(\mathbf{r}) \tag{10}$$

Inserting the Cholesky factors into eq 4 gives the energy expression for CDD-MP2:[33]

$$E^{\mathrm{CDD}}_{\mathrm{MP2}} = -\sum_{\alpha} \sum_{\underline{i}\underline{j}\overline{a}\overline{b}} (\underline{i}\overline{a}|\underline{j}\overline{b})^{(\alpha)} [2(\underline{i}\overline{a}|\underline{j}\overline{b})^{(\alpha)} - (\underline{i}\overline{b}|\underline{j}\overline{a})^{(\alpha)}] \tag{11}$$

where

$$(\underline{i}\overline{a}|\underline{j}\overline{b})^{(\alpha)} = \sum_{\mu\nu\lambda\sigma} \underline{L}^{(\alpha)}_{\mu\underline{i}} \overline{L}^{(\alpha)}_{\nu\overline{a}} \underline{L}^{(\alpha)}_{\lambda\underline{j}} \overline{L}^{(\alpha)}_{\sigma\overline{b}} (\mu\nu|\lambda\sigma) \tag{12}$$

In the RI-CDD-MP2 method,[33] the ERIs are in addition approximated by RI:[40]

$$(\underline{i}\overline{a}|\underline{j}\overline{b})^{(\alpha)} \approx \sum_{PQ} (\underline{i}\overline{a}|P)^{(\alpha)} [\mathbf{V}^{-1}]_{PQ} (Q|\underline{j}\overline{b})^{(\alpha)} \tag{13}$$

$$V_{PQ} = \iint \chi_P(\mathbf{r}_1) \frac{1}{r_{12}} \chi_Q(\mathbf{r}_2)\, \mathrm{d}\mathbf{r}_1 \mathrm{d}\mathbf{r}_2 \tag{14}$$

where the indices PQ... denote auxiliary basis functions.

2.2. Calculation and Transformation of Three-Center Integrals.

In the new method presented in this work, RI with an erfc-attenuated Coulomb metric[25,46,50] is used for approximating the ERIs:

$$(\mu\nu|\lambda\sigma) \approx \sum_{PQ} (\mu\nu\vdots P) \tilde{C}_{PQ} (Q\vdots\lambda\sigma) \tag{15}$$

$$(\mu\nu\vdots P) = \iint \chi_\mu(\mathbf{r}_1)\chi_\nu(\mathbf{r}_1) \frac{\mathrm{erfc}(\omega r_{12})}{r_{12}} \chi_P(\mathbf{r}_2)\, \mathrm{d}\mathbf{r}_1 \mathrm{d}\mathbf{r}_2 \tag{16}$$

$$\tilde{C}_{PQ} = \sum_{RS} [\tilde{\mathbf{V}}^{-1}]_{PR} V_{RS} [\tilde{\mathbf{V}}^{-1}]_{SQ} \tag{17}$$

$$\tilde{V}_{PR} = \iint \chi_P(\mathbf{r}_1) \frac{\mathrm{erfc}(\omega r_{12})}{r_{12}} \chi_R(\mathbf{r}_2)\, \mathrm{d}\mathbf{r}_1 \mathrm{d}\mathbf{r}_2 \tag{18}$$

The erfc metric depends on the attenuation parameter ω and interpolates between the Coulomb metric and the overlap metric depending on the choice of ω. It was shown that the attenuated Coulomb metric with a value of 0.1 for ω provides comparable accuracy as the Coulomb metric without sacrificing useful sparsity in the three-center integrals.[51] Due to the central role of the local metric, we will in the following denote the new MP2 method as ω-RI-CDD-MP2. We want to stress that ω should not be viewed as an empirical parameter but as a threshold, because lowering ω allows one to systematically reduce the deviations from the Coulomb metric; an exact agreement with the Coulomb metric is obtained for ω equal to zero.

The computation of the three-center integrals from eq 16 would scale quadratically, if only the sparsity of $\mu\nu$-shell pairs is exploited. Due to the attenuation with the complementary error function, these integrals decay quickly with increasing bra-ket separation. We exploit this distance decay by using the rigorous distance-including screening based on integral partition bounds (IPB) from Thompson and Ochsenfeld;[48] by default, our implementation employs the approximate bound aIPB and neglects integrals below a threshold $\vartheta_{3\mathrm{c}}$. With aIPB screening, only an asymptotically linear-scaling number of three-center integrals in the AO basis needs to be computed.

The three-center integrals in the atomic orbital basis need to be transformed to the basis of Cholesky-pseudo-MOs for each Laplace point. This could be done by applying the transformations shown in eqs 19 and 20:

$$(\underline{i}\nu\vdots P)^{(\alpha)} = \sum_{\mu} \underline{L}^{(\alpha)}_{\mu\underline{i}} (\mu\nu\vdots P) \tag{19}$$

$$(\underline{i}\overline{a}\vdots P)^{(\alpha)} = \sum_{\nu} \overline{L}^{(\alpha)}_{\nu\overline{a}} (\underline{i}\nu\vdots P)^{(\alpha)} \tag{20}$$

A more efficient way to obtain the transformed integrals involves the Cholesky factors of the ground state density matrix **P**:

$$P_{\mu\nu} = \sum_{i} L_{\mu i} L_{\nu i} \tag{21}$$

Using **L**, the transformation from eq 19 can be replaced by a succession of two transformations:

$$(i\nu\vdots P) = \sum_{\mu} L_{\mu i} (\mu\nu\vdots P) \tag{22}$$

$$(\underline{i}\nu\vdots P)^{(\alpha)} = \sum_{i} T^{(\alpha)}_{\underline{i}i} (i\nu\vdots P) \tag{23}$$

The matrix $\mathbf{T}^{(\alpha)}$ employed in eq 23 transforms the occupied Cholesky-MOs to occupied Cholesky-pseudo-MOs. It can be derived by considering the following identity for the pseudodensity matrix $\underline{\mathbf{P}}^{(\alpha)}$:[52]

$$\underline{\mathbf{P}}^{(\alpha)} = \underline{\mathbf{P}}^{(\alpha)}\mathbf{SP} = \mathbf{PS}\underline{\mathbf{P}}^{(\alpha)}\mathbf{SP} \tag{24}$$

Equation 24 is closely related to the well-known idempotency condition for the density matrix:[53]

$$\mathbf{P} = \mathbf{PSP} = \mathbf{PSPSP} \tag{25}$$

After inserting the Cholesky factors of the occupied density matrix and the occupied pseudodensity matrix into eq 24, one can identify the sought transformation matrix as the product $\underline{\mathbf{L}}^{T(\alpha)}\mathbf{SL}$:

$$\underline{\mathbf{P}}^{(\alpha)} = \underline{\mathbf{L}}^{(\alpha)}\underline{\mathbf{L}}^{T(\alpha)} = \mathbf{L}\underbrace{\mathbf{L}^{T}\mathbf{S}\underline{\mathbf{L}}^{(\alpha)}}_{\mathbf{T}^{T(\alpha)}}\underbrace{\underline{\mathbf{L}}^{T(\alpha)}\mathbf{SL}}_{\mathbf{T}^{(\alpha)}}\mathbf{L}^{T} \tag{26}$$

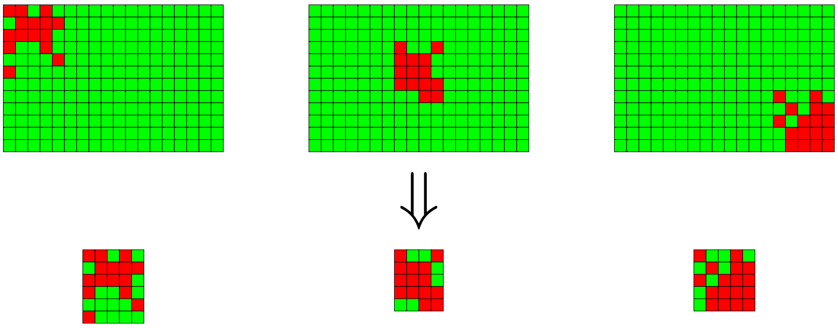


Figure 1. Illustration of the natural blocking format. A third-order tensor is stored as an array of matrices (top row). Red cells correspond to significant entries, while green cells correspond to insignificant entries. When natural blocking is applied, all columns and rows containing only insignificant entries are deleted (bottom row).

Thus, the elements of $\mathbf{T}^{(\alpha)}$ are given by

$$T^{(\alpha)}_{\underline{i}i} = (\underline{\mathbf{L}}^{T(\alpha)}\mathbf{SL})_{\underline{i}i} \tag{27}$$

The advantage of the presented transformation sequence is that eq 23 requires only a sum over occupied MOs instead of a sum over AOs as in eq 19. Due to the smaller number of occupied MOs, eq 23 is thus less computationally costly to evaluate, especially for extended basis sets. The additional transformation shown in eq 22 is also less expensive than the transformation in eq 19, because it has to be carried out only once and not for every Laplace point.

In order to reduce the scaling of the integral transformations, we employ the natural blocking approach from Jung et al.,[24,25] which provides a sparse data format for the three-center integrals. In general, a third-order tensor is often represented as a one-dimensional array of matrices; we will in the following denote these matrices as "slices" of the tensor. The index that labels the one-dimensional matrix array will be called the "slow index". One way to exploit sparsity in a third-order tensor would be to use a block-sparse data format for each slice. In contrast, in the natural blocking approach, entire rows and columns of the slices are deleted if they contain only insignificant values below a given threshold.[24,25] This is illustrated in Figure 1. The threshold used for deleting rows and columns will be called the "natural blocking threshold" ϑ_{NB}. Following Jung et al.,[25] we employ a separate threshold $\vartheta^{iaP}_{\mathrm{NB}}$ for the fully transformed $(\underline{i}\bar{a}{\vdots}P)^{(\alpha)}$ integrals. The advantage of natural blocking is that each slice remains a single, relatively large matrix after deleting rows and columns; in a block-sparse approach, in contrast, the slice is divided up into a large number of small blocks. As the relative efficiency of common matrix multiplications implementations is in general higher for larger matrices, a smaller computational overhead in multiplications involving the slices can be obtained with natural blocking. In order to efficiently exploit sparsity in the natural blocking format, it is important that the third-order tensor has a suitable sparsity pattern. In particular, for each value of the slow index, only $\mathcal{O}(1)$ rows and columns should be significant. For the three-center integrals with a local RI metric, this is fulfilled for both the atomic orbital basis and a basis of localized MOs, because all three indices couple only over short distances. In cases where the slices still have significant sparsity after row and column deletion, one might consider to also use a block-sparse matrix format for these slices; however, we have not exploited this so far.

As a specific example for the application of natural blocking, consider the $(P{\vdots}\mu\nu)$ integrals. If the integrals are ordered such that P is the slow index, each value of P is associated with an $N_{\mathrm{AO}} \times N_{\mathrm{AO}}$ slice, where N_{AO} denotes the number of AOs. In a large molecule, the $(P{\vdots}\mu\nu)$ integrals will only have significant values, if both χ_μ and χ_ν are spatially close to χ_P. Therefore, many elements of the $\mu\nu$ slice will be insignificant; in particular, many rows and columns will contain only negligible elements. By deleting these, the size of the slice can be reduced to $N_{\mathrm{AO(P)}} \times N_{\mathrm{AO(P)}}$, where $N_{\mathrm{AO(P)}}$ is the number of significant AOs for the given P. For bookkeeping purposes, it is important to store the indices of all significant rows (and columns) for a particular P in a list $\{\mu\}_P$. More precisely, the list $\{\mu\}_P$ can be build according to

$$\{\mu\}_P \equiv \{\mu | \max_{\nu} |(P{\vdots}\mu\nu)| > \vartheta_{\mathrm{NB}}\} \tag{28}$$

One situation, in which this list is needed, is the following transformation:

$$(P{\vdots}i\nu) = \sum_{\mu} L_{i\mu}(P{\vdots}\mu\nu) \tag{29}$$

This transformation is carried out with matrix multiplications of the $\mu\nu$ slices with the coefficient matrix $\mathbf{L}$. Before the multiplication of one particular $\mu\nu$ slice with the matrix $\mathbf{L}$ can be performed, all columns in $\mathbf{L}$, whose index μ is not in the list $\{\mu\}_P$, need to be deleted. This reduces the size of $\mathbf{L}$ to $N_{\mathrm{occ}} \times N_{\mathrm{AO(P)}}$. Due to the reduction of the dimensions of the involved matrices, the multiplication can be accelerated significantly. This illustrates how computational savings can be realized with natural blocking.

In an analogous manner to $\{\mu\}_P$, several other lists of significant orbital combinations are constructed and used during the integral transformations. An overview of them is shown in Table 1. The information about the significance of orbital pairs can also be used to create "inverted" lists; the list $\{\underline{i}\}^{(\alpha)}_{\bar{a}}$, e.g., can be obtained by "inversion" of $\{\bar{a}\}^{(\alpha)}_{\underline{i}}$. Furthermore, the letters used in our notation are interchangeable as long as they refer to the same type of orbital; the list $\{\bar{a}\}^{(\alpha)}_{\underline{i}}$, e.g., is the same as $\{\bar{b}\}^{(\alpha)}_{\underline{i}}$.

In contrast to Jung et al.,[24,25] we also employ several rigorous upper bounds for the three-center integrals. Using these, we can exploit sparsity at an earlier stage during the sequence of integral transformation. We also build sparse lists based on the integral bounds, for which we will use a slightly different notation with square brackets such as $[\bar{a}]^{(\alpha)}_{\underline{i}}$. For

Table 1. Overview of All Employed Lists of Significant Orbital Combinations, Which Are Used for Natural Blocking[a]

list of orbital pairs	integral tensor
$\{\bar{a}\}_{\underline{i}}^{(\alpha)}$	$(\underline{i}\bar{a}\vdots P)^{(\alpha)}$
$\{i\}_P$	$(P\vdots i\nu)$
$\{\nu\}_P$	$(P\vdots \mu\nu)$
$\{\underline{i}\}_P^{(\alpha)}$	$(P\vdots \underline{i}\bar{a})^{(\alpha)}$
$\{\bar{a}\}_P^{(\alpha)}$	$(P\vdots \underline{i}\bar{a})^{(\alpha)}$

[a] The lists are generated by inspecting entries of the corresponding integral tensors shown in the right column in analogy to eq 28. The indices of the integral tensors are sorted such that the leftmost index is the "slow index".

computing the upper bounds, the matrices **M** and **N** defined in eqs 30 and 31 are needed:

$$M_{\mu\nu} = \max_P |(\mu\nu\vdots P)| \tag{30}$$

$$N_{\mu P} = \max_\nu |(\mu\nu\vdots P)| \tag{31}$$

The entries of **M** and **N** are accumulated during the computation of the three-center integrals in the AO basis. Using these matrices, upper bounds for the transformed integrals can be obtained by multiplication with the absolute value of the MO or pseudo-MO coefficients. One example is shown in eq 32:

$$|(i\nu\vdots P)| \leq N_{iP} = \sum_\mu |L_{\mu i}| N_{\mu P} \tag{32}$$

The values N_{iP} of the transformed **N** matrix can then be used to construct the list $[i]_P$:

$$[i]_P \equiv \{i | N_{iP} > \vartheta_{\mathrm{NB}}\} \tag{33}$$

All employed lists based on upper bounds are shown in Table 2.

Table 2. Lists of Significant Orbital Pairs, Which Are Generated by Using Upper Bounds for the Three-Center Integrals

list of orbital pairs	upper bound
$[i]_P$	$N_{iP} = \sum_\mu \lvert L_{\mu i}\rvert N_{\mu P} \geq \lvert(i\nu\vdots P)\rvert$
$[\underline{i}]_P^{(\alpha)}$	$N_{\underline{i}P} = \sum_\mu \lvert \underline{L}_{\mu\underline{i}}^{(\alpha)}\rvert N_{\mu P} \geq \lvert(\underline{i}\nu\vdots P)^{(\alpha)}\rvert$
$[\bar{a}]_P^{(\alpha)}$	$N_{\bar{a}P} = \sum_\mu \lvert \bar{L}_{\mu\bar{a}}^{(\alpha)}\rvert N_{\mu P} \geq \lvert(\bar{a}\nu\vdots P)^{(\alpha)}\rvert$
$[\bar{a}]_{\underline{i}}^{(\alpha)}$	$M_{\underline{i}\bar{a}} = \sum_{\mu\nu} \lvert \underline{L}_{\mu\underline{i}}^{(\alpha)}\rvert M_{\mu\nu} \lvert\bar{L}_{\nu\bar{a}}^{(\alpha)}\rvert \geq \lvert(\underline{i}\bar{a}\vdots P)^{(\alpha)}\rvert$

With the sparse lists from Tables 1 and 2, the integral transformations can be carried out with a linear-scaling computational effort. The full algorithm for the transformations is shown in Figure 2. During the transformations, the integrals are ordered such that the auxiliary index P is the slow index. This allows one to perform all transformations as multiplications of the slices and the coefficient matrices **L**, $\mathbf{T}^{(\alpha)}$, and $\bar{\mathbf{L}}^{(\alpha)}$. After the transformations, the integrals are reordered such that $\underline{i}$ becomes the slow index; this ordering is more convenient for the computation of the energy.

```
for all $P$ do
    build $\{\nu\}_P$ and $[i]_P$ lists
    for all $i \in [i]_P$ and $\nu \in \{\nu\}_P$ do
        $(P\vdots i\nu) = \sum_{\mu\in\{\mu\}_P} L_{\mu i}\,(P\vdots \mu\nu)$
    end for
    build $\{i\}_P$ list
end for

for all $\alpha$ do
    build $[\underline{i}]_P^{(\alpha)}$ and $[\bar{a}]_P^{(\alpha)}$ lists
    for all $P$ do
        for all $\underline{i} \in [\underline{i}]_P^{(\alpha)}$ and $\nu \in \{\nu\}_P$ do
            $(P\vdots \underline{i}\nu)^{(\alpha)} = \sum_{i\in\{i\}_P^{(\alpha)}} T_{\underline{i}i}^{(\alpha)}\,(P\vdots i\nu)$
        end for

        for all $\underline{i} \in [\underline{i}]_P^{(\alpha)}$ and $\bar{a} \in [\bar{a}]_P^{(\alpha)}$ do
            $(P\vdots \underline{i}\bar{a})^{(\alpha)} = \sum_{\nu\in\{\nu\}_P} \bar{L}_{\nu\bar{a}}^{(\alpha)}\,(P\vdots \underline{i}\nu)^{(\alpha)}$
        end for
        build $\{\underline{i}\}_P^{(\alpha)}$ and $\{\bar{a}\}_P^{(\alpha)}$ lists
    end for
    reorder integrals: $(P\vdots \underline{i}\bar{a})^{(\alpha)} \rightarrow (\underline{i}\bar{a}\vdots P)^{(\alpha)}$
    build $\{\bar{a}\}_{\underline{i}}^{(\alpha)}$ list
    compute energy contribution (see Figures 3 and 4)
end for
```

Figure 2. Algorithm for the transformation of the three-center integrals. The leftmost index in the three-center integrals is the "slow index".

2.3. Computation of the MP2 Coulomb Energy.

In contrast to the implementation of RI-CDD-MP2 of Maurer et al.,[33] our new implementation treats the Coulomb and exchange contributions to the MP2 energy separately using specialized algorithms. As we will show, this is particularly beneficial in combination with a local RI metric. A separate treatment of Coulomb and exchange contributions has already been proposed in related work by Beuerle et al.[54] for low-scaling beyond-RPA methods and by Helmich-Paris et al.[55] for relativistic RI-CDD-MP2; in both works, also, a local RI metric was employed.

For the Coulomb contribution

$$E_{\mathrm{MP2}}^{\mathrm{C}} = \sum_\alpha E_{\mathrm{MP2}}^{\mathrm{C},(\alpha)} = -2\sum_\alpha \sum_{\underline{i}\underline{j}\bar{a}\bar{b}} (\underline{i}\bar{a}|\underline{j}\bar{b})^{(\alpha)}(\underline{i}\bar{a}|\underline{j}\bar{b})^{(\alpha)} \tag{34}$$

a matrix $\mathbf{Z}^{(\alpha)}$ is computed by contracting three-center integrals over the pseudo-MO indices:

$$Z_{PQ}^{(\alpha)} = \sum_{\underline{i}\bar{a}} (\underline{i}\bar{a}\vdots P)^{(\alpha)}(\underline{i}\bar{a}\vdots Q)^{(\alpha)} \tag{35}$$

Next, $\mathbf{Z}^{(\alpha)}$ is multiplied with the matrix $\tilde{\mathbf{C}}$ defined in eq 17:

$$D_{PQ}^{(\alpha)} = \sum_R Z_{PR}^{(\alpha)} \tilde{C}_{RQ} \tag{36}$$

From $\mathbf{D}^{(\alpha)}$, the contribution of the current Laplace point α to the MP2 Coulomb energy can be computed via the following equation:

$$E_{\mathrm{MP2}}^{\mathrm{C},(\alpha)} = -2\sum_{PQ} D_{PQ}^{(\alpha)} D_{QP}^{(\alpha)} \tag{37}$$

The contraction in eq 35 has the highest formal scaling (N^4), which is nevertheless lower than the formal scaling of the MP2 exchange (N^5). The use of natural blocking allows one to exploit both the sparsity of the $\underline{i}\bar{a}$-charge densities and the sparsity resulting from the local metric as shown in the algorithm in Figure 3. Thus, eq 35 can be evaluated with asymptotically linear scaling. Note that the presented algorithm would be suboptimal in combination with a Coulomb metric as used in the original RI-CDD-MP2 method, because the asymptotic scaling of eq 35 would be cubic with a relatively large prefactor.

```
$\mathbf{Z}^{(\alpha)} = 0$

for all $\underline{i}$ do
    for all $P \in \{P\}_{\underline{i}}^{(\alpha)}$ and $Q \in \{Q\}_{\underline{i}}^{(\alpha)}$ do
        $Z_{PQ}^{(\alpha)} \mathrel{+}= \sum_{\bar{a} \in \{\bar{a}\}_{\underline{i}}^{(\alpha)}} (\underline{i}\bar{a} \vdots P)^{(\alpha)} (\underline{i}\bar{a} \vdots Q)^{(\alpha)}$
    end for
end for

for all $P$ and $Q$ do
    $D_{PQ}^{(\alpha)} = \sum_R Z_{PR}^{(\alpha)} \tilde{C}_{RQ}$
end for

$E_{\text{MP2}}^{\text{C},(\alpha)} = -2 \sum_{PQ} D_{PQ}^{(\alpha)} D_{QP}^{(\alpha)}$
```

Figure 3. Algorithm for the computation of the Coulomb contribution to the ω-RI-CDD-MP2 energy.

The matrix multiplication in eq 36 scales cubically with small prefactor; note that the scaling could be reduced to quadratic by exploiting the sparsity of $\mathbf{Z}^{(\alpha)}$. Since $\mathbf{D}^{(\alpha)}$ is a dense matrix, eq 37 is a low-prefactor quadratic scaling step.

2.4. Computation of the MP2 Exchange Energy.

The algorithm used to evaluate the exchange contribution to the ω-RI-CDD-MP2 energy is shown in Figure 4. In this algorithm, the exchange energy contributions from individual $\underline{ij}$ pairs are computed according to

$$E_{\text{MP2}}^{\text{X},(\alpha),\underline{ij}} = \sum_{\bar{a}\bar{b}} (\underline{i}\bar{a}|\underline{j}\bar{b})^{(\alpha)} (\underline{i}\bar{b}|\underline{j}\bar{a})^{(\alpha)} \tag{38}$$

and summed up to obtain the total exchange energy $E_{\text{MP2}}^{\text{X}}$:

$$E_{\text{MP2}}^{\text{X}} = \sum_{\alpha} E_{\text{MP2}}^{\text{X},(\alpha)} = \sum_{\alpha} \sum_{\underline{ij}} E_{\text{MP2}}^{\text{X},(\alpha),\underline{ij}} \tag{39}$$

For each $\underline{ij}$ pair, only a subset of virtual pseudo-MOs is included in the sum of eq 38. The significant virtual pseudo-MOs are selected by Schwarz screening. The Schwarz upper bound for an individual four-center integral is given by

$$|(\underline{i}\bar{a}|\underline{j}\bar{b})^{(\alpha)}| \le Q_{\underline{i}\bar{a}}^{(\alpha)} Q_{\underline{j}\bar{b}}^{(\alpha)} \tag{40}$$

where the Schwarz factors are defined as follows:

$$Q_{\underline{i}\bar{a}}^{(\alpha)} = \sqrt{(\underline{i}\bar{a}|\underline{i}\bar{a})^{(\alpha)}} \tag{41}$$

The Schwarz factors are computed on the fly during the algorithm for the MP2 exchange (see Figure 4) using the RI approximation:

$$Q_{\underline{i}\bar{a}}^{(\alpha)} \approx \sqrt{\sum_{PQ} (\underline{i}\bar{a} \vdots P)^{(\alpha)} \tilde{C}_{PQ} (Q \vdots \underline{i}\bar{a})^{(\alpha)}} \tag{42}$$

The Schwarz estimates for the integral products appearing in the MP2 exchange are given by

```
for all $\underline{i}$ do
    for all $\bar{a} \in \{\bar{a}\}_{\underline{i}}^{(\alpha)}$ and $Q \in (Q)_{\underline{i}}^{(\alpha)}$ do
        $\tilde{B}_{\underline{i}\bar{a}}^{Q(\alpha)} = \sum_{P \in \{P\}_{\underline{i}}^{(\alpha)}} (\underline{i}\bar{a} \vdots P)^{(\alpha)} \tilde{C}_{PQ}$
    end for

    for all $\bar{a} \in \{\bar{a}\}_{\underline{i}}^{(\alpha)}$ and $\bar{b} \in \{\bar{b}\}_{\underline{i}}^{(\alpha)}$ do
        $(\underline{i}\bar{a}|\underline{i}\bar{b})^{(\alpha)} = \sum_{Q \in (Q)_{\underline{i}}^{(\alpha)} \cap \{Q\}_{\underline{i}}^{(\alpha)}} \tilde{B}_{\underline{i}\bar{a}}^{Q(\alpha)} (Q \vdots \underline{i}\bar{b})^{(\alpha)}$
    end for
    $E_{\text{MP2}}^{\text{X},(\alpha),\underline{i}\,\underline{i}} = \sum_{\bar{a} \in \{\bar{a}\}_{\underline{i}}^{(\alpha)}} \sum_{\bar{b} \in \{\bar{b}\}_{\underline{i}}^{(\alpha)}} (\underline{i}\bar{a}|\underline{i}\bar{b})^{(\alpha)} (\underline{i}\bar{b}|\underline{i}\bar{a})^{(\alpha)}$

    for all $\bar{a} \in \{\bar{a}\}_{\underline{i}}^{(\alpha)}$ do
        $Q_{\underline{i}\bar{a}}^{(\alpha)} = \sqrt{(\underline{i}\bar{a}|\underline{i}\bar{a})^{(\alpha)}}$
    end for

    $\underline{ij}$-prescreening (see eqs 44 to 46)

    for all significant $\underline{j} < \underline{i}$ do
        $E_{\text{MP2}}^{\text{X},(\alpha),\underline{ij}} = 0$
        determine significant $\bar{a}$ and $\bar{b}$ by
            Schwarz screening (see eq 43)
        for all significant $\bar{a}$ and $\bar{b}$ do
            $(\underline{i}\bar{a}|\underline{j}\bar{b})^{(\alpha)} = \sum_{Q \in (Q)_{\underline{i}}^{(\alpha)} \cap \{Q\}_{\underline{j}}^{(\alpha)}} \tilde{B}_{\underline{i}\bar{a}}^{Q(\alpha)} (Q \vdots \underline{j}\bar{b})^{(\alpha)}$
            $E_{\text{MP2}}^{\text{X},(\alpha),\underline{ij}} \mathrel{+}= \sum_{\bar{a}\bar{b}} (\underline{i}\bar{a}|\underline{j}\bar{b})^{(\alpha)} (\underline{i}\bar{b}|\underline{j}\bar{a})^{(\alpha)}$
        end for
        $E_{\text{MP2}}^{\text{X},(\alpha),\underline{ji}} = E_{\text{MP2}}^{\text{X},(\alpha),\underline{ij}}$
    end for
    $E_{\text{MP2}}^{\text{X},(\alpha)} = \sum_{\underline{i}\,\underline{j}} E_{\text{MP2}}^{\text{X},(\alpha),\underline{ij}}$
end for
```

Figure 4. Algorithm for the computation of the exchange energy contribution.

$$|(\underline{i}\bar{a}|\underline{j}\bar{b})^{(\alpha)} (\underline{i}\bar{b}|\underline{j}\bar{a})^{(\alpha)}| \le Q_{\underline{i}\bar{a}}^{(\alpha)} Q_{\underline{j}\bar{b}}^{(\alpha)} Q_{\underline{i}\bar{b}}^{(\alpha)} Q_{\underline{j}\bar{a}}^{(\alpha)} \tag{43}$$

Due to the structure of the MP2 exchange, the right-hand side of eq 43 not only provides a rigorous upper bound to the exchange energy contribution but also captures correctly its exponential decay behavior. We therefore use the product of Schwarz factors from eq 43 to screen virtual pseudo-MOs for a given $\underline{ij}$ pair and to neglect integrals below a threshold denoted as $\vartheta_{\text{schwarz}}$. With this screening, only an asymptotically linear-scaling number of four-center integrals needs to be computed.

For a given $\underline{ij}$ pair, only the virtual pseudo-MOs in $\{\bar{a}\}_{\underline{j}}^{(\alpha)}$ are considered in the screening. In order to further reduce the screening overhead, significant $\underline{ij}$ pairs are selected in a prescreening procedure, which is also based on Schwarz estimates. The exchange energy contribution from an $\underline{ij}$ pair given by eq 38 can be rigorously bound by the expression on the right-hand side of the following equation:

$$|E_{\text{MP2}}^{\text{X},(\alpha),\underline{ij}}| \le \sum_{\bar{a}\bar{b}} Q_{\underline{i}\bar{a}}^{(\alpha)} Q_{\underline{j}\bar{b}}^{(\alpha)} Q_{\underline{i}\bar{b}}^{(\alpha)} Q_{\underline{j}\bar{a}}^{(\alpha)} \tag{44}$$

This upper bound does in general not provide an accurate *quantitative* estimate of the $\underline{ij}$ pair energy, because all contributions on the right-hand side of eq 44 are positive in contrast to the right-hand side of eq 38, where positive and negative contributions to the sum can cancel. Nevertheless, it is useful as a *qualitative* measure for the importance of an $\underline{ij}$ pair.

Table 3. Mean Absolute Deviations (MAD), Maximum Absolute Deviations (MAX), and Root Mean Square Deviation (RMSD) Compared to Canonical RI-MP2 in ω-RI-CDD-MP2 Calculations on the L7 Test Set[64] with Different Values for the Attenuation Parameter ω[a]

	Δ_{abs}			Δ_{int}		
ω	MAD	RMSD	MAX	MAD	RMDS	MAX
0.50	1.48×10^{-0}	1.83×10^{-0}	3.83×10^{-0}	4.70×10^{-1}	6.59×10^{-1}	1.18×10^{-0}
0.20	2.70×10^{-1}	3.52×10^{-1}	8.19×10^{-1}	1.57×10^{-1}	2.18×10^{-1}	3.77×10^{-1}
0.10	4.06×10^{-2}	5.73×10^{-2}	1.49×10^{-1}	3.77×10^{-2}	5.09×10^{-2}	8.39×10^{-2}
0.05	4.95×10^{-3}	7.47×10^{-3}	2.04×10^{-2}	5.73×10^{-3}	7.56×10^{-3}	1.22×10^{-2}
0.01	3.54×10^{-5}	4.95×10^{-5}	1.41×10^{-4}	4.49×10^{-5}	4.82×10^{-5}	7.61×10^{-5}
0.00	9.23×10^{-7}	1.53×10^{-6}	4.52×10^{-6}	1.89×10^{-6}	2.79×10^{-6}	5.36×10^{-6}

[a] For the octadecane monomers and dimer, 7 Laplace points were used; 10 Laplace points were used for all other systems. All screening and natural blocking thresholds were set to zero. Δ_{abs} denotes the error in absolute energies of monomers and dimers; equivalent monomers were only considered once. Δ_{int} denotes the error in interaction energies. All values are given in kcal/mol. The remaining errors for $\omega = 0.00$ arise from the Laplace transformation.

The expression from eq 44 can be implemented very efficiently by first performing a matrix multiplication of the matrix of Schwarz integrals with its transpose:

$$A_{\underline{i}\underline{j}}^{(\alpha)} = \sum_{\bar{a}\in\{\bar{a}\}_{\underline{i}}^{(\alpha)}} Q_{\underline{i}\bar{a}}^{(\alpha)} Q_{\bar{a}\underline{j}}^{(\alpha)} \tag{45}$$

Then, the squares of the entries of $\mathbf{A}^{(\alpha)}$ can be used to screen $\underline{i}\underline{j}$ pairs:

$$|E_{\text{MP2}}^{\text{X},(\alpha),\underline{i}\underline{j}}| \leq A_{\underline{i}\underline{j}}^{(\alpha)2} \tag{46}$$

An $\underline{i}\underline{j}$ pair is neglected, if the right-hand side of eq 46 is smaller than a threshold denoted as $\vartheta_{\underline{i}\underline{j}}$.

For all $\underline{i}\underline{j}\bar{a}\bar{b}$ combinations that are significant according to the presented screening approach, four-center integrals are built in two steps. In the first step, a third-order tensor $\tilde{B}^{(\alpha)}$ is computed by transforming the auxiliary index of the three-center integrals with the matrix $\tilde{\mathbf{C}}$ defined in eq 17:

$$\tilde{B}_{\underline{i}\bar{a}}^{Q(\alpha)} = \sum_{P} (\underline{i}\bar{a}{:}P)^{(\alpha)} \tilde{C}_{PQ} \tag{47}$$

In the second step, the $\tilde{B}^{(\alpha)}$ tensor is contracted with three-center integrals over the auxiliary index:

$$(\underline{i}\bar{a}|\underline{j}\bar{b})^{(\alpha)} = \sum_{Q} \tilde{B}_{\underline{i}\bar{a}}^{Q(\alpha)} (Q{:}\underline{j}\bar{b})^{(\alpha)} \tag{48}$$

Even though the number of four-center integrals that needs to be formed scales linearly, both steps from eqs 47 and 48 would scale quadratically, if implemented naïvely. The reason for this is the quadratic scaling number of significant elements in the $\tilde{B}^{(\alpha)}$ tensor. This can be explained from eq 47, where the indices $\underline{i}$ and P are coupled only over short distances due to the erfc metric; $\underline{i}$ and Q, in contrast, couple over large distances, because the matrix $\tilde{\mathbf{C}}$ is densely populated. It turns out, however, that the index Q in eqs 47 and 48 can be restricted to a subset of the space of auxiliary functions. In order to realize this, consider the MP2 exchange energy expression with the RI plugged in:

$$E_{\text{MP2}}^{\text{X},(\alpha)} = \sum_{\underline{i}\bar{a}\underline{j}\bar{b}} \sum_{PQRS} (\underline{i}\bar{a}{:}P)^{(\alpha)} \tilde{C}_{PQ} (Q{:}\underline{j}\bar{b})^{(\alpha)} (\underline{i}\bar{b}{:}R)^{(\alpha)} \tilde{C}_{RS} (S{:}\underline{j}\bar{a})^{(\alpha)} \tag{49}$$

The coupling of $\underline{i}$ to both $\bar{a}$ and $\bar{b}$ decays exponentially because of the appearing $\underline{i}\bar{a}$- and $\underline{i}\bar{b}$-charge densities. $\bar{a}$ and $\bar{b}$ also appear next to the index $\underline{j}$ on the ket-side of the four-center integrals, where they are coupled to Q and S by the short-ranged erfc-function. By combining these couplings, one can restrict the space of auxiliary functions χ_Q and χ_S that need to be included for a given $\underline{i}$. For a particular $\underline{i}$, a list $(Q)_{\underline{i}}^{(\alpha)}$ is constructed from the union of auxiliary function indices Q in all lists $\{Q\}_{\bar{a}}^{(\alpha)}$ for all $\bar{a}$ in $\{\bar{a}\}_{\underline{i}}^{(\alpha)}$. For readers familiar with the sparse-map formalism of Pinski et al.,[34] it shall be mentioned that this would correspond to a "chaining operation" of the lists $\{\bar{a}\}_{\underline{i}}^{(\alpha)}$ and $\{Q\}_{\bar{a}}^{(\alpha)}$. With these restrictions for the index Q, eqs 47 and 48 can be evaluated with an asymptotically linear-scaling computational effort.

3. COMPUTATIONAL DETAILS

The described method was implemented in a development version of the quantum chemistry program FermiONs++.[56−58] Unless stated explicitly, the def2-SVP basis set[59] was used in combination with the corresponding auxiliary basis set.[60] Shell pairs with an overlap of less than 10^{-12} are omitted from the calculation. An integral screening threshold of 10^{-10} is used during the SCF. The DIIS method[61] is employed for accelerating SCF convergence. The SCF energy is converged to 10^{-7} and the norm of the commutator ($\mathbf{FPS}-\mathbf{SPF}$), where $\mathbf{F}$ is the Fock matrix and $\mathbf{S}$ is the overlap matrix, to 10^{-6}. Unless mentioned explicitly, five Laplace quadrature points are used. The Laplace points are determined using a minimax algorithm as described in ref 62; the number of integration points is reduced automatically in our implementation if the fitting interval is small, and no improved accuracy can be obtained with more integration points. In the calculation of pseudodensity matrices, Fermi-shifting as proposed by Ayala and Scuseria is applied.[15] The pseudodensity matrices are orthogonalized prior to the pivoted Cholesky decomposition as described in ref 51; afterward, the orthogonalization is reverted. The frozen-core approximation is employed in all calculations. For comparison, calculations are carried out with the RI-CDD-MP2 method[33] in the standard formulation, which scales cubically in the asymptotic limit, and with the implementation of canonical RI-MP2 in FermiONs++. All timings were performed using 20 threads on dual-processor Intel Xeon CPU E5-2630 v4 @ 2.20 GHz machines with 256 GB RAM and a solid-state drive (SSD) with a capacity of 1.7 TB.

Table 4. Benchmark Calculations for the aIPB Screening Threshold (ϑ_{3c}) and the Schwarz Screening Threshold ($\vartheta_{\mathrm{schwarz}}$)[a]

ϑ_{3c}				$\vartheta_{\mathrm{schwarz}}$			
value	MAD	RMSD	MAX	value	MAD	RMSD	MAX
10^{-5}	3.58×10^{-1}	1.16×10^{-0}	3.85×10^{-0}	10^{-8}	1.08×10^{-0}	1.51×10^{-0}	3.28×10^{-0}
10^{-6}	1.18×10^{-4}	1.99×10^{-4}	4.73×10^{-4}	10^{-9}	8.62×10^{-2}	1.09×10^{-1}	2.51×10^{-1}
10^{-7}	1.18×10^{-5}	2.43×10^{-5}	7.29×10^{-5}	10^{-10}	6.68×10^{-3}	7.86×10^{-3}	1.61×10^{-2}
10^{-8}	1.36×10^{-6}	2.42×10^{-6}	6.58×10^{-6}	10^{-11}	5.29×10^{-4}	6.06×10^{-4}	1.01×10^{-3}

[a] Mean absolute deviations (MAD), maximum absolute deviations (MAX), and root mean square deviation (RMSD) on the test set from Table 7 are shown. The reference energies were computed using ω-RI-CDD-MP2 without integral screening and natural blocking. 7 Laplace points are used for $[S_8]_5$; 8 Laplace points were used in all other calculations. All values are given in kcal/mol.

4. RESULTS

4.1. Accuracy of the Introduced Approximations.

Among the used approximations, we first consider the Laplace transformation of the MP2 denominator. This approximation has been used extensively, and numerous benchmarks on its accuracy exist;[11,15,44,45,62,63] in most cases, 5−8 Laplace points are sufficient in order to reach μHartree accuracy. As the Laplace transformation for MP2 is a well-established approximation, we did not analyze it further.

Next, we focus on the error introduced by the erfc-attenuated Coulomb metric. The accuracy of this approximation has been studied in detail for SOS-MP2,[25] direct RPA,[51] and full MP2 on small molecules.[46] In order to analyze the influence of the local metric on the accuracy of full MP2 including exchange on larger systems, we compare calculations with ω-RI-CDD-MP2 to canonical RI-MP2 with the standard Coulomb metric. The calculations were performed on the L7 test set,[64] and the results are shown in Table 3. When ω approaches zero, the deviations from the values computed with the Coulomb metric quickly decrease for both absolute energies and interaction energies. For $\omega = 0.1$, mean absolute deviations well below 0.1 kcal/mol are obtained. This is in line with the findings reported by Luenser et al.,[51] who showed that a value of 0.1 enables highly accurate direct RPA calculations. We thus set ω to 0.1 in all ω-RI-CDD-MP2 calculations shown in the following.

The remaining parameters were determined by carrying out benchmark calculations on a set of 11 molecules from the integral screening test set from ref 65. The used molecules all have between 40 and 146 atoms; the average number of atoms amounts to 98. Only the parameter of interest was varied in these calculations, while the others were set to zero, in order to isolate the effect of each parameter. Correlation energies computed with ω-RI-CDD-MP2 without integral screening and without natural blocking are used as references, which excludes the Laplace transformation and the RI as potential sources of error. The influence of the employed integral screening thresholds on the accuracy is analyzed in Tables 4 and 5. For each screening, the accuracy can be improved systematically by lowering the threshold. Suitable values for accurate calculations on large molecules were determined to be 10^{-6} for the aIPB threshold, 10^{-9} for the Schwarz screening threshold, and 10^{-6} for the $\underline{ij}$ prescreening threshold, respectively, as they lead to mean errors below 0.1 kcal/mol.

Table 5. Benchmark Calculations for the Threshold for Prescreening $\underline{ij}$ pairs ($\vartheta\underline{ij}$)[a]

$\vartheta\underline{ij}$			
value	MAD	RMSD	MAX
10^{-4}	1.27×10^{-0}	1.33×10^{-0}	2.08×10^{-0}
10^{-5}	1.00×10^{-1}	1.08×10^{-1}	1.68×10^{-1}
10^{-6}	1.07×10^{-2}	1.13×10^{-2}	1.75×10^{-2}
10^{-7}	8.79×10^{-4}	9.46×10^{-4}	1.53×10^{-3}

[a] Mean absolute deviations (MAD), maximum absolute deviations (MAX), and root mean square deviation (RMSD) on the test set from Table 7 are shown. The reference energies were computed using ω-RI-CDD-MP2 without integral screening and natural blocking. 7 Laplace points are used for $[S_8]_5$; 8 Laplace points are used in all other calculations. All values are given in kcal/mol.

In Table 6, the errors introduced by natural blocking are displayed. Thresholds of $\vartheta_{\mathrm{NB}} = 10^{-6}$ and $\vartheta_{\mathrm{NB}}^{iaP} = 5.0 \times 10^{-6}$ are sufficient to obtain mean errors below 0.1 kcal/mol and are thus employed in all calculations with def2-SVP basis shown in the following. The same set of values has also been used by Jung et al.[25] For ϑ_{NB}, a significant decrease in accuracy can be observed in Table 6 upon increasing the threshold from 10^{-6} to 10^{-5}; a similar effect occurs for ϑ_{3c} in Table 4. This is caused by relatively large errors for graphite$_{96}$ and CNT$_{80}$, which are both systems with strongly delocalized electronic structure and a high degree of symmetry. Due to the symmetry, a large number of integrals with nearly identical values are neglected if the thresholds are raised above a certain value. One thus has to apply care in choosing thresholds for systems of this kind.

The influence of the integral screening and natural blocking thresholds on the accuracy was also studied on the S22 test set,[66] and the corresponding data is shown in the Supporting Information. Due to the small size of the molecules in the S22 test set, only a few contributions are neglected by integral screening or natural blocking and the obtained errors are therefore significantly smaller than the errors obtained with the employed test set containing large molecules.

The combined influence of the determined thresholds is analyzed in Table 7. A mean error of 0.103 kcal/mol compared to canonical RI-MP2 is obtained, which indicates that the employed set of thresholds is suitable to obtain chemically accurate results also for larger molecular systems. Therefore, we use this particular set of thresholds for the timings with the def2-SVP basis set shown in Section 4.2. In Table 7, also, the errors obtained with RI-CDD-MP2 using the recommended thresholds from ref 33 are shown. The mean error obtained with RI-CDD-MP2 is more than twice as large. Thus, we conclude that the use of these settings in the timings shown in Section 4.2 does not introduce a biased advantage for ω-RI-CDD-MP2 in the comparison with RI-CDD-MP2. In Table 8, calculations with a set of suitable thresholds are shown that allow highly accurate computations with a def2-TZVP basis set.

4.2. Scaling Behavior and Efficiency.

Timings on linear n-alkanes are shown in Figure 5 and Table 9. The wall time required for the ω-RI-CDD-MP2 calculations is smaller than

Table 6. Benchmark Calculations for the Natural Blocking Thresholds ϑ_{NB} and ϑ_{NB}^{iaP}[a]

ϑ_{NB}				ϑ_{NB}^{iaP}			
value	MAD	RMSD	MAX	value	MAD	RMSD	MAX
1.0×10^{-5}	4.51×10^{-1}	1.09×10^{-0}	3.54×10^{-0}	1.0×10^{-5}	7.69×10^{-1}	2.12×10^{-0}	6.99×10^{-0}
1.0×10^{-6}	4.99×10^{-4}	1.02×10^{-3}	3.28×10^{-3}	5.0×10^{-6}	4.08×10^{-2}	9.55×10^{-2}	3.10×10^{-1}
1.0×10^{-7}	4.18×10^{-6}	1.00×10^{-5}	3.26×10^{-5}	1.0×10^{-6}	2.08×10^{-4}	2.86×10^{-4}	6.50×10^{-4}
				1.0×10^{-7}	5.70×10^{-6}	1.11×10^{-5}	3.02×10^{-5}

[a] Mean absolute deviations (MAD) and maximum absolute deviations (MAX) on the test set from Table 7 are shown. The reference energies were computed using ω-RI-CDD-MP2 without integral screening and natural blocking. 7 Laplace points are used for $[S_8]_5$; 8 Laplace points are used in all other calculations. All values are given in kcal/mol.

Table 7. Errors Compared to Canonical RI-MP2 with a Coulomb Metric in ω-RI-CDD-MP2 Calculations ($\Delta_{\omega\text{-RI-CDD}}$) and RI-CDD-MP2 Calculations ($\Delta_{\text{RI-CDD}}$)[a]

molecule	$\Delta_{\omega\text{-RI-CDD}}$	$\Delta_{\text{RI-CDD}}$
amylose$_4$	1.69×10^{-2}	9.89×10^{-2}
angiotensin	9.08×10^{-2}	2.39×10^{-1}
β-carotene	1.85×10^{-2}	5.69×10^{-2}
CNT$_{80}$	1.04×10^{-1}	3.56×10^{-1}
diamond$_{102}$	9.89×10^{-2}	1.30×10^{-1}
DNA$_2$	1.15×10^{-1}	2.20×10^{-1}
graphite$_{96}$	2.56×10^{-1}	2.36×10^{-1}
LiF$_{72}$	7.64×10^{-2}	1.28×10^{-1}
polyethyne$_{64}$	1.01×10^{-1}	5.22×10^{-2}
polyyne$_{64}$	1.25×10^{-1}	5.20×10^{-2}
$[S_8]_5$	1.33×10^{-1}	9.70×10^{-1}
MAD	1.03×10^{-1}	2.31×10^{-1}
MRD [%]	0.0019	0.0051

[a] All calculations employ a def2-SVP basis set. For ω-RI-CDD-MP2, the following thresholds are used: $\vartheta_{3c} = 10^{-6}$, $\vartheta_{\text{schwarz}} = 10^{-9}$, $\vartheta_{ij} = 10^{-6}$, $\vartheta_{NB} = 10^{-6}$, and $\vartheta_{NB}^{iaP} = 5 \times 10^{-6}$. For RI-CDD-MP2, a QQR screening threshold of 10^{-9} and a block-sparse linear algebra threshold of 10^{-6} are used. 7 Laplace points are used for $[S_8]_5$; 8 Laplace points are used in all other calculations. All values are given in kcal/mol. MAD and MRD denote mean absolute and mean relative deviation, respectively. The MRD values are given in percent of the total RI-MP2 correlation energy.

Table 8. Errors Compared to Canonical RI-MP2 with a Coulomb Metric in ω-RI-CDD-MP2 Calculations ($\Delta_{\omega\text{-RI-CDD}}$) and RI-CDD-MP2 Calculations ($\Delta_{\text{RI-CDD}}$)[a]

molecule	$\Delta_{\omega\text{-RI-CDD}}$	$\Delta_{\text{RI-CDD}}$
amylose$_4$	1.87×10^{-2}	2.22×10^{-2}
angiotensin	6.90×10^{-2}	7.08×10^{-2}
β-carotene	2.75×10^{-2}	2.07×10^{-2}
CNT$_{80}$	1.76×10^{-1}	7.53×10^{-2}
diamond$_{102}$	3.91×10^{-2}	3.55×10^{-2}
DNA$_2$	6.87×10^{-2}	7.84×10^{-2}
graphite$_{96}$	1.32×10^{-1}	4.58×10^{-2}
LiF$_{72}$	3.20×10^{-2}	5.01×10^{-3}
polyethyne$_{64}$	2.22×10^{-2}	1.71×10^{-2}
polyyne$_{64}$	1.55×10^{-1}	2.78×10^{-2}
$[S_8]_5$	1.09×10^{-2}	4.36×10^{-1}
MAD	6.83×10^{-2}	7.58×10^{-2}
MRD [%]	0.00098	0.00143

[a] The def2-TZVP basis set was used in all calculations. For ω-RI-CDD-MP2, the following thresholds are used: $\vartheta_{3c} = 10^{-7}$, $\vartheta_{\text{schwarz}} = 10^{-10}$, $\vartheta_{ij} = 10^{-6}$, $\vartheta_{NB} = 10^{-7}$, and $\vartheta_{NB}^{iaP} = 10^{-6}$. For RI-CDD-MP2, a QQR screening threshold of 10^{-10} and a block-sparse linear algebra threshold of 10^{-8} are used. 7 Laplace points are used for $[S_8]_5$; 8 Laplace points are used in all other calculations. All values are given in kcal/mol. MAD and MRD denote mean absolute and mean relative deviation, respectively. The MRD values are given in percent of the total RI-MP2 correlation energy.

the wall time used for the RI-CDD-MP2 reference in all of these calculations, and the relative differences rise significantly for increasing chain lengths. This suggests that ω-RI-CDD-MP2 has both a lower prefactor and reduced effective scaling compared to RI-CDD-MP2. For comparison, timings with canonical RI-MP2 are shown. For the smallest alkanes, ω-RI-CDD-MP2 is slower than RI-MP2 because of the screening overhead and the required loop over Laplace points. Due to the lower scaling of ω-RI-CDD-MP2, the situation is reversed for large alkanes; the crossover occurs between 40 and 60 carbon atoms. The largest calculation was done on $C_{300}H_{602}$ with 7210 basis functions and required a compute time of 2400 s, which illustrates that the ω-RI-CDD-MP2 method exploits the available sparsity in an efficient manner. The scaling is analyzed in detail in Table 9. As shown, the effective scaling steadily decreases for alkanes with up to 200 carbon atoms, where the minimal scaling of 1.37 is reached. This shows that the scaling of the time-dominating steps has been reduced to be linear. Between a chain length of 200 and 300 carbon atoms, the effective scaling increases slightly to 1.56. This can be explained by the remaining asymptotically cubic and quadratic scaling steps. For some of these steps, the scaling could be reduced further; the computation of the pseudodensity matrices according to eqs 5 and 6, e.g., could be made to

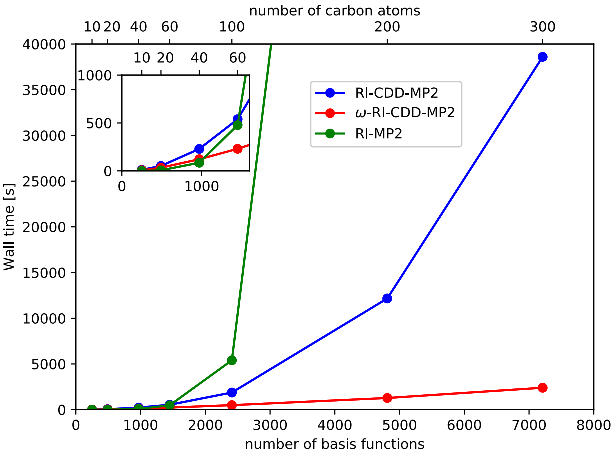


Figure 5. Timings on linear *n*-alkanes with up to 300 carbon atoms. The ω-RI-CDD-MP2 method (red) is compared to RI-CDD-MP2 (blue) and canonical RI-MP2 (green). In all calculations, the def2-SVP basis set is employed in combination with the thresholds from Table 7. The inset shows the timings for the alkanes with up to 60 carbon atoms.

Table 9. Total Wall Times in Seconds and Scaling in Calculations for Linear *n*-Alkanes with the def2-SVP Basis Set and 5 Laplace Points[a]

#carbon atoms	#basisf.	wall time	scaling
10	250	7	
20	490	33	2.32
40	970	121	1.90
60	1450	230	1.59
100	2410	494	1.50
200	4810	1277	1.37
300	7210	2400	1.56

[a] The scaling exponents are computed relative to the calculation on the previous system.

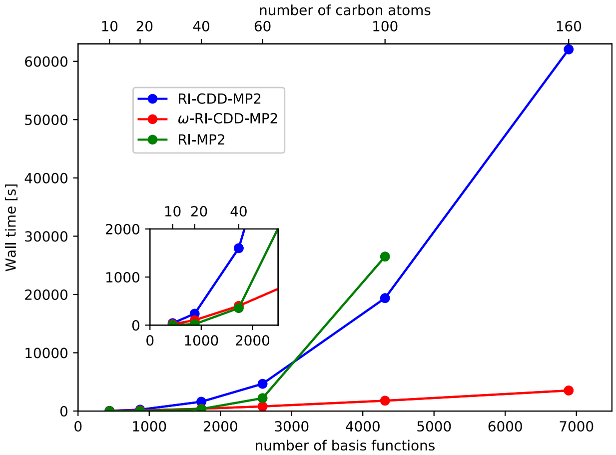


Figure 7. Timings on linear *n*-alkanes with up to 160 carbon atoms. The def2-TZVP basis set in combination with the thresholds from Table 8 was used in all calculations. The ω-RI-CDD-MP2 method (red) is compared to RI-CDD-MP2 (blue) and canonical RI-MP2 (green). The inset shows the timings for the alkanes with up to 40 carbon atoms.

be linear scaling by expressing the pseudodensities in terms of matrix exponentials[67] and using sparse linear algebra in their computation. Other steps such as the inversion of the matrix of two-center integrals in eq 18 or the matrix multiplications from eq 17 are intrinsically cubic scaling. Nevertheless, all steps with a scaling $>O(M)$ have a low prefactor and only become relevant for extremely large systems.

In Figure 6, the number of computed three- and four-center integrals is shown for the largest considered alkanes. The number of significant integrals that need to be computed scales close to linearly, which illustrates the effectiveness of the employed integral screening. Also, the number of significant $\underline{ij}$ pairs in the exchange contribution and the number of elements in the lists $(Q)_{\underline{i}}$ show near linear scaling. The fact that the number of untransformed and transformed three-center integrals scales linearly allows one to also reduce the RAM and disk space requirements and the I/O overhead to linear.

Figure 7 shows calculations for linear *n*-alkanes with the def2-TZVP basis set. As for the def2-SVP basis set, significant speed-ups compared to RI-CDD-MP2 are observed with the new method. The effective scaling between the largest systems with 100 and 160 carbon atoms amounts to 1.46. This demonstrates that the ω-RI-CDD-MP2 method scales favorably with a small prefactor also for high-quality basis sets.

In order to illustrate the applicability of the new method to realistic systems such as large biomolecules, we also performed calculations on DNA strands. The corresponding timings are shown in Figure 8 and Table 10. Also for the DNA strands, significant speed-ups compared to RI-CDD-MP2 are obtained with the new ω-RI-CDD-MP2 method. The crossover between canonical RI-MP2 and ω-RI-CDD-MP2 occurs between AT$_2$ and AT$_4$. The largest ω-RI-CDD-MP2 calculation was performed on a DNA strand with 16 base pairs, 1052 atoms, 11 230 basis functions, and 37 248 auxiliary functions and required a computation time of 35.6 h on a single node (for the specifications, see Computational Details). This demonstrates that the ω-RI-CDD-MP2 method allows for efficient MP2 calculations for large biomolecular systems. Table 10 also shows how much individual steps contribute to the total wall time. For most systems, the computation of the exchange energy is the rate-limiting step, but also, the transformations of the three-center integrals require a significant fraction of the overall compute time. Note that the separate treatment of Coulomb and exchange contributions in the ω-RI-CDD-MP2 method allows for even more efficient SOS-MP2 calculations by simply omitting the exchange contributions. On the basis of the timings from Table 10, one can expect a ω-SOS-RI-CDD-MP2 method to be faster by roughly a factor of 2. The effective scaling for the DNA strands is also shown in Table 10. The effective scaling decreases for the larger DNA strands and reaches a value of 1.90 between AT$_8$ and AT$_{16}$. This demonstrates that subquadratic scaling can be achieved also for realistic systems.

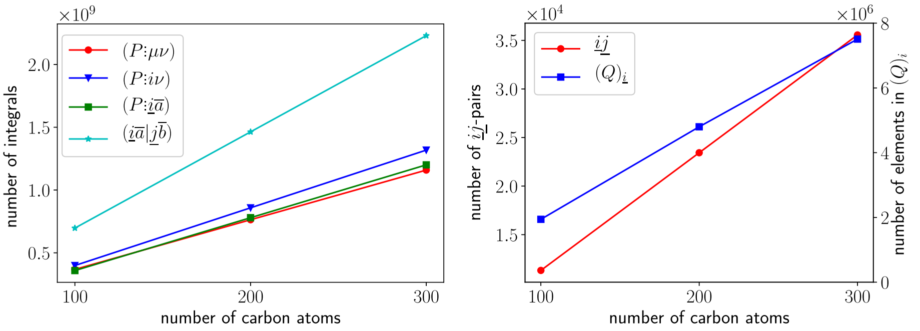


Figure 6. Left: Numbers of computed three-center and four-center integrals for linear alkanes with 100 to 300 carbon atoms. Right: The number of significant $\underline{ij}$ pairs from the $\underline{ij}$ prescreening and the number of elements in the lists $(Q)_{\underline{i}}$ are shown. All values refer to the first Laplace point only.

https://dx.doi.org/10.1021/acs.jctc.0c00600
J. Chem. Theory Comput. 2020, 16, 6856−6868

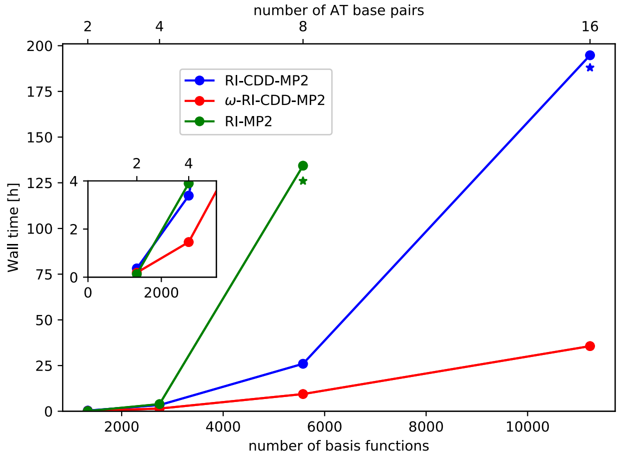


Figure 8. Timings on DNA strands with 2 to 16 adenine−thymine base pairs. The ω-RI-CDD-MP2 method (red) is compared to RI-CDD-MP2 (blue) and canonical RI-MP2 (green). The RI-CDD-MP2 calculation on AT_{16} could not be performed due to disk space limitations; the corresponding wall time marked with a blue star was extrapolated on the basis of the effective scaling between AT_4 and AT_8. The RI-MP2 calculation for AT_8 was extrapolated assuming an N^5 scaling. The inset shows the timings for AT_2 and AT_4.

Table 10. Total Wall Times and Wall Times of the Individual Computation Steps in Calculating Adenine−Thymine Base Pairs AT_n with the def2-SVP Basis Set and 5 Laplace Points[a]

	AT_2	AT_4	AT_8	AT_{16}
wall time	681	5252	33 851	128 274
t_{exchange}	298	2396	13 671	61 318
t_{Coulomb}	45	386	2094	7449
$t_{\text{transform}}$	259	2109	16 550	54 157
t_{3c}	70	303	1186	3315
scaling		2.82	2.63	1.90

[a] All timings are given in seconds. t_{exchange} and t_{Coulomb} denote the time used for the exchange and Coulomb contributions to the correlation energy, respectively. t_{3c} and $t_{\text{transform}}$ denote the wall time used for computing the AO three-center integrals and transforming them into the basis of Cholesky-pseudo-MOs, respectively. The scaling exponents are computed relative to the calculation on the previous system.

5. CONCLUSION

We presented a new method for the efficient computation of MP2 energies denoted as ω-RI-CDD-MP2. Local MOs obtained from a Cholesky decomposition of density and pseudodensity matrices and an erfc-attenuated Coulomb metric used for the RI approximation provide a high degree of sparsity in the occurring intermediates. This sparsity is exploited with efficient integral screening techniques and a sparse linear algebra approach called natural blocking. It was demonstrated in test calculations that the errors introduced by the used approximations can be systematically decreased by tightening the corresponding thresholds. Timings on alkane chains and DNA strands were performed with thresholds that allow one to obtain chemically accurate correlation energies also for large molecular systems. It was shown that the method displays effective subquadratic scaling behavior in conjunction with a small prefactor for sufficiently large and extended molecules. For both alkanes and DNA strands, significant speed-ups compared to the RI-CDD-MP2 method are observed. The largest considered DNA system with 1052 atoms, 11 230 basis functions, and 37 248 auxiliary functions required a computation time of 35.6 h on a single node, which illustrates the potential of the method for applications to large biomolecular systems.

Further improvements might be possible by using iterative orbital localization techniques in order to increase the locality of Cholesky-MOs and Cholesky-pseudo-MOs. The applicability of the method might be extended by developing efficient parallelization schemes or by adapting the method to GPUs. The basis set error could be reduced by combining the method with explicitly correlated F12 approaches. Since the method relies only on integral screening and sparse linear algebra and not on any kind of domain approximation or fragmentation, no problems with discontinuities in the potential energy surfaces are to be expected, which makes the accurate and reliable computation of gradients possible. In future work, the development of analytical derivatives would therefore be desirable in order to allow for accurate and efficient computations of properties of large molecules.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.0c00600.

Benchmark calculations on the S22 test set[66] (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Christian Ochsenfeld − *Chair of Theoretical Chemistry, Department of Chemistry, University of Munich (LMU), 81377 Munich, Germany*; orcid.org/0000-0002-4189-6558; Email: christian.ochsenfeld@uni-muenchen.de

Authors

Michael Glasbrenner − *Chair of Theoretical Chemistry, Department of Chemistry, University of Munich (LMU), 81377 Munich, Germany*

Daniel Graf − *Chair of Theoretical Chemistry, Department of Chemistry, University of Munich (LMU), 81377 Munich, Germany*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jctc.0c00600

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors acknowledge financial support by the "Deutsche Forschungsgemeinschaft" (DFG) via the SFB 1309-32587107 and the cluster of excellence (EXC2111-390814868) "Munich Center for Quantum Science and Technology" (MCQST). M.G. thanks the "Studienstiftung des Deutschen Volkes" for a graduate fellowship. C.O. acknowledges additional financial support as a Max-Planck-Fellow at MPI-FKF Stuttgart.

■ REFERENCES

(1) Møller, C.; Plesset, M. S. Note on an approximation treatment for many-electron systems. *Phys. Rev.* **1934**, *46*, 618.

(2) Cremer, D. Møller–Plesset perturbation theory: from small molecule methods to methods for thousands of atoms. *WIREs Comput. Mol. Sci.* **2011**, *1*, 509–530.

(3) 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*, 9095–9102.

(4) 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*, 9793–9802.

(5) Parr, R. G.; Yang, W. Density-functional theory of the electronic structure of molecules. *Annu. Rev. Phys. Chem.* **1995**, *46*, 701–728.

(6) Grimme, S. Semiempirical hybrid density functional with perturbative second-order correlation. *J. Chem. Phys.* **2006**, *124*, No. 034108.

(7) Goerigk, L.; Moellmann, J.; Grimme, S. Computation of accurate excitation energies for large organic molecules with double-hybrid density functionals. *Phys. Chem. Chem. Phys.* **2009**, *11*, 4611–4620.

(8) Goerigk, L.; Grimme, S. Double-hybrid density functionals. *WIREs Comput. Mol. Sci.* **2014**, *4*, 576–600.

(9) Pulay, P. Localizability of dynamic electron correlation. *Chem. Phys. Lett.* **1983**, *100*, 151–154.

(10) Pulay, P.; Saebø, S. Orbital-invariant formulation and second-order gradient evaluation in Møller–Plesset perturbation theory. *Theor. Chim. Acta* **1986**, *69*, 357–368.

(11) Häser, M. Møller–Plesset (MP2) perturbation theory for large molecules. *Theor. Chim. Acta* **1993**, *87*, 147–173.

(12) Martínez, T. J.; Carter, E. A. Pseudospectral Møller–Plesset perturbation theory through third order. *J. Chem. Phys.* **1994**, *100*, 3631–3638.

(13) Hampel, C.; Werner, H.-J. Local treatment of electron correlation in coupled cluster theory. *J. Chem. Phys.* **1996**, *104*, 6286–6297.

(14) Maslen, P.; Head-Gordon, M. Non-iterative local second order Møller–Plesset theory. *Chem. Phys. Lett.* **1998**, *283*, 102–108.

(15) Ayala, P. Y.; Scuseria, G. E. Linear scaling second-order Møller–Plesset theory in the atomic orbital basis for large molecular systems. *J. Chem. Phys.* **1999**, *110*, 3660–3671.

(16) Maslen, P. E.; Head-Gordon, M. Noniterative local second order Møller–Plesset theory: Convergence with local correlation space. *J. Chem. Phys.* **1998**, *109*, 7093–7099.

(17) Schütz, M.; Hetzer, G.; Werner, H.-J. Low-order scaling local electron correlation methods. I. Linear scaling local MP2. *J. Chem. Phys.* **1999**, *111*, 5691–5705.

(18) Lee, M. S.; Maslen, P. E.; Head-Gordon, M. Closely approximating second-order Møller–Plesset perturbation theory with a local triatomics in molecules model. *J. Chem. Phys.* **2000**, *112*, 3592–3601.

(19) Hetzer, G.; Schütz, M.; Stoll, H.; Werner, H.-J. 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*, 9443–9455.

(20) Saebø, S.; Pulay, P. A low-scaling method for second order Møller–Plesset calculations. *J. Chem. Phys.* **2001**, *115*, 3975–3983.

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

(22) Werner, H.-J.; Manby, F. R.; Knowles, P. J. Fast linear scaling second-order Møller–Plesset perturbation theory (MP2) using local and density fitting approximations. *J. Chem. Phys.* **2003**, *118*, 8149–8160.

(23) Lambrecht, D. S.; Doser, B.; Ochsenfeld, C. Rigorous integral screening for electron correlation methods. *J. Chem. Phys.* **2005**, *123*, 184102.

(24) Jung, Y.; Head-Gordon, M. A fast correlated electronic structure method for computing interaction energies of large van der Waals complexes applied to the fullerene-porphyrin dimer. *Phys. Chem. Chem. Phys.* **2006**, *8*, 2831–2840.

(25) 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.

(26) Doser, B.; Lambrecht, D. S.; Ochsenfeld, C. Tighter multipole-based integral estimates and parallel implementation of linear-scaling AO–MP2 theory. *Phys. Chem. Chem. Phys.* **2008**, *10*, 3335–3344.

(27) Doser, B.; Lambrecht, D. S.; Kussmann, J.; Ochsenfeld, C. Linear-scaling atomic orbital-based second-order Møller–Plesset perturbation theory by rigorous integral screening criteria. *J. Chem. Phys.* **2009**, *130*, No. 064107.

(28) Zienau, J.; Clin, L.; Doser, B.; Ochsenfeld, C. Cholesky-decomposed densities in Laplace-based second-order Møller–Plesset perturbation theory. *J. Chem. Phys.* **2009**, *130*, 204112.

(29) Kristensen, K.; Høyvik, I.-M.; Jansik, B.; Jørgensen, P.; Kjærgaard, T.; Reine, S.; Jakowski, J. MP2 energy and density for large molecular systems with internal error control using the Divide-Expand-Consolidate scheme. *Phys. Chem. Chem. Phys.* **2012**, *14*, 15706–15714.

(30) 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.

(31) Hohenstein, E. G.; Parrish, R. M.; Martínez, T. J. Tensor hypercontraction density fitting. I. Quartic scaling second-and third-order Møller–Plesset perturbation theory. *J. Chem. Phys.* **2012**, *137*, No. 044103.

(32) 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*, No. 014101.

(33) 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.

(34) Pinski, P.; Riplinger, C.; Valeev, E. F.; Neese, F. Sparse maps—A systematic infrastructure for reduced-scaling electronic structure methods. I. An efficient and simple linear scaling local MP2 method that uses an intermediate basis of pair natural orbitals. *J. Chem. Phys.* **2015**, *143*, No. 034108.

(35) Nagy, P. R.; Samu, G.; Kállay, M. An integral-direct linear-scaling second-order Møller–Plesset approach. *J. Chem. Theory Comput.* **2016**, *12*, 4897–4914.

(36) Baudin, P.; Ettenhuber, P.; Reine, S.; Kristensen, K.; Kjærgaard, T. Efficient linear-scaling second-order Møller–Plesset perturbation theory: The divide–expand–consolidate RI-MP2 model. *J. Chem. Phys.* **2016**, *144*, No. 054102.

(37) Pham, B. Q.; Gordon, M. S. Compressing the Four-Index Two-Electron Repulsion Integral Matrix using the Resolution-of-the-Identity Approximation Combined with the Rank Factorization Approximation. *J. Chem. Theory Comput.* **2019**, *15*, 2254–2264.

(38) Pham, B. Q.; Gordon, M. S. Hybrid distributed/shared memory model for the RI-MP2 method in the fragment molecular orbital framework. *J. Chem. Theory Comput.* **2019**, *15*, 5252–5258.

(39) Barca, G. M.; McKenzie, S. C.; Bloomfield, N. J.; Gilbert, A. T.; Gill, P. M. Q-MP2-OS: Møller–Plesset Correlation Energy by Quadrature. *J. Chem. Theory Comput.* **2020**, *16*, 1568–1577.

(40) 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.

(41) Parrish, R. M.; Hohenstein, E. G.; Martínez, T. J.; Sherrill, C. D. Tensor hypercontraction. II. Least-squares renormalization. *J. Chem. Phys.* **2012**, *137*, 224106.

(42) Riplinger, C.; Neese, F. An efficient and near linear scaling pair natural orbital based local coupled cluster method. *J. Chem. Phys.* **2013**, *138*, No. 034106.

(43) Riplinger, C.; Pinski, P.; Becker, U.; Valeev, E. F.; Neese, F. Sparse maps—A systematic infrastructure for reduced-scaling elec-

tronic structure methods. II. Linear scaling domain based pair natural orbital coupled cluster theory. *J. Chem. Phys.* **2016**, *144*, No. 024109.

(44) Almlöf, J. Elimination of energy denominators in Møller–Plesset perturbation theory by a Laplace transform approach. *Chem. Phys. Lett.* **1991**, *181*, 319–320.

(45) Häser, M.; Almlöf, J. Laplace transform techniques in Møller–Plesset perturbation theory. *J. Chem. Phys.* **1992**, *96*, 489–494.

(46) Jung, Y.; Sodt, A.; Gill, P. M.; Head-Gordon, M. Auxiliary basis expansions for large-scale electronic structure calculations. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 6692–6697.

(47) Häser, M.; Ahlrichs, R. Improvements on the direct SCF method. *J. Comput. Chem.* **1989**, *10*, 104–111.

(48) Thompson, T. H.; Ochsenfeld, C. Integral partition bounds for fast and effective screening of general one-, two-, and many-electron integrals. *J. Chem. Phys.* **2019**, *150*, No. 044101.

(49) Higham, N. J. Cholesky factorization. *Wiley Interdiscip. Rev. Comput. Stat.* **2009**, *1*, 251–254.

(50) 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.

(51) 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.

(52) Graf, D.; Beuerle, M.; Schurkus, H. F.; Luenser, A.; Savasci, G.; Ochsenfeld, C. Accurate and Efficient Parallel Implementation of an Effective Linear-Scaling Direct Random Phase Approximation Method. *J. Chem. Theory Comput.* **2018**, *14*, 2505–2515.

(53) McWeeny, R. Hartree-Fock theory with nonorthogonal basis functions. *Phys. Rev.* **1959**, *114*, 1528.

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

(55) Helmich-Paris, B.; Repisky, M.; Visscher, L. Relativistic Cholesky-decomposed density matrix MP2. *Chem. Phys.* **2019**, *518*, 38–46.

(56) Kussmann, J.; Ochsenfeld, C. Pre-selective screening for matrix elements in linear-scaling exact exchange calculations. *J. Chem. Phys.* **2013**, *138*, 134114.

(57) 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.

(58) Kussmann, J.; Ochsenfeld, C. Hybrid CPU/GPU integral engine for strong-scaling ab initio methods. *J. Chem. Theory Comput.* **2017**, *13*, 3153–3159.

(59) Weigend, F.; Ahlrichs, R. Balanced basis sets 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.

(60) 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.

(61) Pulay, P. Convergence acceleration of iterative sequences. The case of SCF iteration. *Chem. Phys. Lett.* **1980**, *73*, 393–398.

(62) Kaltak, M.; Klimeš, J.; Kresse, G. Low scaling algorithms for the random phase approximation: Imaginary time and Laplace transformations. *J. Chem. Theory Comput.* **2014**, *10*, 2498–2507.

(63) 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*, No. 044112.

(64) Sedlak, R.; Janowski, T.; Pitoňák, M.; Řezáč, J.; Pulay, P.; Hobza, P. Accuracy of Quantum Chemical Methods for Large Noncovalent Complexes. *J. Chem. Theory Comput.* **2013**, *9*, 3364–3374.

(65) 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.

(66) Jurečka, P.; Šponer, J.; Černý, J.; Hobza, P. 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*, 1985–1993.

(67) 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.

Supplementary Information for: Efficient Reduced-Scaling Second-Order Møller-Plesset Perturbation Theory with Cholesky-Decomposed Densities and an Attenuated Coulomb Metric

Michael Glasbrenner, Daniel Graf, and Christian Ochsenfeld

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

October 12, 2020

In Tab. 1, the results of benchmark calculations on the S22 test set are shown. Note that two different references are used. All calculations with ω-RI-CDD-MP2 employ RI with an attenuated Coulomb metric and an ω-value of 0.1 and furthermore employ a Laplace-transformation. The reference denoted as ω-RI-CDD-MP2$_{\text{exact}}$ also contains these two sources of error, but no additional approximations. The deviations from ω-RI-CDD-MP2$_{\text{exact}}$ thus show how large the errors due to integral screening and natural blocking are.

For all integral screening and natural blocking thresholds considered in Tab. 1, fast convergence towards the ω-RI-CDD-MP2$_{\text{exact}}$ reference is observed if the thresholds are lowered. The deviations from canonical RI-MP2, which are also shown in Tab. 1, are significantly larger and converge to a constant value upon lowering the thresholds. The remaining deviations are mostly caused by the use of an attenuated Coulomb metric instead of a standard Coulomb metric for the RI.

Table 1: Benchmark calculations on all monomers and dimers from the S22 test set for all ϑ-thresholds. Mean absolute viations (MAD) and maximum absolute deviations (MAX) are shown. The errors are computed either relative to canonical RI-MP2 or to ω-RI-CDD-MP2$_{\text{exact}}$, which refers to the ω-RI-CDD-MP2 method with Laplace transformation and attenuated Coulomb metric ($\omega = 0.1$) as the only approximations. 7 Laplace-points are used for $[S_8]_5$; 8 Laplace-points in all other calculations. The def2-SVP basis set is employed in all calculations. All values are given in kcal/mol.

		Error vs. ω-RI-CDD-MP2$_{\text{exact}}$		Error vs. RI-MP2	
threshold	value	MAD	MAX	MAD	MAX
$\vartheta_{3\text{c}}$	10^{-5}	1.17×10^{-5}	5.22×10^{-5}	1.01×10^{-2}	2.59×10^{-2}
	10^{-6}	1.10×10^{-6}	6.84×10^{-6}	1.01×10^{-2}	2.58×10^{-2}
	10^{-7}	1.06×10^{-7}	6.28×10^{-7}	1.01×10^{-2}	2.58×10^{-2}
	10^{-8}	2.66×10^{-8}	4.39×10^{-7}	1.01×10^{-2}	2.58×10^{-2}
$\vartheta_{\text{schwarz}}$	10^{-8}	8.64×10^{-3}	7.22×10^{-2}	1.14×10^{-2}	5.54×10^{-2}
	10^{-9}	5.58×10^{-4}	5.00×10^{-3}	9.61×10^{-3}	2.57×10^{-2}
	10^{-10}	3.51×10^{-5}	3.58×10^{-4}	1.00×10^{-2}	2.58×10^{-2}
	10^{-11}	2.38×10^{-6}	2.09×10^{-5}	1.01×10^{-2}	2.58×10^{-2}
$\vartheta_{\underline{ij}}$	10^{-4}	7.75×10^{-2}	3.14×10^{-1}	6.78×10^{-2}	2.98×10^{-1}
	10^{-5}	7.46×10^{-3}	3.08×10^{-2}	6.14×10^{-3}	1.99×10^{-2}
	10^{-6}	6.16×10^{-4}	2.75×10^{-3}	9.54×10^{-3}	2.53×10^{-2}
	10^{-7}	6.03×10^{-5}	2.92×10^{-4}	1.00×10^{-2}	2.58×10^{-2}
ϑ_{NB}	10^{-4}	4.68×10^{-2}	5.75×10^{-1}	4.70×10^{-2}	5.66×10^{-1}
	10^{-5}	1.21×10^{-5}	3.30×10^{-4}	1.00×10^{-2}	2.58×10^{-2}
	10^{-6}	8.56×10^{-8}	1.82×10^{-6}	1.01×10^{-2}	2.58×10^{-2}
	10^{-7}	3.04×10^{-8}	6.90×10^{-7}	1.01×10^{-2}	2.58×10^{-2}
$\vartheta_{\text{NB}}^{iaP}$	10^{-4}	1.54×10^{-2}	5.72×10^{-1}	2.06×10^{-2}	5.63×10^{-1}
	10^{-5}	8.37×10^{-6}	2.88×10^{-4}	1.01×10^{-2}	2.58×10^{-2}
	10^{-6}	4.56×10^{-8}	1.07×10^{-6}	1.01×10^{-2}	2.58×10^{-2}
	10^{-7}	3.14×10^{-8}	4.39×10^{-7}	1.01×10^{-2}	2.58×10^{-2}

3.2 Low-Scaling Tensor Hypercontraction in the Cholesky Molecular Orbital Basis Applied to Second-Order Møller–Plesset Perturbation Theory

F.H. Bangerter, M. Glasbrenner, C. Ochsenfeld.
J. Chem. Theory Comput. **17**, 211 (2021).

Abstract: We employ various reduced scaling techniques to accelerate the recently developed least-squares tensor hypercontraction (LS-THC) approximation [Parrish, R. M., Hohenstein, E. G., Martínez, T. J., Sherrill, C. D. *J. Chem. Phys. 137*, 224106 (**2012**)] for electron repulsion integrals (ERIs) and apply it to second-order Møller–Plesset perturbation theory (MP2). The grid-projected ERI tensors are efficiently constructed using a localized Cholesky molecular orbital basis from density-fitted integrals with an attenuated Coulomb metric. Additionally, rigorous integral screening and the natural blocking matrix format are applied to reduce the complexity of this step. By recasting the equations to form the quantized representation of the 1/r operator Z into the form of a system of linear equations, the bottleneck of inverting the grid metric via pseudoinversion is removed. This leads to a reduced scaling THC algorithm and application to MP2 yields the (sub-)quadratically scaling THC-ω-RI-CDD-SOS-MP2 method. The efficiency of this method is assessed for various systems including DNA fragments with over 8000 basis functions and the subquadratic scaling is illustrated.

Reprinted with permission from:

F.H. Bangerter, M. Glasbrenner, C. Ochsenfeld.
“Low-Scaling Tensor Hypercontraction in the Cholesky Molecular Orbital Basis Applied to Second-Order Møller–Plesset Perturbation Theory.”
J. Chem. Theory Comput. **17**, 211 (2021).

`https://pubs.acs.org/doi/abs/10.1021/acs.jctc.0c00934`

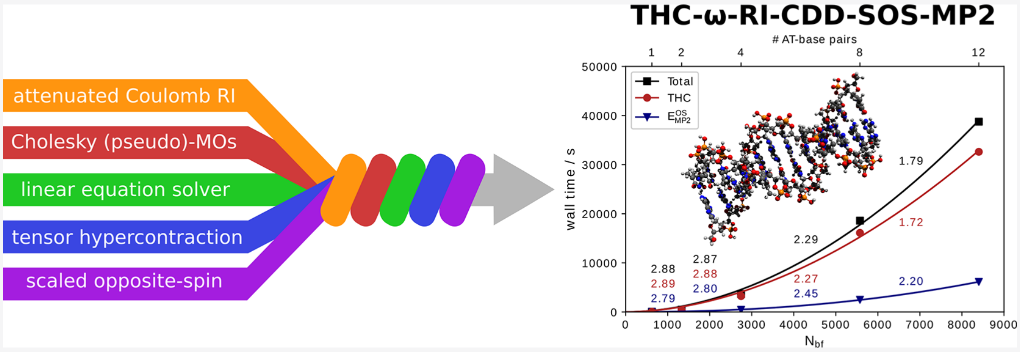
JCTC Journal of Chemical Theory and Computation

pubs.acs.org/JCTC Article

Low-Scaling Tensor Hypercontraction in the Cholesky Molecular Orbital Basis Applied to Second-Order Møller–Plesset Perturbation Theory

Felix H. Bangerter, Michael Glasbrenner, and Christian Ochsenfeld*

Cite This: *J. Chem. Theory Comput.* 2021, 17, 211−221 Read Online



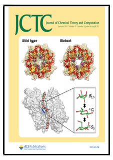
ABSTRACT: We employ various reduced scaling techniques to accelerate the recently developed least-squares tensor hypercontraction (LS-THC) approximation [Parrish, R. M., Hohenstein, E. G., Martínez, T. J., Sherrill, C. D. *J. Chem. Phys.* *137*, 224106 (**2012**)] for electron repulsion integrals (ERIs) and apply it to second-order Møller−Plesset perturbation theory (MP2). The grid-projected ERI tensors are efficiently constructed using a localized Cholesky molecular orbital basis from density-fitted integrals with an attenuated Coulomb metric. Additionally, rigorous integral screening and the natural blocking matrix format are applied to reduce the complexity of this step. By recasting the equations to form the quantized representation of the $1/r$ operator $\mathbf{Z}$ into the form of a system of linear equations, the bottleneck of inverting the grid metric via pseudoinversion is removed. This leads to a reduced scaling THC algorithm and application to MP2 yields the (sub-)quadratically scaling THC-ω-RI-CDD-SOS-MP2 method. The efficiency of this method is assessed for various systems including DNA fragments with over 8000 basis functions and the subquadratic scaling is illustrated.

1. INTRODUCTION

One of the simplest quantum chemical methods to recapture the electron correlation neglected by Hartree−Fock (HF) theory is second-order Møller−Plesset perturbation theory (MP2),[1] which has been extensively studied in the past decades.[2] Today, it remains popular as a post-HF method with affordable cost and also as part of double-hybrid density functionals (DHDF).[3,4] It has further proven to provide accurate results for molecular properties such as NMR chemical shifts.[5] The computational cost of conventional MP2 is governed by the transformation of the electron repulsion integrals (ERI) from atomic orbital (AO) to molecular orbital (MO) basis, which scales as $O(N^5)$ with the number of basis functions N. The formal fifth power scaling of the computational cost and the $O(N^4)$ storage requirements severely restrict the size of computationally accessible molecules. Naturally, considerable effort has been put into reducing the computational cost of MP2 calculations. Methods aiming at reducing the scaling behavior of MP2 can be grouped into two different categories: first, methods aiming at exploiting the sparsity of the MP2 problem, either of intrinsic nature such as in the AO-MP2 formulation[6−8] or created by ansatz as in local MP2 methods[9−18] and, second, methods based on tensor decompositions of the ERIs involved.[19−27]

The pioneering work of Pulay and Sæbø[9−11] on the local treatment of correlation in the 1980s served as a stepping stone for the development of the family of local correlation (LC) methods by Werner, Schütz, and others.[12−18] In the LC methods developed by Werner and Schütz,[12−15,17] the occupied space is spanned by localized molecular orbitals (LMOs), which are obtained from the canonical orbitals by localization

Received: September 8, 2020
Published: December 29, 2020



https://dx.doi.org/10.1021/acs.jctc.0c00934

techniques, such as Foster−Boys[28] or Pipek−Mezey,[29] while the virtual space is spanned by nonorthogonal projected AOs (PAOs), which are obtained from the AO basis by projecting out the occupied orbitals. Using the locality created by a standard localization procedure for the occupied MOs and the inherent locality of the PAOs, low- and linear-scaling algorithms can be devised. The latter has been demonstrated for MP2,[13,17] the singles and doubles coupled cluster method (CCSD),[12,14] and even for additional perturbative triple excitations (CCSD(T)).[15,30] In order to overcome the sometimes large PAO domains needed to converge the correlation energy, Neese *et al.*[31] proposed a compression of the virtual space by means of pair natural orbitals (PNOs). In combination with careful prescreening and a multipole expansion for distant orbital pairs, Werner *et al.*[32] developed the PNO-LMP2 method that scales as $\mathcal{O}(M)$ with respect to the molecule size M and has good parallel efficiency. Conceptually similar is the domain-based local PNO (DLPNO) implementation of local MP2 by Pinski *et al.*,[33] which was extended to a linear scaling implementation of the DLPNO-CCSD(T) method.[34−36] An alternative approach to this is the divide-expand-consolidate (DEC) group of methods,[34−36] which achieve massive parallelization[37] and therefore allow the computation of very large molecules by partitioning of the orbital space.

In spite of their favorable scaling and efficiency, local correlation methods suffer from distinct drawbacks. One of the most severe disadvantages is that the domains may change as a function of the molecular geometry, which can reduce the accuracy and may even lead to discontinuities in the potential energy surface.[38,39] Thus, instead of localizing the orbital space, a second approach to reduce the scaling of the MP2 algorithm is to recast the MP2 equations into the AO basis. Making use of the Laplace transformation for the orbital energy denominator occurring in the canonical MP2 formulation introduced by Almlöf and Häser,[6,40,41] in combination with efficient integral screening, our group developed an asymptotically linear scaling AO-MP2 method.[8] To overcome the large prefactor associated with the AO-based formulation, the RI-CDD-MP2 method[42] was introduced, which uses a pivoted Cholesky decomposition of the pseudo-density matrices (CDD) to obtain a localized Cholesky pseudo-MO basis. RI-CDD-MP2 significantly reduces the prefactor of the MP2 calculation while scaling asymptotically cubic if no further approximations are applied.[42]

In contrast to the above, methods aiming at decomposing the tensors involved in the calculations usually do not lower the scaling exponent of the method they are applied to, but rather lower the prefactor of the respective method. The most established approach in the family of tensor decomposition methods is the density fitting or resolution-of-the-identity ansatz,[19,43−45] which approximates a generalized charge density $\rho_{\mu\nu}(\mathbf{r}_1) = \chi_\mu(\mathbf{r}_1)\chi_\nu(\mathbf{r}_1)$ with an auxiliary basis $\{\chi_\alpha(\mathbf{r}_1)\}$. Closely related is the Cholesky decomposition (CD) approach[20−22,46−50] of the ERIs, in which the auxiliary basis is constructed on-the-fly as an orthogonalized subset of the basis function space $\{\chi_\mu\}$. Both RI and CD achieve a factorization of the fourth-order ERI tensor into two third-order tensors. Today, RI is routinely applied to Møller−Plesset perturbation theory,[51−53] density functional theory (DFT),[54,55] the random phase approximation (RPA),[56,57] or coupled cluster (CC)[58] theory. Further decomposition of the ERIs into two second-order and one third-order tensor is possible within the pseudospectral approach,[23−25] in which the integral over one electronic coordinate is solved analytically and the other by numerical quadrature. Conceptually related are the chain-of-spheres approach (COSX)[59] and the seminumerical exact-exchange formalism (sn-LinK)[60] that make use of the same kind of decomposition. Taking this approach one step further and attempting a two-sided decomposition of the ERI, one arrives at the recently developed tensor hypercontraction (THC)[27,61−63] factorization. Within the THC framework, it is possible to approximate the fourth-order ERI tensor as a product of just five factor matrices, reducing the highest order tensor necessary to represent the full ERI to only two. While different flavors of THC have been proposed, like parallel factors THC (PF-THC),[26] which achieves the factorization by canonical polyadic decomposition (CPD) of the three-center overlap integrals within the overlap metric RI formalism, we focus our discussion on the more accurate least-squares THC (LS-THC)[27] approach. In passing, we note that the atomic-batched tensor decomposition by Schmitz *et al.*[64,65] can also be viewed as a THC-like factorization starting from RI-approximated ERIs with further decomposition by singular value decomposition of the third-order subtensors belonging to atom pairs. THC attempts a quantization of the spatial coordinates via $\mathbf{r}_1 \rightarrow \{\mathbf{r}_P\}$ and $\mathbf{r}_2 \rightarrow \{\mathbf{r}_Q\}$ and replacing the singular grid operator $r_{12}^{-1} \rightarrow \{r_{PQ}^{-1}\}$ with a renormalized operator Z^{PQ}. Together with the RI approximation, LS-THC has proven to be a flexible and efficient framework and has been successfully applied to exact exchange,[66] MP2,[26,27,62,63,66−69] MP3,[26,66] CCSD,[70] second-order approximated CC (CC2),[71] as well as equation-of-motion CC2 (EOM-CC2),[72] RPA,[73] and second-order complete active space perturbation theory (CASPT2).[74]

In this work, we aim to combine different linear scaling techniques to obtain an efficient and low-scaling THC algorithm. In particular, we extend the RI-LS-THC formalism developed by Martínez and co-workers[27,61−63] using an attenuated Coulomb metric[75] for the THC fitting procedure. Together with a localized Cholesky-MO basis, we seek to efficiently exploit the sparsity in the three-center integrals for the subsequent contractions. In our integral code, we make use of efficient integral screening based on the recently developed integral partition bounds (IPB).[76] Furthermore, we employ the natural blocking scheme originally developed by Jung *et al.*[77] for the transformation to the Cholesky-MO basis and subsequent projection onto the quadrature grid. In addition to thereby lowering the scaling of the formally most expensive step, i.e., the formation of the grid-projected ERI tensor, we present ways to circumvent the computational bottleneck of forming the pseudoinverse of the grid metric necessary for building the quantized and renormalized grid-projected potential operator. We do this by decomposition of the grid metric matrix followed by a forward and a backward substitution step. This is in contrast to the local-THC approximation by Song *et al.*[63] in which the molecule is partitioned into fragments to reduce the complexity of the inversion. After contraction of one-half of the projected potential operator, obtained by the linear solve step, with its transpose and the two-center RI integrals, the familiar $\mathbf{Z}$ tensor is obtained. Finally, we apply our efficient THC algorithm to MP2, for which the energy evaluation is performed in a localized Cholesky pseudo-MO basis using block-sparse linear algebra to efficiently exploit sparsity and reduce the scaling of this step. Together with the scaled opposite-spin (SOS)[78] approximation, we present the (sub-)quadratically scaling THC-ω-RI-CDD-SOS-MP2 method.

2. THEORY

2.1. Notation. We briefly summarize the relevant indices below:

- $\mu, \nu, \lambda, \sigma$: atomic orbital indices belonging to the AO basis $\{\chi_\mu\}$ of size N_{bf}.
- $\alpha, \beta, \gamma, \delta$: auxiliary basis indices belonging to the density fitting basis $\{\chi_\alpha\}$ of size N_{aux} (usually $N_{\mathrm{aux}} \approx 3 \cdot N_{\mathrm{bf}}$).
- P, Q, R, S: grid point indices belonging to the LS-THC grid of size N_{grid} (usually $N_{\mathrm{grid}} \approx 3 \cdot N_{\mathrm{aux}}$).
- i, j, k: occupied molecular orbital indices belonging to the MO basis $\{\phi_i\}$ of size N_{occ}.
- a, b, c: virtual molecular orbital indices belonging to the MO basis $\{\phi_a\}$ of size N_{virt} ($N_{\mathrm{virt}} \gg N_{\mathrm{occ}}$).
- κ: index of the Laplace quadrature points for the MP2 energy denominator (usually integration with 5−8 points is sufficiently accurate).

2.2. Review of Tensor Hypercontraction. The LS-THC estimator[27] for approximating the ERIs in the occupied-virtual (OV) subspace of the MO integrals using a physical-space quadrature is defined as

$$\underset{Z}{\arg\min} \sum_{ijab} \frac{1}{2} \left\| \Delta_{iajb} \right\|^2 \tag{1}$$

where the deviation Δ_{iajb} of the THC factorization from the exact ERI tensor is given by

$$\Delta_{iajb} = (ia|jb) - \sum_{PQ} X_i^P X_a^P Z^{PQ} X_j^Q X_b^Q \tag{2}$$

Note that we perform the THC fitting not in the AO space, but instead in the OV subspace of the MO space, since this is the only block of the MO integrals needed for evaluating the MP2 equation. Additionally, since the fitting space is smaller, grids with fewer grid points can be used, or, if the same grids are applied, more accurate results can be expected.[27,62,66] In contrast to the AO-based local-THC approximation by Song *et al.*,[63] we make use of a global description of the THC factorization, i.e., using a single global **Z** tensor, that is not prone to discontinuities of the potential energy surface. The latter is demonstrated in the Supporting Information for the rotation around the 4C−5C bond in vitamin K$_2$ as a representative example.

As is already evident from eqs 1 and 2, THC provides a more flexible tensor decomposition than all methods discussed above as it effectively unpins all four MO indices. A direct consequence of this is that fewer indices need to be carried over during tensor contractions, providing means for the formulation of lower-scaling algorithms. As shown by Parrish *et al.*,[27] differentiating eq 1 with respect to **Z** and subsequently solving the stationary condition for **Z** gives the analytic solution

$$Z^{PQ} = \sum_{P'Q'} [\mathbf{S}^{-1}]^{PP'} E^{P'Q'} [\mathbf{S}^{-1}]^{QQ'} \tag{3}$$

where the grid metric **S** is defined as

$$S^{PP'} = \sum_{i'a'} X_{i'}^P X_{i'}^{P'} X_{a'}^P X_{a'}^{P'} \tag{4}$$

and the grid-projected ERI tensor **E** is defined as

$$E^{P'Q'} = \sum_{ijab} X_i^{P'} X_a^{P'} (ia|jb) X_j^{Q'} X_b^{Q'} \tag{5}$$

Within the LS-THC framework,[27] the **X** tensors are chosen to be collocation matrices, i.e., results of evaluating the basis functions $\{\chi_\mu\}$ at the real-space grid $\{\mathbf{r}_P\}$ given by

$$X'^P_\mu = \sqrt[4]{w_P}\,\chi_\mu(\mathbf{r}_P) \tag{6}$$

where w_P are the associated grid weights. For reasons of numerical stability, we found it useful to additionally balance the collocation matrices **X** according to

$$X_\mu^P = v_P X'^P_\mu \quad \text{with} \quad v_P = \frac{1}{\sqrt{\sum_\mu X'^P_\mu X'^P_\mu}} \tag{7}$$

as proposed by Parrish *et al.*[61] After rebalancing, the collocation matrices are transformed into Cholesky-MO space[79] according to

$$X_i^P = \sum_\mu L_{\mu i} X_\mu^P \tag{8}$$

$$X_a^P = \sum_\nu L_{\nu a} X_\nu^P \tag{9}$$

where $L_{\mu i}$ and $L_{\nu a}$ are the local Cholesky factors obtained from the occupied density matrix **P** and the virtual density matrix **Q**, respectively. We note that here and throughout CDD refers to the Cholesky decomposition of the ground state densities as well as the pseudo-densities in the energy evaluation (eq 30). Since **P** and **Q** are both invariant under orbital rotations, an infinite number of MO coefficient representations of the density matrix exists.[79] The Cholesky factors thus represent one possible set of MO coefficients with the advantage that the Cholesky-MOs are local and the coefficients sparse, which can be exploited during subsequent contractions.

2.3. Construction of the Grid-Projected ERI Tensor E. Inserting the RI factorization of the ERI tensor into eq 5 allows one to reduce the formal scaling of the construction of the **E** tensor to only quartic or $O(N_{\mathrm{occ}} N_{\mathrm{virt}} N_{\mathrm{aux}} N_{\mathrm{grid}})$ to be more precise.[27] Thus, since all other operations in obtaining the THC factorization scale at most cubic, the formation of **E** will still be the rate-determining step in the asymptotic limit. To circumvent this issue in the THC algorithm, we substitute the ERI tensor by a double density-fitted RI approximation with a complementary error function (erfc-) attenuated Coulomb metric[75]

$$(\mu\nu|\lambda\sigma) = \sum_{\alpha\beta\gamma\delta} (\mu\nu:\alpha)[\mathbf{V}^{-1}]_{\alpha\gamma}[\mathbf{J}]_{\gamma\delta}[\mathbf{V}^{-1}]_{\delta\beta}(\beta:\lambda\sigma) \tag{10}$$

where the two- and three-center integrals with the erfc-attenuated Coulomb operator are given by

$$(\mu\nu:\alpha) = \iint_{\mathbb{R}^3} \mathrm{d}\mathbf{r}_1\,\mathrm{d}\mathbf{r}_2\,\chi_\mu(\mathbf{r}_1)\chi_\nu(\mathbf{r}_1)\frac{\mathrm{erfc}(\omega r_{12})}{r_{12}}\chi_\alpha(\mathbf{r}_2) \tag{11}$$

$$V_{\alpha\gamma} = \iint_{\mathbb{R}^3} \mathrm{d}\mathbf{r}_1\,\mathrm{d}\mathbf{r}_2\,\chi_\alpha(\mathbf{r}_1)\frac{\mathrm{erfc}(\omega r_{12})}{r_{12}}\chi_\gamma(\mathbf{r}_2) \tag{12}$$

and $[\mathbf{J}]_{\gamma\delta}$ is the usual two-center Coulomb integral of the auxiliary functions. In the limits of $\lim_{\omega\to 0}$ and $\lim_{\omega\to\infty}$, the attenuated Coulomb operator reduces to the Coulomb metric $1/r_{12}$ or the overlap metric $\delta(r_{12})$, respectively. Tuning the attenuation parameter ω thus provides a way to interpolate between the more accurate Coulomb metric and the overlap metric, which provides greater sparsity in the corresponding integrals. An attenuation strength of $\omega = 0.1$ was shown to yield acceptable accuracy while providing sparsity comparable to the overlap metric.[56] Due to the locality of the erfc-attenuated three-center integrals, their computation can be reduced to linear scaling by applying screening based on the recently developed rigorous integral partition bounds (IPB).[76] For the THC fitting in the Cholesky-MO basis, the three-center AO integrals have first to be transformed to the OV subspace according to

$$(ia \,\vdots\, \alpha) = \sum_{\mu\nu} L_{\mu i} L_{\nu a} (\mu\nu \,\vdots\, \alpha) \tag{13}$$

where $L_{\mu i}$ and $L_{\nu a}$ are again the local Cholesky factors obtained from the occupied density matrix $\mathbf{P}$ and the virtual density matrix $\mathbf{Q}$, respectively. The contraction is carried out stepwise and to reduce the computational complexity of these two steps we make use of the natural blocking matrix format initially proposed by Jung *et al.*,[77] as recently reported in our work on reduced-scaling ω-RI-CDD-MP2.[80] In short, third-order tensors are usually internally stored as arrays of matrices, so-called tensor slices, which can in turn be stored in any sparse matrix format. Within the natural blocking approach,[77] entire rows and columns of these slices are removed if all their values fall below a certain threshold, the so-called natural blocking threshold $\varepsilon_{\mathrm{NB}}$. For the natural blocking format to be efficient, only a constant number of rows and columns of a given slice should be significant. The latter is given using an attenuated Coulomb operator causing coupling over only short distances between the bra and ket and by transformation into the localized Cholesky-MO basis, which reintroduces coupling between the bra indices. To keep track of the significant rows and columns, natural blocking relies on bookkeeping in terms of significance lists. Adopting the notation introduced in our work on ω-RI-CDD-MP2,[80] an example would be the list $\{\mu\}_\alpha$, which is defined as

$$\{\mu\}_\alpha \equiv \{\mu | \max_\nu |(\mu\nu \,\vdots\, \alpha)| > \varepsilon_{\mathrm{NB}}\} \tag{14}$$

and stores all significant AO indices μ for a given auxiliary function α. A second type of lists, which goes beyond the original natural blocking approach introduced by Jung *et al.*,[77] is based on integral screening prior to the AO-to-MO transformation.[80] For this, two matrices $\mathbf{M}$ and $\mathbf{N}$, defined as

$$M_{\mu\nu} = \max_\alpha |(\mu\nu \,\vdots\, \alpha)| \tag{15}$$

$$N_{\nu\alpha} = \max_\mu |(\mu\nu \,\vdots\, \alpha)| \tag{16}$$

are constructed in the three-center integral kernel. With these screening matrices, upper bounds for the transformed integrals can be constructed. A detailed algorithm for the transformation of the three-center integrals to the Cholesky-MO basis is provided in the Supporting Information. Here, we rather focus on the construction of intermediate $\mathbf{Y}$, defined as

$$Y_\alpha^{P'} = \sum_{ia} X_i^{P'} X_a^{P'} (ia \,\vdots\, \alpha) \tag{17}$$

obtained by inserting the ω-RI ansatz from eq 10 into the definition of the $\mathbf{E}$ tensor from eq 5. One advantage of performing the THC fitting in the Cholesky-MO space is that the prefactor of the contraction above will be reduced compared to a formulation in AOs as one occupied index is present. The projection of the RI integrals onto the THC grids then follows the same general procedure as their transformation into Cholesky-MO space. The significance list $[P]_i$, meaning the grid points significant for a given occupied orbital can readily be obtained from the Cholesky-MO-transformed collocation matrix according to

$$[P]_i \equiv \{P | |X_i^P| > \varepsilon_{\mathrm{NB}}\} \tag{18}$$

Due to the local nature of the Cholesky-MOs, for systems with significant bandgaps, the number of significant grid points can be expected to be small and constant for large enough molecules. Using the lists $\{a\}_i$ and $\{\alpha\}_i$, obtained during the integral transformation analogously to eq 14 and the upper bound list $[P]_i$, all tensors from eq 17 will have reduced sizes. Due to the coupling between all integral indices and the locality of the grid-projected Cholesky-MOs, the algorithm shown below for the evaluation of $Y_\alpha^{P'}$ can be expected to be asymptotically linear scaling.

Algorithm 1 ω-RI algorithm for $Y_\alpha^{P'}$

```
procedure BUILD_Y
    $\mathbf{Y} \leftarrow \mathbf{0}_{N_{\mathrm{grid}},N_{\mathrm{aux}}}$
    for all $i$ do
        for all $\alpha \in \{\alpha\}_i$ and $P' \in [P]_i$ do
            $(iP' \,\vdots\, \alpha) \leftarrow \sum_{a \in \{a\}_i} X_a^{P'} (ia \,\vdots\, \alpha)$
            $Y_\alpha^{P'} \mathrel{+}= X_i^{P'} (iP' \,\vdots\, \alpha)$
        end for
    end for
    return Y
end procedure
```

Here we note, that the grid-projected integral tensor $\mathbf{E}$ can be obtained from $\mathbf{Y}$ according to

$$E^{P'Q'} = \sum_{\alpha\beta} Y_\alpha^{P'} \tilde{C}_{\alpha\beta} Y_\beta^{Q'} \tag{19}$$

where $\tilde{C}_{\alpha\beta}$ is the contraction of all two-center RI integrals in eq 10. However, retaining this factorization allows for a more efficient reformulation of obtaining the $\mathbf{Z}$ tensor, as will be discussed below.

2.4. Removing the Pseudoinversion Bottleneck.

Despite the formal quartic scaling computation of the $\mathbf{E}$ tensor, Martínez and co-workers[27,62,63] reported the inversion of the grid metric $\mathbf{S}$ to be the bottleneck of the THC procedure. Since the metric tensor is generally ill-conditioned and singular, pseudoinversion with a cutoff $\varepsilon_{\mathrm{SVD}}$ between 10^{-10} to 10^{-12} is applied to remove the near-linear dependencies.[27,66] Due to the large prefactor of the underlying singular value decomposition (SVD), forming the inverse $\mathbf{S}^{-1}$ is usually the most time-consuming step. Using a local ansatz of the THC factorization[63] based on partitioning of the molecular system was presented as one way of avoiding this.

Here, we present an alternative for which we make use of our factorization of the $\mathbf{E}$ tensor and rewrite eq 3:

$$Z^{PQ} = \sum_{P'Q'} \sum_{\alpha\beta} [\mathbf{S}^{-1}]^{PP'} Y_\alpha^{P'} \tilde{C}_{\alpha\beta} Y_\beta^{Q'} [\mathbf{S}^{-1}]^{QQ'} \tag{20}$$

Taking only the tensors for the P' grid index of eq 20, we define a new intermediate $\mathbf{\Lambda}$, given by

$$\Lambda_\alpha^P \equiv \sum_{P'} [\mathbf{S}^{-1}]^{PP'} Y_\alpha^{P'} \Leftrightarrow Y_\alpha^{P'} = \sum_P S^{PP'} \Lambda_\alpha^P \tag{21}$$

As becomes evident from the right side of eq 21, instead of solving for $\mathbf{\Lambda}$ by inversion of the metric, the equation can also be treated as a system of linear equations with multiple right-hand sides, i.e., $\mathbf{B} = \mathbf{AX}$. Standard procedures to solving linear systems can then be applied, which typically involve a factorization of the coefficient matrix, here $\mathbf{S}$, followed by forward and backward substitution. Since the grid metric is generally singular and thus not symmetric positive definite, Cholesky decomposition cannot be applied for the factorization. However, other matrix decompositions like the modified blocked LDL factorization with diagonal pivoting or LU decomposition with partial pivoting are applicable. Inserting, for example, the pivoted LU decomposition $\mathbf{S} = \mathbf{\Pi LU}$ into eq 21 gives the following triangular systems of linear equations

$$\mathbf{\Pi}^\mathrm{T}\mathbf{Y} = \mathbf{LW} \triangleright \text{solve for } \mathbf{W} \tag{22}$$

$$\mathbf{W} = \mathbf{U\Lambda} \triangleright \text{solve for } \mathbf{\Lambda} \tag{23}$$

which can be solved efficiently by forward and backward substitution, respectively. We note that this procedure was recently also suggested by Matthews,[69] although in a slightly different context, i.e., for the automatic grid generation in LS-THC using a pivoted Cholesky decomposition of the grid metric constructed from a larger parent grid.

The asymptotic scaling of finding the grid metric inverse remains $O(N^3)$, albeit with a significantly lower prefactor, which makes this step irrelevant for the overall scaling for all but the largest molecules. It should be emphasized that applying linear solvers instead of pseudoinversion removes the bias of choosing a cutoff parameter ε_SVD while also providing numerical stability comparable to pseudoinversion. For large enough molecules with a significant HOMO−LUMO gap, the metric $\mathbf{S}$ will become sparse and diagonally dominant with appropriate reordering, which allows for the use of specialized algorithms for banded matrices. The latter allows us to reduce the scaling of this step to quadratic, as will be discussed in Section 3.

Having computed one side of the projected potential operator $\mathbf{\Lambda}$, the familiar $\mathbf{Z}$ tensor is obtained by contraction with its transpose and the two-center RI integrals $\tilde{\mathbf{C}}$ according to

$$Z^{PQ} = \sum_{\alpha\beta} \Lambda_\alpha^P \tilde{C}_{\alpha\beta} \Lambda_\beta^Q \tag{24}$$

After the construction of the $\mathbf{X}$ and $\mathbf{Z}$ tensors, they can be used to approximate ERIs in various correlation methods.

2.5. Evaluation of the THC-CDD-MP2 Energy.

The THC-CDD-MP2 equations[26] are based on the AO formulation of the MP2 energy, as originally proposed by Almlöf and Häser,[6,40,41] which makes use of the Laplace transformation for the energy denominator

$$E_\mathrm{MP2}^\mathrm{AO} = -\sum_\kappa \sum_{\mu\nu\lambda\sigma} \sum_{\mu'\nu'\lambda'\sigma'} \underline{P}_{\mu\mu'} \overline{P}_{\nu\nu'} \underline{P}_{\lambda\lambda'} \overline{P}_{\sigma\sigma'} (\mu\nu|\lambda\sigma) \cdot [2(\mu'\nu'|\lambda'\sigma') - (\mu'\sigma'|\lambda'\nu')] \tag{25}$$

where $\underline{\mathbf{P}}$ and $\overline{\mathbf{P}}$ are the occupied and virtual pseudo-density matrices, respectively. It should be noted that the Laplace transformation is especially important for the THC format as it allows to unpin all four orbital indices and thus to make full use of the THC factorization. The pseudo-densities, defined as

$$\underline{P}_{\mu\mu'} = (\omega^{(\kappa)})^{1/4} \sum_i C_{\mu i}\, \mathrm{e}^{\varepsilon_i t^{(\kappa)}} C_{\mu' i} \tag{26}$$

$$\overline{P}_{\nu\nu'} = (\omega^{(\kappa)})^{1/4} \sum_a C_{\nu a}\, \mathrm{e}^{-\varepsilon_a t^{(\kappa)}} C_{\nu' a} \tag{27}$$

are contractions of the MO coefficients and the orbital energies at a given Laplace point κ. Inserting the THC factorization of the ERIs into eq 25 and Cholesky decomposing the pseudo-densities, as is done in CDD-MP2,[42] i.e.,

$$\underline{P}_{\mu\mu'} = \sum_i \underline{L}_{\mu i} \underline{L}_{\mu' i} \tag{28}$$

$$\overline{P}_{\nu\nu'} = \sum_a \overline{L}_{\nu a} \overline{L}_{\nu' a} \tag{29}$$

the THC-CDD-MP2 energy equation is obtained

$$E_\mathrm{MP2}^\mathrm{THC\text{-}CDD} = -\sum_\kappa \sum_{PQRS} \sum_{ijab} \underline{X}_i^P \overline{X}_a^P Z^{PQ} \underline{X}_j^Q \overline{X}_b^Q \cdot [2\underline{X}_i^R \overline{X}_a^R Z^{RS} \underline{X}_j^S \overline{X}_b^S - \underline{X}_i^R \overline{X}_b^R Z^{RS} \underline{X}_j^S \overline{X}_a^S] \tag{30}$$

Since the Cholesky pseudo-MOs span the same space as the canonical and the Cholesky-MOs, the THC fitting can be performed once in the Cholesky-MO basis and does not have to be repeated for every Laplace point. $\underline{\mathbf{X}}$ and $\overline{\mathbf{X}}$ are the collocation matrices transformed into the Cholesky pseudo-MO basis given by

$$\underline{X}_i^P = \sum_\mu \underline{L}_{\mu i} X_\mu^P \tag{31}$$

$$\overline{X}_a^P = \sum_\nu \overline{L}_{\nu a} X_\nu^P \tag{32}$$

The benefit of evaluating the THC-MP2 energy in the Cholesky pseudo-MO basis is that the collocation matrices are expected to be sparse and contractions involving them can efficiently be done using sparse linear algebra. Since the $\mathbf{Z}$ tensor is the grid representation of the long-ranged $1/r_{12}$ operator, it is generally not sparse. For this reason, the Coulomb- and exchange-like parts of the MP2 energy are evaluated by first contracting all possible collocation matrices, only including the $\mathbf{Z}$ tensors in the last step, to carry the sparsity in $\mathbf{X}$ through most of the computation. By doing this, we obtain an algorithm with reduced scaling compared to the formal quartic scaling of this step. Detailed algorithms for the evaluation of the Coulomb- and exchange-like contributions to the MP2 energy are provided in the Supporting Information. It should also be noted that the computation times of the evaluation of the Coulomb-like energy are nearly independent of the choice of MOs because the time-determining step involves the contraction of the THC $\mathbf{Z}$ matrix with an intermediate that is invariant under orbital rotation. A more detailed discussion of this matter is included as a separate section in the Supporting Information. For our timings in Section 3, we make use of the SOS approximation[78] of the MP2 energy, considering only the Coulomb-like term

$$E_\mathrm{SOS\text{-}MP2}^\mathrm{THC\text{-}CDD} = -c_\mathrm{OS} \sum_\kappa \sum_{PQRS} \sum_{ijab} \underline{X}_i^P \overline{X}_a^P Z^{PQ} \underline{X}_j^Q \overline{X}_b^Q \cdot \underline{X}_i^R \overline{X}_a^R Z^{RS} \underline{X}_j^S \overline{X}_b^S \tag{33}$$

where c_OS is the opposite-spin scaling factor, for which we use 1.3 as reported by Jung *et al.*[78] The SOS approximation reduces the formal scaling of the energy evaluation to cubic once the **X** and **Z** tensors are obtained. The actual scaling will, however, have a reduced scaling exponent due to the aforementioned use of sparse linear algebra, see Section 3.2. For a formal comparison of THC-RI-SOS-MP2 against different SOS-MP2 methods, see the corresponding section in the Supporting Information.

3. RESULTS AND DISCUSSION

We implemented the above-described THC algorithm together with the THC-RI-CDD-MP2 method within our quantum chemistry package FermiONs++.[81−83] The cc-pVDZ and cc-pVTZ basis sets[84] are used in combination with the corresponding auxiliary basis sets for all RI calculations and the corresponding optimized THC grids.[67] For the phosphorus atoms occurring in the calculations on DNA, the fluorine grids were used. We demonstrate in the Supporting Information that this does not lead to an additional error by benchmarking against a small set of selected phosphorus species. DIIS acceleration[85] is used for the preceding SCF calculations, for which the SCF energy is converged to 10^{-12} and the commutator $\mathbf{FPS} - \mathbf{SPF}$ to 5×10^{-9} for the S22 test set[86] and the L7 test set,[87] together with 10 Laplace points for the integration of the orbital energy denominator. For the assessment of the scaling, the thresholds 10^{-7} and 10^{-6} were used, respectively, in combination with seven Laplace points, which is expected to provide sufficient accuracy in the submilli-Hartree regime. Fermi shifting is applied to the formation of the pseudo-density matrices as proposed by Ayala and Scuseria.[7] For the evaluation of the energy equations, we employ a block-sparse matrix format with a block threshold of 10^{-7} for the Frobenius norm of a block. All timings are done on an Intel Xeon E5-2667 v4 (3.20 GHz) CPU node with 256 GB RAM and three 960 GB SSD drives.

3.1. Accuracy of the THC-ω-RI-CDD-SOS-MP2 Method.

We begin with the analysis of the accuracy of the linear solver algorithm for the solution of eq 21 compared to the pseudoinversion ($\varepsilon_{\text{SVD}} = 10^{-12}$) suggested by Martínez and co-workers.[27,61−63] For this, benchmark calculations on the S22 test set were performed with our implementation of the THC-RI-CDD-SOS-MP2 method using the cc-pVDZ and cc-pVTZ basis sets. The modified Cholesky decomposition (LDL) was chosen as the linear solver routine, but we found that the solution based on LU decomposition provides virtually identical results. Table 1 summarizes the deviations of the absolute energies for the S22 monomers and dimers as well as the deviations of the relative energies for the complexes.

Table 1. Deviation of the Linear Solver Algorithm Compared to Pseudoinversion for the S22 Test Set

	ΔE_{abs}[a]		ΔE_{rel}[b]	
S22	MAD	MAX	MAD	MAX
cc-pVDZ	0.4	3.8	<0.001	0.002
cc-pVTZ	2.0	7.8	0.003	0.011

[a] Deviations in μH. [b] Deviations in kcal/mol.

It can be seen that the mean absolute deviations (MAD) of the absolute energies and the maximum error (MAX) are on the order of 10^{-6} H. As a consequence, the relative energies are also accurately reproduced. It should be emphasized that these deviations are not with respect to a reference method, like, e.g., RI-SOS-MP2, but to the pseudoinversion variant of THC, meaning that the errors are more likely due to cutoff errors based on the SVD threshold. The latter will be illustrated by comparing the results against RI-SOS-MP2 in the following.

After having established that implicitly forming the inverse of the grid metric matrix can be accurately done by the linear solution of eq 21, we now assess the accuracy of our THC-RI-CDD-SOS-MP2 method against the reference RI-SOS-MP2 method. For this, we compare two different variants of the THC algorithm against the reference with the cc-pVDZ and cc-pVTZ basis sets. The first variant uses the regular Coulomb metric and a tight natural blocking threshold of 10^{-10}, henceforth termed THC-RI-CDD-SOS-MP2 and the second variant uses the attenuated Coulomb metric from eq 10 with an attenuation strength ω of 0.1 and a looser natural blocking threshold of 10^{-6}, from here on referred to as THC-ω-RI-CDD-SOS-MP2. In Table 2, the absolute and relative errors are compared to the reference RI-SOS-MP2 method for the S22 and the larger L7 test sets.

Table 2. Error of THC-RI-CDD-SOS-MP2 and THC-ω-RI-CDD-SOS-MP2 Compared to Canonical RI-SOS-MP2 for the S22 and L7 Test Set Using the cc-pVXZ/cc-pVXZ-RI ($X \in \{D,T\}$) Basis Sets

			ΔE_{abs}[a]		ΔE_{rel}[b]	
test set	basis	THC variant	MAD	MAX	MAD	MAX
S22	cc-pVDZ	RI	8.8	38.9	0.003	0.013
		ω-RI	13.9	57.9	0.008	0.030
	cc-pVTZ	RI	2.0	14.0	0.001	0.010
		ω-RI	7.8	26.0	0.005	0.015
L7	cc-pVDZ	RI	72.8	284.8	0.027	0.067
		ω-RI	50.3	183.3	0.045	0.129
	cc-pVTZ	RI	50.6	187.3	0.010	0.030
		ω-RI	45.4	137.0	0.022	0.049

[a] Deviations in μH. [b] Deviations in kcal/mol.

Table 2 shows that for the S22 test set both THC variants provide good accuracies for the absolute energies with a MAD on the order of 10^{-6} H, which shows that the THC factorization only introduces an insignificant error compared to the well-established RI error.[27] Furthermore, the use of an attenuated Coulomb metric does not worsen the mean error significantly. The maximum errors are larger but still on the order of 10^{-5} H and in all cases correspond to aromatic molecules, like the nucleobases uracil or thymine, in the test set. The latter can be attributed to the fact that delocalized electronic structures are more difficult to fit with small grids in the LS-THC procedure, which has previously been reported by Parrish *et al.*[27] for acenes. Since the mean errors for the absolute energies are small, the relative energies are also accurately reproduced. For the L7 test set, comprised of larger molecules, the errors are worse compared to the results for the S22 test set and are roughly an order of magnitude larger. This is somewhat expected since the L7 test set represents a collection of worst case molecules, in a sense that all complexes involve large, strongly delocalized electronic structures. Remarkably, for this case, the errors for the cc-pVTZ basis set are smaller than for the cc-pVDZ basis set. This can be rationalized by the fact that for the triple-ζ basis set larger integration grids are used, which allow for better fitting of the delocalized orbitals. Furthermore, for the absolute energies, the errors decrease when using the attenuated Coulomb metric and integral screening, most likely due to favorable error cancellation. For the interaction energies, the same trend of

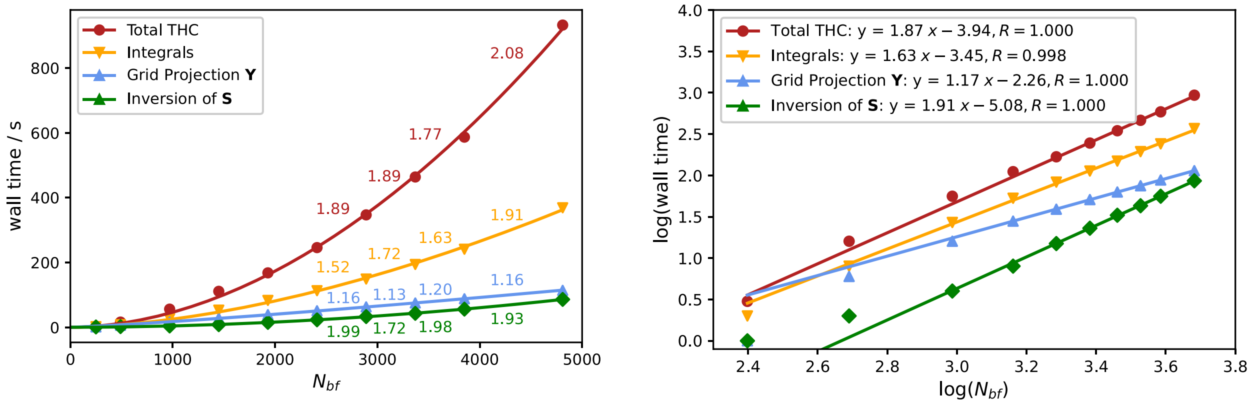
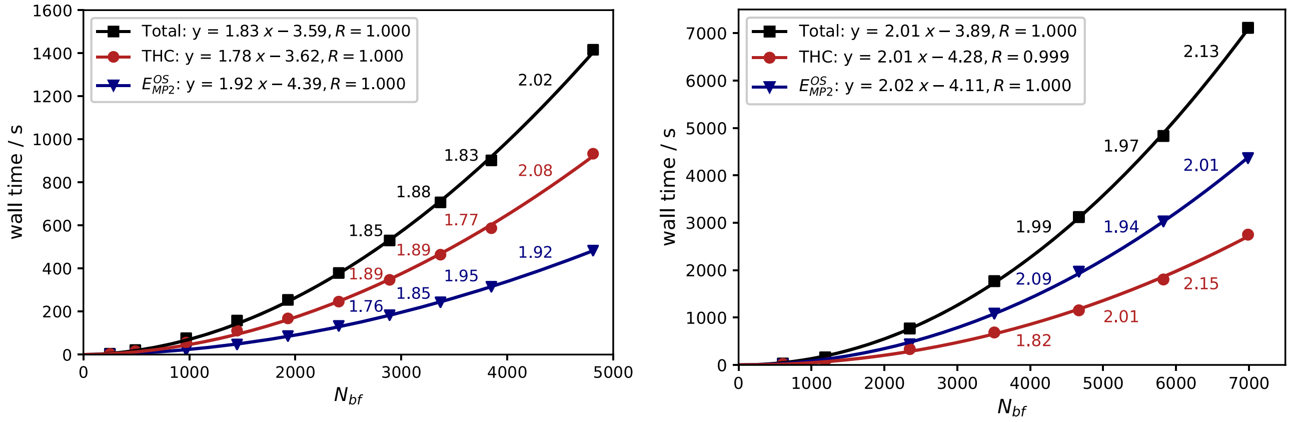


Figure 1. Total execution time of the THC algorithm (red) and individual contributions from the most expensive constituting steps, i.e., the formation of the RI integrals (orange), the construction of intermediate **Y** (blue), and the inversion of the grid metric **S** (green) for linear alkanes C_nH_{2n+2} using the cc-pVDZ basis set (left) and the corresponding double logarithmic plot (right). The colored numbers correspond to the scaling exponent with respect to the preceding fragment, and the scaling for the largest molecules was determined by linear regression starting from $C_{60}H_{122}$ ($N_{bf} = 1450$) up to $C_{200}H_{402}$ ($N_{bf} = 4810$).



(a) cc-pVDZ (b) cc-pVTZ

Figure 2. Plots of the total execution times for the THC-ω-RI-CDD-SOS-MP2 method (black) for linear alkanes C_nH_{2n+2} as well as the individual contributions from the THC fitting (red) and the computation of the MP2 energy (blue) for the cc-pVDZ (left) and cc-pVTZ basis sets (right). The scaling exponent for the longest chains, starting from $C_{60}H_{122}$, was determined by linear regression in the double logarithmic plot, and the colored numbers correspond to the scaling with respect to the preceding fragment.

increasing errors as compared to the S22 test set is observed. Nonetheless, the deviations are significantly lower than 1 kcal/mol, which can be considered to be chemical accuracy. To conclude, the THC-ω-RI-CDD-SOS-MP2 method introduces no significant errors compared to the THC-RI-CDD-SOS-MP2 implementation and provides good accuracy for the variety of compounds present in the test sets under investigation. We also note that our THC algorithm is not only able to accurately reproduce final MP2 energies but also provides good accuracy for important intermediate quantities such as MP2 amplitudes, which is demonstrated in the Supporting Information.

3.2. Performance Analysis. In the following, the performance of the here developed THC-ω-RI-CDD-SOS-MP2 method is assessed in terms of computational complexity and asymptotic scaling behavior. First, we focus on the THC routine itself, i.e., without the computation of the MP2 energy to investigate the most expensive steps of forming the THC factorized integrals. For this, we consider the construction of the RI integrals in the Cholesky-MO basis, the formation of the grid-projected RI integrals **Y** (eq 17), which is the formally highest scaling step and the inversion—or rather the linear solution to find the implicit inverse—of the grid metric **S** (eq 21). In Figure 1, this analysis is exemplarily carried out for linear alkanes up to $C_{200}H_{402}$ using the cc-pVDZ basis set. On the left, the computation times of the total THC procedure (red) and the individual steps are plotted against the number of basis functions.

It can be seen that the construction and contraction of the RI integrals dominate the scaling behavior. Initially, subquadratic scaling is observed for the RI integrals, which increases with bigger molecule sizes. This is caused by the fact that, for the largest chains, the inversion of the two-center integrals in eq 10 becomes computationally significant and starts to dominate the overall scaling of the construction of the RI integrals. We note that, for chemically relevant molecules, with a less sparse structure, this crossover is expected to be for larger molecule sizes. As the inversion is asymptotically cubic scaling, this worsens the overall scaling of this step, while the formation of the three-center integrals becomes close to linear scaling with the screening outlined in Section 2.3. As the same screening in combination with the natural blocking approach for the tensor contractions is applied to the construction of the **Y** tensor, this step also becomes close to linear scaling. For the linear solution of eq 21, a specialized band matrix version of the LU decomposition (LAPACK: xgbsv) was used. For this, reverse Cuthill–McKee[88] reordering of the atoms is performed in order to minimize the bandwidth of the resulting grid-projected overlap matrix. The bandwidth can then be determined numerically by the algorithm presented in the Supporting Information and converted to the LAPACK band storage format. We found that, after a certain chain length, here $C_{40}H_{82}$, the bandwidth is constant and the number of elements in **S** only increases linearly, which results in an overall quadratic scaling of this step. Even if no specialized linear solvers are applied, the prefactor of this step is small, such that it only makes up a small fraction of the total computation time. In total, using an attenuated Coulomb metric in the RI integrals and natural blocking for the corresponding contractions in combination with a localized Cholesky basis and specialized algorithms for the inversion of **S**, the overall scaling of obtaining the THC factorization becomes quadratic. It should, however, be noted that, for even larger molecule sizes, the scaling will eventually be dominated by asymptotically cubic scaling steps, such as the inversion of the two-center integrals or the formation of **Z** (eq 24), which involves only dense matrices.

Next, we focus the discussion on the overall scaling of the evaluation of the SOS-MP2 energy equation with the THC algorithm presented above. We first again consider linear alkanes as a model system to determine the asymptotic scaling behavior. Figure 2 shows the corresponding plots of the computation times against the number of basis functions for the cc-pVDZ (left) and cc-pVTZ basis sets (right). For this, the total wall times for the evaluation of the THC-ω-RI-CDD-SOS-MP2 equations are shown (black) as well as the contributions from obtaining the THC factorization (red) and the calculation of the SOS-MP2 energy (blue).

For both basis sets, roughly quadratic scaling is observed for the THC-ω-RI-CDD-SOS-MP2 method. Through the use of sparse matrix algebra, the evaluation of the SOS-MP2 energy can efficiently be performed since all quantities are sparse—except for the **Z** tensor—due to the transformation into the Cholesky pseudo-MO basis. It can be seen that for the smaller double-ζ basis set, obtaining the THC factorization is the rate-determining step and the energy evaluation only makes up 1/3 of the total computation time. This highlights an advantage of using a global THC formulation as, for the local-THC approximation, the evaluation of the SOS-MP2 energy seems to be dominating.[63] The latter is especially problematic when using more expensive correlation methods than MP2. For the larger triple-ζ basis set, however, the evaluation of the energy equations also becomes rate-determining for our global THC algorithm. Here, larger grids are employed and the basis set includes more diffuse functions, resulting in less sparse intermediates in the energy evaluation. Overall, for linear alkanes, as a sparse model, both the THC factorization and the energy evaluation are quadratic scaling with a low prefactor for the considered molecule sizes.

Even though linear alkanes provide a valuable model for assessing the efficiency of screening algorithms and the scaling of a method, we now focus on a class of molecules that is more representative for practical applications in the form of adenine–thymine base pair stacks $(AT)_n$ ($n \in \{1,2,4,8,12\}$). The same calculations as in Figure 2 with the cc-pVDZ basis set were performed for the DNA fragments, and the results are shown in Figure 3.

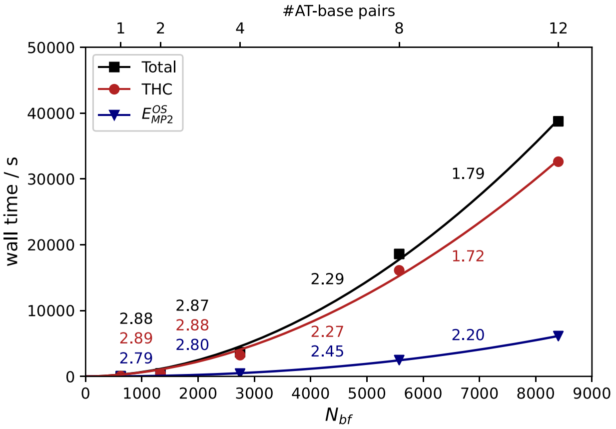


Figure 3. Total execution time of the THC-ω-RI-CDD-SOS-MP2 method (black) for $(AT)_n$ base pair stacks as well as the individual contributions from the THC fitting (red) and the computation of the MP2 energy (blue) for the cc-pVDZ basis set. The colored numbers correspond to the scaling with respect to the preceding fragment.

Here, all observations discussed before for linear alkanes still hold. The evaluation of the SOS-MP2 energy equations approaches quadratic scaling for the largest DNA fragment while obtaining the THC factorization reaches subquadratic scaling. The latter is due to the fact that, for the less sparse DNA systems, as compared to linear alkanes, the cubic scaling steps cease to be significant and the scaling is dominated by the linear scaling integral contractions. Thus, the whole THC-ω-RI-CDD-SOS-MP2 method also becomes subquadratically scaling. The more globular structure of DNA, however, causes a later onset of subquadratic scaling (linear alkanes: $N_{bf} \approx 1000$, DNA: $N_{bf} \approx 6000$) because only for the largest fragments the sparsity in the tensors involved becomes sufficiently large to be exploited by screening and sparse matrix algebra. It can be expected, that for even larger fragments the scaling behavior of the energy evaluation will reduce and approach the scaling observed for linear alkanes. For the largest fragment, i.e., $(AT)_{12}$, detailed wall times for the individual steps of the THC-ω-RI-CDD-SOS-MP2 method are presented in Table 3.

Table 3. Detailed Computation Times for the Individual Steps of the THC-ω-RI-CDD-SOS-MP2 Method for an $(AT)_{12}$ DNA Fragment (cc-pVDZ)

step	wall time (s)	%
integrals (eqs 11–13)	17,768	45.9
X (eqs 6–9)	208	0.5
Y (eq 17)	11,970	30.9
S (eq 4)	410	1.1
$\mathbf{S}^{-1}$ (eq 21)	960	2.5
Z (eq 24)	683	1.8
$E_{SOS-MP2}$ (eq 33)	6132	15.8
total	38,745	

As can be seen from Table 3, the operations involving the RI integrals, i.e., their construction and transformation into the MO basis as well as the grid projection to obtain intermediate **Y**, make up roughly 75% of the total computation time. However, we were able to show that these steps have been made linear scaling through the use of an attenuated Coulomb metric and the local Cholesky-MO basis in combination with integral screening (see Figure 1), with the exception of the inversion of the two-center integrals. Furthermore, the previously reported bottleneck of the THC procedure, i.e., the inversion of the grid metric **S**, was removed by the use of linear solvers, such that now this step only contributes 2.5% of the total computation time for the largest DNA fragment under consideration. Additionally, here the other cubic scaling steps, i.e., obtaining the **Z** tensor and inverting the two-center integrals, only contribute a small percentage of the total wall time and are not rate-determining. Finally, the evaluation of the SOS-MP2 energy equation only requires roughly 15% of the total wall time.

4. CONCLUSIONS

In the present work, we showed that, through a combination of different reduced-scaling methods, the computational complexity of the LS-THC formalism can significantly be reduced: By combining the locality of Cholesky orbitals with an erfc-attenuated Coulomb metric and appropriate integral screening for the RI integrals, all contractions involving these integrals can be performed in linear time complexity. This especially allowed us to perform the formally highest scaling step, i.e., the projection of the RI integrals onto the THC grids to form intermediate **Y**, in a linear scaling fashion. Additionally, the prefactor of the usually most expensive step, i.e., forming the inverse of the grid metric **S**, can be significantly reduced by reformulation of the underlying equations and application of linear equation solvers. All of this leads to a subquadratically scaling formulation of the LS-THC procedure for chemically relevant molecules. We note that this still involves cubic scaling steps, like the inversion of the two-center integrals or the final contraction to form the **Z** tensor; however, these steps have small enough prefactors to only show up for very large molecule sizes. Applying our THC method to SOS-MP2 leads to an overall subquadratically scaling MP2 method, which we term THC-ω-RI-CDD-SOS-MP2.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.0c00934.

> A more detailed description of the algorithm for the transformation of the RI integrals to the Cholesky-MO basis using the natural blocking formalism and the computation of the THC-CDD-MP2 energy; a comparison between using canonical and Cholesky (pseudo)-MOs and a comparison of THC-RI-SOS-MP2 against different SOS-MP2 methods; and details regarding the accuracy of final MP2 energies and MP2 amplitudes (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Christian Ochsenfeld – *Chair of Theoretical Chemistry, Department of Chemistry, University of Munich (LMU), D-81377 Munich, Germany*; orcid.org/0000-0002-4189-6558; Email: christian.ochsenfeld@cup.uni-muenchen.de

Authors

Felix H. Bangerter – *Chair of Theoretical Chemistry, Department of Chemistry, University of Munich (LMU), D-81377 Munich, Germany*

Michael Glasbrenner – *Chair of Theoretical Chemistry, Department of Chemistry, University of Munich (LMU), D-81377 Munich, Germany*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jctc.0c00934

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors acknowledge financial support by the "Deutsche Forschungsgemeinschaft" (DFG) via the SFB 1309-32587107 and the cluster of excellence (EXC2111-390814868) "Munich Center for Quantum Science and Technology" (MCQST). F.H.B. thanks the "Fonds der Chemischen Industrie" (FCI) for a graduate fellowship. M.G. thanks the "Studienstiftung des Deutschen Volkes" for a graduate fellowship. C.O. acknowledges additional financial support as a Max-Planck-Fellow at MPI-FKF Stuttgart.

■ REFERENCES

(1) Møller, C.; Plesset, M. S. Note on an approximation treatment for many-electron systems. *Phys. Rev.* **1934**, *46*, 618−622.

(2) Cremer, D. Møller-Plesset perturbation theory: from small molecule methods to methods for thousands of atoms. *WIREs Comput. Mol. Sci.* **2011**, *1*, 509−530.

(3) Grimme, S. Semiempirical hybrid density functional with perturbative second-order correlation. *J. Chem. Phys.* **2006**, *124*, No. 034108.

(4) Schwabe, T.; Grimme, S. Double-hybrid density functionals with long-range dispersion corrections: Higher accuracy and extended applicability. *Phys. Chem. Chem. Phys.* **2007**, *9*, 3397−3406.

(5) Flaig, D.; Maurer, M.; Hanni, M.; Braunger, K.; Kick, L.; Thubauville, M.; Ochsenfeld, C. Benchmarking Hydrogen and Carbon NMR Chemical Shifts at HF, DFT, and MP2 Levels. *J. Chem. Theory Comput.* **2014**, *10*, 572−578.

(6) Häser, M.; Almlöf, J. Laplace transform techniques in Møller-Plesset perturbation theory. *J. Chem. Phys.* **1992**, *96*, 489.

(7) Ayala, P. Y.; Scuseria, G. E. Linear scaling second-order Møller-Plesset theory in the atomic orbital basis for large molecular systems. *J. Chem. Phys.* **1999**, *110*, 3660.

(8) Doser, B.; Lambrecht, D. S.; Kussmann, J.; Ochsenfeld, C. Linear-scaling atomic orbital-based second-order Møller-Plesset perturbation theory by rigorous integral screening criteria. *J. Chem. Phys.* **2009**, *130*, No. 064107.

(9) Sæbø, S.; Pulay, P. Local configuration interaction: An efficient approach for larger molecules. *Chem. Phys. Lett.* **1985**, *113*, 13−18.

(10) Sæbø, S.; Pulay, P. Fourth-order Møller-Plessett perturbation theory in the local correlation treatment. I. Method. *J. Chem. Phys.* **1987**, *86*, 914.

(11) Sæbø, S.; Pulay, P. Local treatment of electron correlation. *Annu. Rev. Phys. Chem.* **1993**, *44*, 213−236.

(12) Hampel, C.; Werner, H. J. Local treatment of electron correlation in coupled cluster theory. *J. Chem. Phys.* **1996**, *104*, 6286.

(13) Schütz, M.; Hetzer, G.; Werner, H. J. Low-order scaling local electron correlation methods. I. Linear scaling local MP2. *J. Chem. Phys.* **1999**, *111*, 5691.

(14) Schütz, M.; Werner, H. J. Low-order scaling local electron correlation methods. IV. Linear scaling local coupled-cluster (LCCSD). *J. Chem. Phys* **2001**, *114*, 661.

(15) Schütz, M. A new, fast, semi-direct implementation of linear scaling local coupled cluster theory. *Phys. Chem. Chem. Phys.* **2002**, *4*, 3941−3947.

(16) Lee, M. S.; Maslen, P. E.; Head-Gordon, M. Closely approximating second-order Møller-Plesset perturbation theory with a local triatomics in molecules model. *J. Chem. Phys.* **2000**, *112*, 3592.

(17) Werner, H. J.; Manby, F. R.; Knowles, P. J. Fast linear scaling second-order Møller-Plesset perturbation theory (MP2) using local and density fitting approximations. *J. Chem. Phys.* **2003**, *118*, 8149.

(18) Maslen, P. E.; Head-Gordon, M. Noniterative local second order Møller-Plesset theory: Convergence with local correlation space. *J. Chem. Phys.* **1998**, *109*, 7093.

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

(20) Beebe, N. H. F.; Linderberg, J. Simplifications in the Generation and Transformation of Two-Electron Integrals in Molecular Calculations. *Int. J. Quantum Chem.* **1977**, *12*, 683−705.

(21) Røeggen, I.; Wisløff-Nilssen, E. On the Beebe-Linderberg two-electron integral approximation. *Chem. Phys. Lett.* **1986**, *132*, 154−160.

(22) Weigend, F.; Kattannek, M.; Ahlrichs, R. Approximated electron repulsion integrals: Cholesky decomposition versus resolution of the identity methods. *J. Chem. Phys.* **2009**, *130*, 164106.

(23) Ringnalda, M. N.; Belhadj, M.; Friesner, R. A. Pseudospectral Hartree-Fock theory: Applications and algorithmic improvements. *J. Chem. Phys.* **1990**, *93*, 3397−3407.

(24) Martínez, T. J.; Mehta, A.; Carter, E. A. Pseudospectral full configuration interaction. *J. Chem. Phys.* **1992**, *97*, 1876.

(25) Murphy, R. B.; Beachy, M. D.; Friesner, R. A.; Ringnalda, M. N. Pseudospectral localized Møller-Plesset methods: Theory and calculation of conformational energies. *J. Chem. Phys.* **1995**, *103*, 1481−1490.

(26) Hohenstein, E. G.; Parrish, R. M.; Martínez, T. J. Tensor hypercontraction density fitting I. Quartic scaling second- and third-order Møller-Plesset perturbation theory. *J. Chem. Phys.* **2012**, *137*, No. 044103.

(27) Parrish, R. M.; Hohenstein, E. G.; Martínez, T. J.; Sherrill, C. D. Tensor hypercontraction II. Least-squares renormalization. *J. Chem. Phys.* **2012**, *137*, 224106.

(28) Foster, J. M.; Boys, S. F. Canonical Configurational Interaction Procedure. *Rev. Mod. Phys.* **1960**, *32*, 300−302.

(29) Pipek, J.; Mezey, P. G. A fast intrinsic localization procedure applicable for ab initio and semiempirical linear combination of atomic orbital wave functions. *J. Chem. Phys.* **1989**, *90*, 4916.

(30) Riplinger, C.; Pinski, P.; Becker, U.; Valeev, E. F.; Neese, F. Sparse maps - A systematic infrastructure for reduced-scaling electronic structure methods. II. Linear scaling domain based pair natural orbital coupled cluster theory. *J. Chem. Phys.* **2016**, *144*, No. 024109.

(31) Neese, F.; Wennmohs, F.; Hansen, A. Efficient and accurate local approximations to coupled-electron pair approaches: An attempt to revive the pair natural orbital method. *J. Chem. Phys.* **2009**, *130*, 114108.

(32) Werner, H. J.; Knizia, G.; Krause, C.; Schwilk, M.; Dornbach, M. Scalable electron correlation methods I.: PNO-LMP2 with linear scaling in the molecular size and near-inverse-linear scaling in the number of processors. *J. Chem. Theory Comput.* **2015**, *11*, 484−507.

(33) Pinski, P.; Riplinger, C.; Valeev, E. F.; Neese, F. Sparse maps - A systematic infrastructure for reduced-scaling electronic structure methods. I. An efficient and simple linear scaling local MP2 method that uses an intermediate basis of pair natural orbitals. *J. Chem. Phys.* **2015**, *143*, No. 034108.

(34) Høyvik, I. M.; Kristensen, K.; Jansik, B.; Jørgensen, P. The divide-expand-consolidate family of coupled cluster methods: Numerical illustrations using second order Møller-Plesset perturbation theory. *J. Chem. Phys.* **2012**, *136*, No. 014105.

(35) Baudin, P.; Ettenhuber, P.; Reine, S.; Kristensen, K.; Kjærgaard, T. Efficient linear-scaling second-order Møller-Plesset perturbation theory: The divide-expand-consolidate RI-MP2 model. *J. Chem. Phys.* **2016**, *144*, No. 054102.

(36) Ettenhuber, P.; Baudin, P.; Kjærgaard, T.; Jørgensen, P.; Kristensen, K. Orbital spaces in the divide-expand-consolidate coupled cluster method. *J. Chem. Phys.* **2016**, *144*, 164116.

(37) Kristensen, K.; Kjærgaard, T.; Høyvik, I. M.; Ettenhuber, P.; Jørgensen, P.; Jansik, B.; Reine, S.; Jakowski, J. The divide-expand-consolidate MP2 scheme goes massively parallel. *Mol. Phys.* **2013**, *111*, 1196−1210.

(38) Russ, N. J.; Crawford, T. D. Potential energy surface discontinuities in local correlation methods. *J. Chem. Phys.* **2004**, *121*, 691.

(39) Mata, R. A.; Werner, H. J. Calculation of smooth potential energy surfaces using local electron correlation methods. *J. Chem. Phys.* **2006**, *125*, 184110.

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

(41) Häser, M. Møller-Plesset (MP2) perturbation theory for large molecules. *Theor. Chim. Acta* **1993**, *87*, 147−173.

(42) 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.

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

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

(45) 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.

(46) Røeggen, I.; Johansen, T. Cholesky decomposition of the two-electron integral matrix in electronic structure calculations. *J. Chem. Phys.* **2008**, *128*, 194107.

(47) Boström, J.; Aquilante, F.; Pedersen, T. B.; Lindh, R. Ab initio density fitting: Accuracy assessment of auxiliary basis sets from cholesky decompositions. *J. Chem. Theory Comput.* **2009**, *5*, 1545−1553.

(48) Aquilante, F.; Gagliardi, L.; Pedersen, T. B.; Lindh, R. Atomic Cholesky decompositions: A route to unbiased auxiliary basis sets for density fitting approximation with tunable accuracy and efficiency. *J. Chem. Phys.* **2009**, *130*, 154107.

(49) Aquilante, F.; Delcey, M. G.; Pedersen, T. B.; Galván, I. F.; Lindh, R. Inner projection techniques for the low-cost handling of two-electron integrals in quantumchemistry. *Mol. Phys.* **2017**, *115*, 2052−2064.

(50) Folkestad, S. D.; Kjønstad, E. F.; Koch, H. An efficient algorithm for Cholesky decomposition of electron repulsion integrals. *J. Chem. Phys.* **2019**, *150*, 194112.

(51) 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.

(52) Bernholdt, D. E.; Harrison, R. J. Largescale correlated electronic structure calculations: The RI-MP2 method on parallel computers. *Chem. Phys. Lett.* **1996**, *250*, 477−484.

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

(54) Černý, J.; Jurečka, P.; Hobza, P.; Valdés, H. Resolution of identity density functional theory augmented with an empirical dispersion term (RI-DFT-D): A promising tool for studying isolated small peptides. *J. Phys. Chem. A* **2007**, *111*, 1146−1154.

(55) Grimme, S.; Neese, F. Double-hybrid density functional theory for excited electronic states of molecules. *J. Chem. Phys.* **2007**, *127*, 154116.

(56) 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.

(57) Graf, D.; Beuerle, M.; Ochsenfeld, C. Low-Scaling Self-Consistent Minimization of a Density Matrix Based Random Phase Approximation Method in the Atomic Orbital Space. *J. Chem. Theory Comput.* **2019**, *15*, 4468−4477.

(58) Epifanovsky, E.; Zuev, D.; Feng, X.; Khistyaev, K.; Shao, Y.; Krylov, A. I. General implementation of the resolution-of-the-identity and Cholesky representations of electron repulsion integrals within coupled-cluster and equation-of-motion methods: Theory and benchmarks. *J. Chem. Phys.* **2013**, *139*, 134105.

(59) Neese, F.; Wennmohs, F.; Hansen, A.; Becker, U. Efficient, approximate and parallel Hartree-Fock and hybrid DFT calculations. A

'chain-of-spheres' algorithm for the Hartree-Fock exchange. *Chem. Phys.* **2009**, *356*, 98−109.

(60) Laqua, H.; Thompson, T. H.; Kussmann, J.; Ochsenfeld, C. Highly Efficient, Linear-Scaling Seminumerical ExactExchange Method for Graphic Processing Units. *J. Chem. Theory Comput.* **2020**, *16*, 1456−1468.

(61) Parrish, R. M.; Hohenstein, E. G.; Martínez, T. J.; Sherrill, C. D. Discrete variable representation in electronic structure theory: Quadrature grids for least-squares tensor hypercontraction. *J. Chem. Phys.* **2013**, *138*, 194107.

(62) Song, C.; Martínez, T. J. Atomic orbital-based SOS-MP2 with tensor hypercontraction. I. GPU-based tensor construction and exploiting sparsity. *J. Chem. Phys.* **2016**, *144*, 174111.

(63) Song, C.; Martínez, T. J. Atomic orbital-based SOS-MP2 with tensor hypercontraction. II. Local tensor hypercontraction. *J. Chem. Phys.* **2017**, *146*, No. 034104.

(64) Schmitz, G.; Madsen, N. K.; Christiansen, O. Atomic-batched tensor decomposed two-electron repulsion integrals. *J. Chem. Phys.* **2017**, *146*, 134112.

(65) Schmitz, G.; Christiansen, O. Assessment of the overlap metric in the context of RI-MP2 and atomic batched tensor decomposed MP2. *Chem. Phys. Lett.* **2018**, *701*, 7−14.

(66) Lee, J.; Lin, L.; Head-Gordon, M. Systematically Improvable Tensor Hypercontraction: Interpolative Separable Density Fitting for Molecules Applied to Exact Exchange, Second- and Third-Order Møller-Plesset Perturbation Theory. *J. Chem. Theory Comput.* **2019**, *16*, 243−263.

(67) Kokkila Schumacher, S. I. L.; Hohenstein, E. G.; Parrish, R. M.; Wang, L.-P.; Martínez, T. J. Tensor Hypercontraction Second-Order Møller-Plesset Perturbation Theory: Grid Optimization and Reaction Energies. *J. Chem. Theory Comput.* **2015**, *11*, 3042−3052.

(68) Song, C.; Martínez, T. J. Analytical gradients for tensor hypercontracted MP2 and SOS-MP2 on graphical processing units. *J. Chem. Phys.* **2017**, *147*, 161723.

(69) Matthews, D. A. Improved Grid Optimization and Fitting in Least Squares Tensor Hypercontraction. *J. Chem. Theory Comput.* **2020**, *16*, 1382−1385.

(70) Parrish, R. M.; David Sherrill, C.; Hohenstein, E. G.; Kokkila, S. I. L.; Martínez, T. J. Communication: Acceleration of coupled cluster singles and doubles via orbital-weighted least-squares tensor hypercontraction. *J. Chem. Phys.* **2014**, *140*, 181102.

(71) Hohenstein, E. G.; Kokkila, S. I. L.; Parrish, R. M.; Martínez, T. J. Quartic scaling second-order approximate coupled cluster singles and doubles via tensor hypercontraction: THC-CC2. *J. Chem. Phys.* **2013**, *138*, 124111.

(72) Hohenstein, E. G.; Kokkila, S. I. L.; Parrish, R. M.; Martínez, T. J. Tensor hypercontraction equation-of-motion second-order approximate coupled cluster: Electronic excitation energies in O(N4) time. *J. Phys. Chem. B* **2013**, *117*, 12972−12978.

(73) Schutski, R.; Zhao, J.; Henderson, T. M.; Scuseria, G. E. Tensor-structured coupled cluster theory. *J. Chem. Phys.* **2017**, *147*, 184113.

(74) Song, C.; Martínez, T. J. Reduced scaling CASPT2 using supporting subspaces and tensor hyper-contraction. *J. Chem. Phys.* **2018**, *149*, No. 044108.

(75) 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.

(76) Thompson, T. H.; Ochsenfeld, C. Integral partition bounds for fast and effective screening of general one-, two-, and manyelectron integrals. *J. Chem. Phys.* **2019**, *150*, No. 044101.

(77) 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.

(78) 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.

(79) Aquilante, F.; Bondo Pedersen, T.; Sánchez De Merás, A.; Koch, H. Fast noniterative orbital localization for large molecules. *J. Chem. Phys.* **2006**, *125*, 174101.

(80) Glasbrenner, M.; Graf, D.; Ochsenfeld, C. Efficient reduced-scaling second-order Møller-Plesset perturbation theory with Cholesky-decomposed densities and an attenuated Coulomb metric. *J. Chem. Theory Comput.* **2020**, *16*, 6856−6868.

(81) Kussmann, J.; Ochsenfeld, C. Preselective screening for matrix elements in linear-scaling exact exchange calculations. *J. Chem. Phys.* **2013**, *138*, 134114.

(82) 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.

(83) Kussmann, J.; Ochsenfeld, C. Hybrid CPU/GPU Integral Engine for Strong-Scaling Ab Initio Methods. *J. Chem. Theory Comput.* **2017**, *13*, 3153−3159.

(84) Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007.

(85) Pulay, P. Convergence acceleration of iterative sequences. the case of SCF iteration. *Chem. Phys Lett.* **1980**, *73*, 393−398.

(86) Jurečka, P.; Šponer, J.; Černý, J.; Hobza, P. 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*, 1985−1993.

(87) Sedlak, R.; Janowski, T.; Pitoňák, M.; Řezáč, J.; Pulay, P.; Hobza, P. Accuracy of quantum chemical methods for large noncovalent complexes. *J. Chem. Theory Comput.* **2013**, *9*, 3364−3374.

(88) Cuthill, E.; McKee, J. Reducing the bandwidth of sparse symmetric matrices. *Proceedings of the 1969 24th national conference*; ACM *1969*, 157−172.

Supporting Information: Low-scaling Tensor Hypercontraction in the Cholesky Molecular Orbital Basis Applied to Second-Order Møller-Plesset Perturbation Theory

Felix H. Bangerter, Michael Glasbrenner, and Christian Ochsenfeld*

Chair of Theoretical Chemistry, Department of Chemistry, University of Munich (LMU),
Butenandtstrasse 5-13, D-81377 Munich, Germany

E-mail: christian.ochsenfeld@cup.uni-muenchen.de

Contents

1 Transformation of the Three-center Integrals

Following up on our work on ω-RI-CDD-MP2,[S1] we employ the natural blocking matrix format[S2] in combination with screening based on integral partition bounds (IPB)[S3] for the transformation of the three-center RI integrals into Cholesky-MO basis. In short, in the three-center AO integral kernel the matrices $\mathbf{M}$ and $\mathbf{N}$,

$$M_{\mu\nu} = \max_{\alpha} |(\mu\nu \vdots \alpha)|, \tag{1}$$

$$N_{\mu\alpha} = \max_{\nu} |(\mu\nu \vdots \alpha)|, \tag{2}$$

are constructed. The screening matrices $\mathbf{M}$ and $\mathbf{N}$ can then be turned into upper bounds for the transformation into Cholesky-MO basis by contraction with the absolute Cholesky factors $L_{\mu i}$ and $L_{\mu a}$ according to

$$|(i\nu \vdots \alpha)| \leq N_{i\alpha} = \sum_{\mu} |L_{\mu i}| N_{\mu\alpha}, \tag{3}$$

$$|(a\nu \vdots \alpha)| \leq N_{a\alpha} = \sum_{\mu} |L_{\mu a}| N_{\mu\alpha}. \tag{4}$$

The elements of $N_{i\alpha}$ and $N_{a\alpha}$ can then be used to construct significance lists

$$[i]_{\alpha} \equiv \{i | N_{i\alpha} > \epsilon_{\mathrm{NB}}\}, \tag{5}$$

$$[a]_{\alpha} \equiv \{a | N_{a\alpha} > \epsilon_{\mathrm{NB}}\}. \tag{6}$$

With these significance lists the Cholesky-MO transformation can be carried out in linear-scaling fashion using the algorithm presented below.

Algorithm 1 Transform the three-center RI integrals into Cholesky-MO basis.

```
procedure TRANSFORMRI
    for all $\alpha$ do
        for all $i \in [i]_\alpha$ and $\nu \in \{\nu\}_\alpha$ do
            $(\alpha \vdots i\nu) = \sum_{\mu \in \{\mu\}_\alpha} L_{\mu i}(\alpha \vdots \mu\nu)$
        end for
        build $\{i\}_\alpha$ list
    end for
    for all $\alpha$ do
        for all $i \in \{i\}_\alpha$ and $a \in [a]_\alpha$ do
            $(\alpha \vdots ia) = \sum_{\nu \in \{\nu\}_\alpha} L_{\nu a}(\alpha \vdots i\nu)$
        end for
        build $\{a\}_\alpha$ list
    end for
    reorder: $(\alpha \vdots ia) \rightarrow (ia \vdots \alpha)$
    build $\{a\}_i$ list
    return $(ia \vdots \alpha)$
end procedure
```

2 Band Matrix Solver for Y

Here we present the pseudocode for determining the numerical bandwidth of a symmetric matrix **A** of size $n \times n$ and converting it into the band matrix storage format of LAPACK. Following the notation of the LAPACK documentation, kl is the number of significant subdiagonals and ku the number of superdiagonals. Since **A** is expected to be symmetric $kl = ku$ and the total bandwidth is given by $2kl + 1$. The algorithm determines the bandwidth by iterating over all subdiagonals and accumulating the elements of a given subdiagonal in the variable `sum`. Once the sum is less than the number of elements in this subdiagonals times a given truncation threshold `thresh`, the subdiagonal is considered to be insignificant.

Algorithm 2 Determine the numerical band width of a symmetric matrix **A**.

```
procedure DETERMINE_NUMERICAL_BANDWIDTH($\mathbf{A}$)
    kl $\leftarrow$ 0
    for sub $\leftarrow$ 1 to sub $<$ n do                    ▷ iterate over all subdiagonals
        sum $\leftarrow$ 0.0
        for k $\leftarrow$ 0 to k $<$ n - sub do     ▷ iterate over all elements in a subdiagonal
            sum $\leftarrow$ sum + $\mathbf{A}$(k + sub, k)
        end for
        if sum $<$ ((n - sub) $\times$ thresh) then
            break
        end if
        kl $\leftarrow$ kl + 1
    end for
    return kl
end procedure
```

Algorithm 3 Convert matrix **A** to LAPACK band matrix storage format.

```
procedure CONVERT_TO_BAND_STORAGE($\mathbf{A}$)
    kl, ku $\leftarrow$ DETERMINE_NUMERICAL_BANDWIDTH($\mathbf{A}$)
    $\mathbf{B}$ $\leftarrow$ $\mathbf{0}$(kl+ku+1, n)
    for j $\leftarrow$ 0 to j $<$ n do
        for i $\leftarrow$ max(0, j-ku) to i $<$ min(n, j + kl) do
            $\mathbf{B}$(ku + i - j, j) $\leftarrow$ $\mathbf{A}$(i, j)
        end for
    end for
    return $\mathbf{B}$
end procedure
```

Finally, we note that for the `xgbsv` routine the resulting matrix **B** has to be copied to a zero-padded matrix with dimensions (2`kl`+`ku`+1)×`n`.

3 Evaluation of the MP2 Energy Equations using THC

Here the pseudocode for the algorithms to compute the THC-CDD-MP2 energy is presented. The algorithms for the Coulomb-like and exchange-like energy contribution are shown separately, but both make use of first contracting all possible MO indices after transformation of the collocation matrices into pseudo-MO basis. Here, we present the algorithm to compute the Coulomb-like contribution $E^{\text{THC-CDD}}_{\text{MP2-J}}$ to the MP2 energy with $\mathcal{O}(N^3)$ scaling. We note, that not all contractions are considered explicitly as some of the intermediates are equivalent.

Algorithm 4 Compute the Coulomb-like MP2 energy $E^{\text{THC-CDD}}_{\text{MP2-J}}$

```
procedure $E^{\text{THC-CDD}}_{\text{MP2-J}} \leftarrow -2 \sum_{\kappa} \sum_{PQRS} \sum_{ijab} \underline{X}_i^P \overline{X}_a^P Z^{PQ} \underline{X}_j^Q \overline{X}_b^Q \cdot \underline{X}_i^R \overline{X}_a^R Z^{RS} \underline{X}_j^S \overline{X}_b^S$
    $E_{\text{MP2-J}} \leftarrow 0.0$
    for $\kappa \leftarrow 1$ to $N_\kappa$ do                                    ▷ Laplace quadrature
        $\underline{X}_i^P \leftarrow \sum_{\mu} \underline{L}_{\mu i} X_{\mu}^P$                      ▷ $\mathcal{O}(N_{\text{grid}} N_{\text{bf}} N_{\text{occ}})$
        $\overline{X}_a^P \leftarrow \sum_{\nu} \overline{L}_{\nu a} X_{\nu}^P$                      ▷ $\mathcal{O}(N_{\text{grid}} N_{\text{bf}} N_{\text{virt}})$
        $A_{(\kappa)}^{PR} \leftarrow \sum_i \underline{X}_i^P \underline{X}_i^R$                      ▷ $\mathcal{O}(N_{\text{grid}}^2 N_{\text{occ}})$
        $B_{(\kappa)}^{PR} \leftarrow \sum_a \overline{X}_a^P \overline{X}_a^R$                      ▷ $\mathcal{O}(N_{\text{grid}}^2 N_{\text{virt}})$
        $C_{(\kappa)}^{PR} \leftarrow A_{(\kappa)}^{PR} \circ B_{(\kappa)}^{PR}$                      ▷ $\mathcal{O}(N_{\text{grid}}^2)$
        $D_{(\kappa)}^{PS} \leftarrow \sum_R C_{(\kappa)}^{PR} Z^{RS}$                      ▷ $\mathcal{O}(N_{\text{grid}}^3)$
        $E_{\text{MP2-J}} \mathrel{+}= -2 \sum_{PS} D_{(\kappa)}^{PS} D_{(\kappa)}^{SP}$                      ▷ $\mathcal{O}(N_{\text{grid}}^2)$
    end for
end procedure
```

Next, the algorithm to compute the exchange-like contribution $E^{\text{THC-CDD}}_{\text{MP2-K}}$ to the MP2 energy with $\mathcal{O}(N^4)$ scaling is presented. To circumvent memory issues, the accumulation of the final $E^{\text{THC-CDD}}_{\text{MP2-K}}$ energy contribution is batched over the occupied index i which avoids the necessity of storing higher than second-order tensors. Again, we note that not all contractions are considered explicitly as some of the intermediates are equivalent. Here, this is especially relevant for intermediates $A_{(\kappa)}^{PR}$ and $A_{(\kappa)}^{QS}$.

Algorithm 5 Compute the exchange-like MP2 energy $E^{\text{THC-CDD}}_{\text{MP2-K}}$

```
procedure $E^{\text{THC-CDD}}_{\text{MP2-K}} \leftarrow \sum_{\kappa} \sum_{PQRS} \sum_{ijab} \underline{X}_i^P \overline{X}_a^P Z^{PQ} \underline{X}_j^Q \overline{X}_b^Q \cdot \underline{X}_i^R \overline{X}_b^R Z^{RS} \underline{X}_j^S \overline{X}_a^S$
    $E_{\text{MP2-K}} \leftarrow 0.0$
    for $\kappa \leftarrow 1$ to $N_\kappa$ do                                   ▷ Laplace quadrature
        $\underline{X}_i^P \leftarrow \sum_{\mu} \underline{L}_{\mu i} X_\mu^P$                                  ▷ $\mathcal{O}(N_{\text{grid}} N_{\text{bf}} N_{\text{occ}})$
        $\overline{X}_a^P \leftarrow \sum_{\nu} \overline{L}_{\nu a} X_\nu^P$                                  ▷ $\mathcal{O}(N_{\text{grid}} N_{\text{bf}} N_{\text{virt}})$
        $A^{PR}_{(\kappa)} \leftarrow \sum_i \underline{X}_i^P \underline{X}_i^R$                                 ▷ $\mathcal{O}(N_{\text{grid}}^2 N_{\text{occ}})$
        $B^{PS}_{(\kappa)} \leftarrow \sum_a \overline{X}_a^P \overline{X}_a^S$                                 ▷ $\mathcal{O}(N_{\text{grid}}^2 N_{\text{virt}})$
        $e^{(\kappa)}_{\text{MP2-K}} \leftarrow 0.0$
        for $i \leftarrow 1$ to $i \leq \text{rk}(\underline{\mathbf{P}})$ do              ▷ Batching over occupied index $i$
            $C^{PS}_{(\kappa)} \leftarrow B^{PS}_{(\kappa)} \underline{X}_i^P(i,:)$                            ▷ $\mathcal{O}(N_{\text{grid}}^2 N_{\text{occ}})$
            $D^{QS}_{(\kappa)} \leftarrow \sum_P Z^{PQ} C^{PS}_{(\kappa)}$                             ▷ $\mathcal{O}(N_{\text{grid}}^3 N_{\text{occ}})$
            $e^{(\kappa)}_{\text{MP2-K}} \mathrel{+}= \sum_{QS} A^{QS}_{(\kappa)} D^{QS}_{(\kappa)} D^{SQ}_{(\kappa)}$                 ▷ $\mathcal{O}(N_{\text{grid}}^2 N_{\text{occ}})$
        end for
        $E_{\text{MP2-K}} \mathrel{+}= e^{(\kappa)}_{\text{MP2-K}}$
    end for
end procedure
```

4 Comparison of Different SOS-MP2 Methods

To put our THC-RI-SOS-MP2 method into context with different SOS-MP2 methods we compare it to LT-RI-SOS-MP2 based on work by Jung *et al.*[S4] and to CD-SOS-MP2 based on work by Aquilante *et al.*[S5] The LT-RI-SOS-MP2 method makes use of the Laplace transformation of the orbital energy denominator of the MP2 energy, which in combination with density-fitted integrals achieves a decoupling of the bra- and ket-indices. As noted in the original publication,[S4] the most expensive step of the method is the contraction of two third-order **B**-tensors to form the final intermediate for the energy evaluation, i.e., **Z**. Asymptotically this step scales as $\mathcal{O}(N_{\text{occ}} N_{\text{virt}} N_{\text{aux}}^2)$ and has to be performed for every Laplace point. The CD-SOS-MP2 method on the other hand is based on a Cholesky decomposition of the MP2 amplitudes while the remaining ERI can either be approximated by RI or also be Cholesky-decomposed. This – similar to LT-RI-SOS-MP2 – achieves a factorization into at most third-order tensors. For CD-SOS-MP2, however, the time-determining steps have to be performed only once as a preprocessing step before the actual energy calculation. Asymptotically both methods are quartic scaling with the relative performance mainly governed by the ratio of these steps. THC-RI-SOS-MP2 is another quartic scaling SOS-MP2 method which like LT-RI-SOS-MP2 uses a Laplace transformation of the orbital energy denominator and a preprocessing step in the form of factorizing the ERIs into THC format. The quartic scaling step in THC-RI-SOS-MP2 is the contraction of the collocation matrices $\mathbf{X}$ with the third-order RI tensors. The actual energy calculation requires at most $\mathcal{O}(N_{\mathrm{grid}}^3)$ scaling steps and therefore offers an advantage over LT-RI-SOS-MP2 and CD-SOS-MP2. The energy expressions of the three methods together with the scaling behavior of the most expensive steps are presented below.

LT-RI-SOS-MP2

$$\begin{aligned} E_{\text{OS-MP2}} &= -2\sum_{\kappa}\sum_{ijab}(\underline{i}\overline{a}|\underline{j}\overline{b})(\underline{i}\overline{a}|\underline{j}\overline{b}) \\ &= -2\sum_{\kappa}\sum_{\alpha\beta}\sum_{ia} B_{\underline{i}\overline{a}}^{\alpha}B_{\underline{i}\overline{a}}^{\beta}\sum_{jb} B_{\underline{j}\overline{b}}^{\alpha}B_{\underline{j}\overline{b}}^{\beta} \\ &= -2\sum_{\kappa}\sum_{\alpha\beta} Z_{\alpha\beta}Z_{\alpha\beta} \end{aligned}$$

B_{ia}^{α}	$\mathcal{O}(N_{\text{bf}}^2N_{\text{occ}}N_{\text{aux}}) + \mathcal{O}(N_{\text{bf}}N_{\text{occ}}N_{\text{virt}}N_{\text{aux}})$
$B_{\underline{i}\overline{a}}^{\alpha}$	$\mathcal{O}(N_{\text{occ}}N_{\text{virt}}N_{\text{aux}})$
$\mathbf{Z}$	$\mathcal{O}(N_{\text{occ}}N_{\text{virt}}N_{\text{aux}}^2)$
E	$\mathcal{O}(N_{\text{aux}}^2)$
memory	$\mathcal{O}(N_{\text{bf}}^2N_{\text{aux}})$[a] or $\mathcal{O}(N_{\text{occ}}N_{\text{virt}}N_{\text{aux}})$

CD-SOS-MP2

$$\begin{aligned} E_{\text{OS-MP2}} &= -2\sum_{ijab} t_{ij}^{ab}(ia|jb) \\ &= -2\sum_{K}\sum_{J}\sum_{ia} R_{ia}^{K}L_{ia}^{J}\sum_{jb} R_{jb}^{K}L_{jb}^{J} \\ &= -2\sum_{K}\sum_{J} Z^{KJ}Z^{KJ} \end{aligned}$$

t_{ij}^{ab}	$\mathcal{O}(N_{\text{occ}}N_{\text{virt}}N_{\text{aux}}N_{\text{chol}}) + \mathcal{O}(N_{\text{occ}}N_{\text{virt}}N_{\text{chol}}^2/2)$
$\mathbf{Z}$	$\mathcal{O}(N_{\text{occ}}N_{\text{virt}}N_{\text{aux}}N_{\text{chol}})$
E	$\mathcal{O}(N_{\text{aux}}N_{\text{chol}})$
memory	$\mathcal{O}(N_{\text{occ}}^2N_{\text{virt}}^2)$[b] or $\mathcal{O}(N_{\text{occ}}N_{\text{virt}}N_{\text{chol}})$

THC-RI-SOS-MP2

$$\begin{aligned} E_{\text{OS-MP2}} &= -2\sum_{\kappa}\sum_{PQRS}\sum_{ijab} \underline{X}_i^P\overline{X}_a^P Z^{PQ}\underline{X}_j^Q\overline{X}_b^Q \cdot \\ &\qquad \underline{X}_i^R\overline{X}_a^R Z^{RS}\underline{X}_j^S\overline{X}_b^S \\ &= -2\sum_{\kappa}\sum_{PQRS} A_{(\kappa)}^{PR}B_{(\kappa)}^{PR}Z^{PQ}Z^{RS}A_{(\kappa)}^{QS}B_{(\kappa)}^{QS} \\ &= -2\sum_{\kappa}\sum_{PQRS} C_{(\kappa)}^{PR}Z^{PQ}Z^{RS}C_{(\kappa)}^{QS} \\ &= -2\sum_{\kappa}\sum_{PS} D_{(\kappa)}^{PS}D_{(\kappa)}^{PS} \end{aligned}$$

$\mathbf{Y}$	$\mathcal{O}(N_{\text{occ}}N_{\text{virt}}N_{\text{aux}}N_{\text{grid}}) + \mathcal{O}(N_{\text{occ}}N_{\text{aux}}N_{\text{grid}})$
$\mathbf{Z}$	$\mathcal{O}(N_{\text{aux}}N_{\text{grid}}^2)$
$\mathbf{A}, \mathbf{B}$	$\mathcal{O}(N_{\text{occ}}N_{\text{grid}}^2)$, $\mathcal{O}(N_{\text{virt}}N_{\text{grid}}^2)$
$\mathbf{D}$	$\mathcal{O}(N_{\text{grid}}^3)$
E	$\mathcal{O}(N_{\text{grid}}^2)$
memory	$\mathcal{O}(N_{\text{bf}}^2N_{\text{aux}})$[a] or $\mathcal{O}(N_{\text{occ}}N_{\text{virt}}N_{\text{aux}})$[a] or $\mathcal{O}(N_{\text{grid}}^2)$

[a] if not implemented with on-the-fly integral calculation/contraction

[b] if t_{ij}^{ab} is not computed on-the-fly and Cholesky decomposition of the t_{ij}^{ab} supermatrix is performed

All methods are formally quartic scaling and their relative performance can be assessed by comparing the time-determining steps of CD-SOS-MP2 and THC-RI-SOS-MP2 against LT-RI-SOS-MP2. For LT-RI-SOS-MP2 we assume an operation count proportional to $N_{\text{occ}}N_{\text{virt}}N_{\text{aux}}^2$, for CD-SOS-MP2 an operation count of $2N_{\text{occ}}N_{\text{virt}}N_{\text{aux}}N_{\text{chol}}+N_{\text{occ}}N_{\text{virt}}N_{\text{chol}}^2/2$ and for THC-RI-SOS-MP2 an operation count of $N_{\text{occ}}N_{\text{virt}}N_{\text{aux}}N_{\text{grid}}$. As explained by

Aquilante *et al.*[S5] the relative performance of CD-SOS-MP2 against LT-RI-SOS-MP2 is governed by the ratio of $\sigma = 2\,N_{\text{chol}}/N_{\text{aux}} + \left(N_{\text{chol}}/N_{\text{aux}}\right)^2/2$ to the number of Laplace points N_κ. For reasonable accuracy the ratio $N_{\text{chol}}/N_{\text{aux}}$ should be greater or equal to one and assuming $N_{\text{chol}}/N_{\text{aux}} = 1$ simplifies σ to $\sigma = 2.5$. Assuming that $N_{\text{chol}}/N_{\text{aux}} = 1$ results in µH accuracy of the final energies, for which in LT-RI-SOS-MP2 7–10 Laplace points would be needed, it can be argued that CD-SOS-MP2 requires roughly 2–4 times less operations. The quartic scaling step of the THC-RI-SOS-MP2 method has to be carried out only once, while in LT-RI-SOS-MP2 this step has to be performed for every Laplace point. Assuming that the number of THC grid points is three times greater than the number of auxiliary functions, this results in a ratio σ' of $\sigma' = 3/N_\kappa$. Assuming that again 7–10 Laplace points are used, THC-RI-SOS-MP2 offers an operation count advantage of roughly 2–3 times and is thus comparable to CD-SOS-MP2. Additionally, one of the strengths of THC-RI-SOS-MP2 is that – if all integral operations are performed on-the-fly – the storage requirements are an order of magnitude smaller than for the other two methods. Furthermore, once the THC factorization is obtained all subsequent steps of evaluating the SOS-MP2 energy are only cubic scaling.

5 Canonical MOs vs Cholesky MOs

The effect of using local Cholesky- and Cholesky pseudo-MOs for the formation of the THC factorized ERIs and the computation of the MP2 energy, respectively, is demonstrated. As outlined in the main part of the paper, our screening routines for both the transformation of the RI integrals into Cholesky-MO basis and the subsequent projection onto the THC grids by contraction with the collocation matrices $\mathbf{X}$ rely on the locality of the MOs used. Furthermore, the evaluation of the MP2 energy relies on sparse linear algebra and therefore also benefits from the locality of the Cholesky pseudo-MOs. To highlight this, the following figures compare computation times for our THC-ω-RI-CDD-SOS-MP2 method, split into contributions from obtaining the THC factorization and the evaluation of the SOS-MP2 energy, using canonical MOs (red) against using Cholesky (pseudo)-MOs (black).

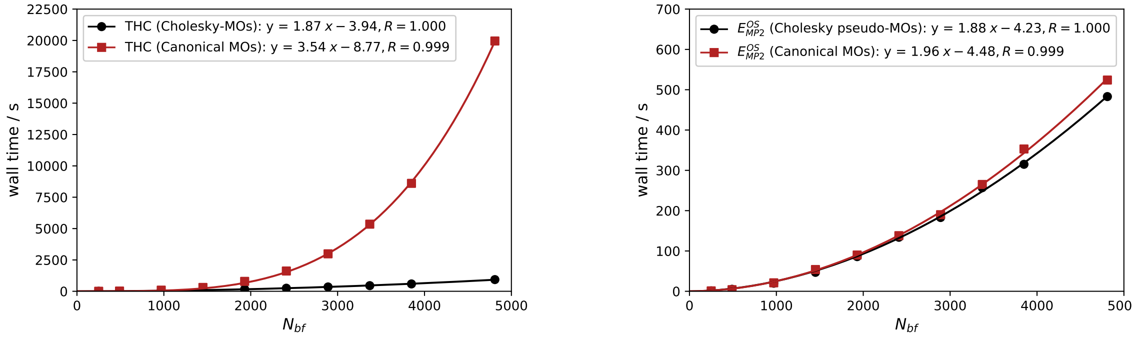


Figure S1: Total execution time of the THC algorithm (left) and the evaluation of the SOS-MP2 energy (right) using Cholesky (pseudo)-MOs (black) and canonical MOs (red) for linear alkanes C_nH_{2n+2} using the cc-pVDZ basis set. The scaling for the largest molecules was determined by linear regression starting from $C_{60}H_{122}$ ($N_{\mathrm{bf}} = 1450$) up to $C_{200}H_{402}$ ($N_{\mathrm{bf}} = 4810$).

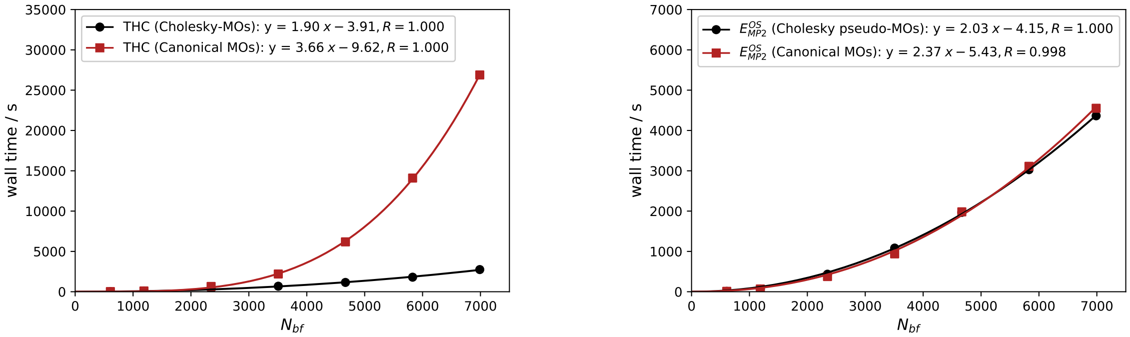


Figure S2: Total execution time of the THC algorithm (left) and the evaluation of the SOS-MP2 energy (right) using Cholesky (pseudo)-MOs (black) and canonical MOs (red) for linear alkanes C_nH_{2n+2} using the cc-pVTZ basis set. The scaling for the largest molecules was determined by linear regression starting from $C_{60}H_{122}$ ($N_{\mathrm{bf}} = 3508$) up to $C_{120}H_{242}$ ($N_{\mathrm{bf}} = 6988$).

The close to quartic scaling of obtaining the THC factorization when using canonical MOs is expected since the non-locality of the MOs causes the contraction of the collocation matrices $\mathbf{X}$ with the three-center RI integrals to approach its formal $\mathcal{O}(N^4)$ scaling. This clearly highlights the efficiency of our screening algorithms for the construction of the $\mathbf{Z}$ matrix. For the energy evaluation, Cholesky pseudo-MOs only provide a speedup for lines 6 and 7 of algorithm 4, i.e., for contractions of the type $\mathbf{X}\mathbf{X}^{\mathrm{T}}$. The product $\mathbf{A} = \underline{\mathbf{X}}\,\underline{\mathbf{X}}^{\mathrm{T}}$ is essentially a representation of the occupied density matrix in the THC grid basis. Analogously, $\mathbf{B} = \overline{\mathbf{X}}\overline{\mathbf{X}}^{\mathrm{T}}$ is a representation of the virtual density matrix. As such $\mathbf{A}$ and $\mathbf{B}$ possess a similar sparsity pattern as the regular density matrices and more importantly should be invariant under rotation of the MOs. The latter means, that $\mathbf{A}$ constructed as $\mathbf{A} = \sum_i \underline{X}_i^P \underline{X}_i^R$ using Cholesky pseudo-MOs and $\mathbf{A}'$ constructed from regular MOs as $\mathbf{A}' = \sum_i e^{\varepsilon_i t^{(\kappa)}} X_i^P X_i^R$ are identical. For this reason, the only step of the evaluation of the MP2 Coulomb-like energy that can be speeded up is the formation of intermediates $\mathbf{A}$ and $\mathbf{B}$, while all subsequent contractions are identical for both types of MOs. Since the contraction $\mathbf{D} = \mathbf{C}\mathbf{Z}$ from algorithm 4 is the time-determining step and $\mathbf{C}$ will have the same sparsity pattern irrespective of the choice of MOs, the overall computation times for the Coulomb-like MP2 energy are very similar (see right sides of figures S1 and S2).

6 Detailed Results for the S22 and L7 Test Set

6.1 Results for the S22 Test Set

Table S1: Absolute energies for the S22 test set monomers computed with our implementation of the RI-SOS-MP2 ($E^{\mathrm{RI}}_{\mathrm{SOS\text{-}MP2}}$) reference method as well as the THC-RI-CDD-SOS-MP2 ($E^{\mathrm{THC\text{-}RI\text{-}CDD}}_{\mathrm{SOS\text{-}MP2}}$) and the THC-$\omega$-RI-CDD-SOS-MP2 ($E^{\mathrm{THC\text{-}\omega\text{-}RI\text{-}CDD}}_{\mathrm{SOS\text{-}MP2}}$) method for the cc-pVDZ basis set. Additionally the deviations (ΔE) of the results from the THC methods to the RI-SOS-MP2 energies are shown.

molecule	$E^{\mathrm{RI}}_{\mathrm{SOS\text{-}MP2}}$ / H	$E^{\mathrm{THC\text{-}RI\text{-}CDD}}_{\mathrm{SOS\text{-}MP2}}$ / H	ΔE / µH	$E^{\mathrm{THC\text{-}\omega\text{-}RI\text{-}CDD}}_{\mathrm{SOS\text{-}MP2}}$ / H	ΔE / µH
2-aminopyridine	−302.703 924	−302.703 927	2.6	−302.703 906	18.5
2-pyridoxine	−322.546 423	−322.546 425	2.1	−322.546 406	16.9
adenine	−465.932 237	−465.932 233	3.1	−465.932 201	35.2
ammonia	−56.384 925	−56.384 925	<0.1	−56.384 928	3.0
benzene	−231.485 902	−231.485 910	8.1	−231.485 889	12.8
ethene	−78.323 148	−78.323 148	<0.1	−78.323 150	1.5
ethyne	−77.082 330	−77.082 330	<0.1	−77.082 333	3.3
formamide	−169.417 178	−169.417 178	0.1	−169.417 173	4.9
formicacid	−189.261 080	−189.261 079	0.4	−189.261 074	5.6
hydrocyanic acid	−93.163 064	−93.163 064	<0.1	−93.163 068	4.5
indole	−362.658 729	−362.658 743	14.4	−362.658 715	13.2
methane	−40.371 390	−40.371 390	<0.1	−40.371 397	6.1
phenol	−306.521 903	−306.521 911	7.5	−306.521 889	14.2
pyrazine	−263.507 029	−263.507 028	1.1	−263.507 005	24.2
thymine	−452.829 481	−452.829 469	12.7	−452.829 455	26.7
uracil	−413.641 115	−413.641 108	7.7	−413.641 090	24.8
water	−76.224 952	−76.224 952	<0.1	−76.224 952	0.4

Table S2: Absolute energies for the S22 test set dimers and complexes computed with our implementation of the RI-SOS-MP2 ($E^{\text{RI}}_{\text{SOS-MP2}}$) reference method as well as the THC-RI-CDD-SOS-MP2 ($E^{\text{THC-RI-CDD}}_{\text{SOS-MP2}}$) and the THC-$\omega$-RI-CDD-SOS-MP2 ($E^{\text{THC-}\omega\text{-RI-CDD}}_{\text{SOS-MP2}}$) method for the cc-pVDZ basis set. Additionally the deviations (ΔE) of the results from the THC methods to the RI-SOS-MP2 energies are shown.

molecule	$E^{\text{RI}}_{\text{SOS-MP2}}$ / H	$E^{\text{THC-RI-CDD}}_{\text{SOS-MP2}}$ / H	ΔE / µH	$E^{\text{THC-}\omega\text{-RI-CDD}}_{\text{SOS-MP2}}$ / H	ΔE / µH
2-pyridoxine + 2-aminopyridine	−625.277 430	−625.277 432	2.1	−625.277 395	35.0
adenine + thymine (stack)	−918.781 508	−918.781 497	11.6	−918.781 494	13.7
adenine + thymine (W.-C.)	−918.788 467	−918.788 458	8.8	−918.788 415	52.6
ammonia dimer	−112.776 724	−112.776 724	<0.1	−112.776 729	4.5
benzene + methane	−271.858 651	−271.858 660	8.6	−271.858 646	5.0
benzene + ammonia	−287.873 967	−287.873 978	10.9	−287.873 962	4.3
benzene dimer (‖)	−462.974 877	−462.974 902	24.8	−462.974 881	4.0
benzene dimer (⊥)	−462.975 643	−462.975 667	24.3	−462.975 635	8.4
benzene + hydrocyanic acid	−324.655 118	−324.655 128	9.9	−324.655 112	6.3
benzene + water	−307.715 837	−307.715 848	12.2	−307.715 830	6.5
ethene dimer	−156.647 632	−156.647 632	0.3	−156.647 630	1.9
ethene + ethyne	−155.407 786	−155.407 786	<0.1	−155.407 789	3.0
formamide dimer	−338.860 613	−338.860 611	2.1	−338.860 603	9.6
formic acid dimer	−378.552 763	−378.552 761	1.6	−378.552 751	12.0
indole + benzene (‖)	−594.150 675	−594.150 713	38.9	−594.150 696	21.2
indole + benzene (⊥)	−594.152 908	−594.152 942	34.5	−594.152 902	5.1
methane dimer	−80.743 019	−80.743 019	<0.1	−80.743 028	8.9
phenol dimer	−613.056 581	−613.056 617	36.4	−613.056 584	3.5
pyrazine dimer	−527.020 386	−527.020 372	13.8	−527.020 350	35.7
uracil dimer (H-bond)	−827.313 142	−827.313 116	25.9	−827.313 084	57.9
uracil dimer (stack)	−827.297 960	−827.297 941	18.3	−827.297 934	25.9
water dimer	−152.460 696	−152.460 696	<0.1	−152.460 695	1.0

Table S3: Absolute energies for the S22 test set monomers computed with our implementation of the RI-SOS-MP2 ($E^{\text{RI}}_{\text{SOS-MP2}}$) reference method as well as the THC-RI-CDD-SOS-MP2 ($E^{\text{THC-RI-CDD}}_{\text{SOS-MP2}}$) and the THC-$\omega$-RI-CDD-SOS-MP2 ($E^{\text{THC-}\omega\text{-RI-CDD}}_{\text{SOS-MP2}}$) method for the cc-pVTZ basis set. Additionally the deviations (ΔE) of the results from the THC methods to the RI-SOS-MP2 energies are shown.

molecule	$E^{\text{RI}}_{\text{SOS-MP2}}$ / H	$E^{\text{THC-RI-CDD}}_{\text{SOS-MP2}}$ / H	ΔE / µH	$E^{\text{THC-}\omega\text{-RI-CDD}}_{\text{SOS-MP2}}$ / H	ΔE / µH
2-aminopyridine	−303.110 460	−303.110 460	0.2	−303.110 465	5.2
2-pyridoxine	−322.966 088	−322.966 092	3.3	−322.966 094	6.0
adenine	−466.526 943	−466.526 939	3.9	−466.526 946	3.0
ammonia	−56.471 514	−56.471 515	<0.1	−56.471 516	1.4
benzene	−231.807 569	−231.807 569	0.7	−231.807 574	5.8
ethene	−78.441 379	−78.441 379	<0.1	−78.441 375	4.5
ethyne	−77.191 318	−77.191 318	<0.1	−77.191 326	7.7
formamide	−169.634 769	−169.634 768	0.8	−169.634 770	1.0
formicacid	−189.491 093	−189.491 093	0.5	−189.491 094	0.3
hydrocyanic acid	−93.276 293	−93.276 293	<0.1	−93.276 305	11.5
indole	−363.150 385	−363.150 381	4.1	−363.150 383	1.6
methane	−40.441 911	−40.441 911	<0.1	−40.441 911	<0.1
phenol	−306.931 762	−306.931 763	0.7	−306.931 766	4.0
pyrazine	−263.847 780	−263.847 780	0.2	−263.847 787	6.6
thymine	−453.403 928	−453.403 930	1.9	−453.403 930	2.8
uracil	−414.156 850	−414.156 843	6.7	−414.156 848	2.3
water	−76.328 144	−76.328 145	<0.1	−76.328 144	0.7

Table S4: Absolute energies for the S22 test set dimers and complexes computed with our implementation of the RI-SOS-MP2 ($E^{\text{RI}}_{\text{SOS-MP2}}$) reference method as well as the THC-RI-CDD-SOS-MP2 ($E^{\text{THC-RI-CDD}}_{\text{SOS-MP2}}$) and the THC-$\omega$-RI-CDD-SOS-MP2 ($E^{\text{THC-}\omega\text{-RI-CDD}}_{\text{SOS-MP2}}$) method for the cc-pVTZ basis set. Additionally the deviations (ΔE) of the results from the THC methods to the RI-SOS-MP2 energies are shown.

molecule	$E^{\text{RI}}_{\text{SOS-MP2}}$ / H	$E^{\text{THC-RI-CDD}}_{\text{SOS-MP2}}$ / H	ΔE / µH	$E^{\text{THC-}\omega\text{-RI-CDD}}_{\text{SOS-MP2}}$ / H	ΔE / µH
2-pyridoxine + 2-aminopyridine	−626.102 751	−626.102 756	3.6	−626.102 760	7.8
adenine + thymine (stack)	−919.953 279	−919.953 283	4.0	−919.953 305	26.0
adenine + thymine (W.-C.)	−919.956 369	−919.956 365	4.3	−919.956 365	3.9
ammonia dimer	−112.948 625	−112.948 625	<0.1	−112.948 624	1.1
benzene + methane	−272.251 679	−272.251 680	0.9	−272.251 685	6.5
benzene + ammonia	−288.282 957	−288.282 956	1.0	−288.282 970	12.5
benzene dimer (‖)	−463.620 183	−463.620 177	6.2	−463.620 192	9.4
benzene dimer (⊥)	−463.619 930	−463.619 927	3.9	−463.619 935	5.3
benzene + hydrocyanic acid	−325.091 880	−325.091 880	<0.1	−325.091 904	23.5
benzene + water	−308.141 724	−308.141 722	1.8	−308.141 728	3.6
ethene dimer	−156.884 286	−156.884 286	<0.1	−156.884 291	5.8
ethene + ethyne	−155.634 944	−155.634 945	1.0	−155.634 952	8.3
formamide dimer	−339.294 087	−339.294 085	2.3	−339.294 087	0.1
formic acid dimer	−379.010 826	−379.010 824	2.4	−379.010 851	24.6
indole + benzene (‖)	−594.967 113	−594.967 112	0.9	−594.967 130	16.6
indole + benzene (⊥)	−594.967 839	−594.967 834	4.7	−594.967 860	20.9
methane dimer	−80.884 303	−80.884 303	<0.1	−80.884 303	0.1
phenol dimer	−613.875 204	−613.875 205	1.2	−613.875 214	9.7
pyrazine dimer	−527.703 274	−527.703 273	1.2	−527.703 298	23.5
uracil dimer (H-bond)	−828.344 550	−828.344 536	14.0	−828.344 537	13.7
uracil dimer (stack)	−828.331 458	−828.331 460	2.1	−828.331 473	15.6
water dimer	−152.665 252	−152.665 252	<0.1	−152.665 250	1.7

6.2 Results for the L7 Test Set

Table S5: Absolute energies for the L7 test set monomers (2nd and 3rd row of each cell) and complexes (1st row of each cell) computed with our implementation of the RI-SOS-MP2 ($E^{\text{RI}}_{\text{SOS-MP2}}$) reference method as well as the THC-RI-CDD-SOS-MP2 ($E^{\text{THC-RI-CDD}}_{\text{SOS-MP2}}$) and the THC-$\omega$-RI-CDD-SOS-MP2 ($E^{\text{THC-}\omega\text{-RI-CDD}}_{\text{SOS-MP2}}$) method for the cc-pVDZ basis set. Additionally the deviations (ΔE) of the results from the THC methods to the RI-SOS-MP2 energies are shown.

molecule	$E^{\text{RI}}_{\text{SOS-MP2}}$ / H	$E^{\text{THC-RI-CDD}}_{\text{SOS-MP2}}$ / H	ΔE / µH	$E^{\text{THC-}\omega\text{-RI-CDD}}_{\text{SOS-MP2}}$ / H	ΔE / µH
circumcoronene	−2528.319 812	−2528.320 097	284.8	−2528.319 995	183.3
+	−465.934 047	−465.934 057	10.7	−465.934 025	21.3
adenine	−2062.358 644	−2062.358 811	167.1	−2062.358 643	1.4
circumcoronene	−2997.209 161	−2997.209 377	215.5	−2997.209 315	153.5
+	−934.805 872	−934.805 864	8.7	−934.805 814	58.6
GC	−2062.358 156	−2062.358 361	205.3	−2062.358 192	36.1
coronene	−1837.897 882	−1837.898 001	118.9	−1837.897 970	87.0
+	−918.932 012	−918.932 097	85.6	−918.932 029	17.4
coronone	−918.932 012	−918.932 097	85.6	−918.932 029	17.4
GC	−1869.612 341	−1869.612 225	115.2	−1869.612 216	124.0
+	−934.793 898	−934.793 872	25.5	−934.793 824	73.7
GC	−934.793 898	−934.793 872	25.5	−934.793 824	73.7
guanine	−1622.921 181	−1622.921 159	21.9	−1622.921 131	50.0
+	−540.978 004	−540.978 009	5.0	−540.977 979	24.9
2×guanine	−1081.938 478	−1081.938 454	24.5	−1081.938 429	49.4
octadecane	−1412.944 380	−1412.944 429	48.8	−1412.944 427	46.8
+	−706.468 122	−706.468 122	0.3	−706.468 123	1.3
octadecane	−706.468 122	−706.468 122	0.3	−706.468 123	1.3
phenylalanine	−2056.636 559	−2056.636 594	35.5	−2056.636 561	1.7
+	−685.518 803	−685.518 816	13.5	−685.518 802	1.0
2×phenylalanine	−1371.076 543	−1371.076 574	30.9	−1371.076 551	7.3

Table S6: Absolute energies for the L7 test set monomers (2nd and 3rd row of each cell) and complexes (1st row of each cell) computed with our implementation of the RI-SOS-MP2 ($E^{\text{RI}}_{\text{SOS-MP2}}$) reference method as well as the THC-RI-CDD-SOS-MP2 ($E^{\text{THC-RI-CDD}}_{\text{SOS-MP2}}$) and the THC-$\omega$-RI-CDD-SOS-MP2 ($E^{\text{THC-}\omega\text{-RI-CDD}}_{\text{SOS-MP2}}$) method for the cc-pVTZ basis set. Additionally the deviations (ΔE) of the results from the THC methods to the RI-SOS-MP2 energies are shown.

molecule	$E^{\text{RI}}_{\text{SOS-MP2}}$ / H	$E^{\text{THC-RI-CDD}}_{\text{SOS-MP2}}$ / H	ΔE / µH	$E^{\text{THC-}\omega\text{-RI-CDD}}_{\text{SOS-MP2}}$ / H	ΔE / µH
circumcoronene	−2531.676 601	−2531.676 470	131.4	−2531.676 470	131.4
+	−466.526 351	−466.526 346	5.7	−466.526 356	4.2
adenine	−2065.114 437	−2065.114 359	77.2	−2065.114 340	96.2
circumcoronene	−3001.162 208	−3001.162 191	17.5	−3001.162 191	17.5
+	−935.987 058	−935.987 033	25.6	−935.987 036	22.1
GC	−2065.114 682	−2065.114 673	9.6	−2065.114 659	23.8
coronene	−1840.399 230	−1840.399 043	187.3	−1840.399 093	137.0
+	−920.176 000	−920.175 898	101.9	−920.175 897	103.2
coronone	−920.176 000	−920.175 898	101.9	−920.175 897	103.2
GC	−1872.005 953	−1872.005 838	115.6	−1872.005 875	77.9
+	−935.988 309	−935.988 264	45.3	−935.988 265	43.9
GC	−935.988 309	−935.988 264	45.3	−935.988 265	43.9
guanine	−1624.970 678	−1624.970 622	55.0	−1624.970 653	24.1
+	−541.658 226	−541.658 214	11.8	−541.658 217	8.2
2×guanine	−1083.305 549	−1083.305 514	35.6	−1083.305 533	16.1
octadecane	−1415.073 527	−1415.073 509	18.1	−1415.073 463	63.9
+	−707.530 305	−707.530 297	8.3	−707.530 314	8.3
octadecane	−707.530 305	−707.530 297	8.3	−707.530 314	8.3
phenylalanine	−2059.357 071	−2059.357 044	27.3	−2059.357 062	8.6
+	−686.425 836	−686.425 827	8.6	−686.425 838	2.5
2×phenylalanine	−1372.890 709	−1372.890 692	17.3	−1372.890 706	3.5

7 Accuracy of THC-MP2 Amplitudes

As shown in the main part of this paper and section 6 of the SI, our THC-ω-RI-CDD-SOS-MP2 method is able to accurately reproduce RI-SOS-MP2 energies with average errors on the order of 10^{-6} H to 10^{-5} H for the investigated S22 and L7 test set. To show that our THC algorithm can also accurately reproduce important intermediate quantities and not only final energies, we compare RI-MP2 amplitudes, given as

$$t_{ij}^{ab}(\text{RI-MP2}) = \sum_{\alpha\beta} \frac{(ia|\alpha)\,[\mathbf{J}^{-1}]_{\alpha\beta}\,(\beta|jb)}{\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b}, \tag{7}$$

against the THC-ω-RI-MP2 amplitudes

$$t_{ij}^{ab}(\text{THC-}\omega\text{-RI-MP2}) = \sum_{PQ} \frac{X_i^P X_a^P Z^{PQ} X_j^Q X_b^Q}{\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b}, \tag{8}$$

as computed with our THC algorithm. The deviation of the THC-ω-RI-MP2 amplitudes from the RI-MP2 amplitudes is quantified by calculating the mean absolute deviation of the diagonal elements given by

$$\Delta t_{ii}^{aa} = \frac{1}{N_\text{occ} N_\text{virt}} \sum_{ia} |t_{ii}^{aa}(\text{RI-MP2}) - t_{ii}^{aa}(\text{THC-}\omega\text{-RI-MP2})| \tag{9}$$

for the molecules and complexes in the S22 test set using the cc-pVDZ and cc-pVTZ basis sets. The same thresholds and settings are used as in the main part of the paper.

Table S7: Deviations of the diagonal elements of the THC-ω-RI-MP2 amplitudes from the RI-MP2 amplitudes for the S22 test set monomers computed with the cc-pVDZ and cc-pVTZ basis sets.

molecule	Δt_{ii}^{aa} (cc-pVDZ) / 10^{-7}	Δt_{ii}^{aa} (cc-pVTZ) / 10^{-7}
2-aminopyridine	2.935	0.345
2-pyridoxine	3.053	0.455
adenine	3.139	0.464
ammonia	3.293	0.725
benzene	2.981	0.565
ethene	1.614	0.343
ethyne	2.742	0.691
formamide	1.571	0.614
formicacid	1.436	0.785
hydrocyanic acid	3.393	1.036
indole	3.194	0.669
methane	1.677	0.325
phenol	2.790	0.276
pyrazine	3.604	0.568
thymine	2.977	0.357
uracil	3.699	0.912
water	4.234	1.122
MAD	2.843	0.603
MAX	4.234	1.122

Table S8: Deviations of the diagonal elements of the THC-ω-RI-MP2 amplitudes from the RI-MP2 amplitudes for the S22 test set dimers and complexes computed with the cc-pVDZ and cc-pVTZ basis sets.

molecule	Δt_{ii}^{aa} (cc-pVDZ) / 10^{-7}	Δt_{ii}^{aa} (cc-pVTZ) / 10^{-7}
2-pyridoxine + 2-aminopyridine	1.529	0.177
adenine + thymine (stack)	1.872	0.180
adenine + thymine (W.-C.)	1.600	0.241
ammonia dimer	0.984	0.235
benzene + methane	2.018	0.366
benzene + ammonia	2.179	0.233
benzene dimer (‖)	2.327	0.145
benzene dimer (⊥)	1.610	0.343
benzene + hydrocyanic acid	2.027	0.260
benzene + water	2.352	0.237
ethene dimer	0.616	0.121
ethene + ethyne	0.761	0.195
formamide dimer	1.669	0.405
formic acid dimer	2.092	0.575
indole + benzene (‖)	2.677	0.133
indole + benzene (⊥)	1.898	0.261
methane dimer	0.651	0.936
phenol dimer	2.358	0.128
pyrazine dimer	2.235	0.271
uracil dimer (H-bond)	2.003	0.379
uracil dimer (stack)	2.253	0.214
water dimer	1.520	0.320
MAD	1.783	0.219
MAX	2.358	0.936

For the MP2 amplitude deviations similar effects as for the MP2 energies can be observed. The THC factorization provides a quantitatively good approximation of the RI-MP2 amplitudes and deviations with the triple-ζ basis set are smaller due to larger THC grids being employed. Errors of the THC-ω-RI-MP2 amplitudes are on the order of 10^{-7} for the cc-pVDZ basis set and on the order of 10^{-8} for the cc-pVTZ basis set across the whole S22 test set. This was expected since the THC grids[S6] used were optimized to best fit density-fitted ERIs.

8 Validation of the Phosphorus Grids

The computation of DNA fragments requires THC grids for the atom types {H, C, N, O, P}. The grids provided by Martínez and coworkers[S6] were however only optimized for 1st- and 2nd-row elements. To still make computations of DNA fragments possible we chose to use the largest atom grids, i.e., the grids for fluorine, for the phosphorus atoms as well. To demonstrate that this does not result in an additional error beyond the regular THC error, we compare the Coulomb-like energies of the THC-RI-CDD-MP2 method against the RI-MP2 reference method for a selection of representative phosphorus species. This mini test set includes P_4, the trivalent PH_3 and the pentavalent H_3PO_4. H_3PO_4 is of special interest here because the DNA backbone contains structurally comparable phosphate moieties. The results are shown in the table below.

Table S9: Comparison of the Coulomb-like energies $E_{\text{MP2-J}}$ for the THC-RI-CDD-MP2 method against RI-MP2 for the validation of the phosphorus grids. All calculations were performed with the cc-pVDZ basis set and the cc-pVDZ-RI fitting basis set.

molecule	$E_{\text{MP2-J}}^{\text{RI}}$ / H	$E_{\text{MP2-J}}^{\text{THC-RI-CDD}}$ / H	ΔE / µH
P_4	-0.357244658	-0.357244828	0.17
PH_3	-0.115739227	-0.115739217	0.01
H_3PO_4	-0.637504139	-0.637504382	0.24

As can be seen from table S9, the errors are well below 1 µH and comparable to the regular THC error for the other atom types. This justifies the use of the fluorine grids for phosphorus atoms and shows that this does not lead to a significant error increase.

9 Demonstrating Continuousness of the Potential Energy Surface for a C−C Bond Rotation in Vitamin K_2

To demonstrate, that the THC-ω-RI-CDD-SOS-MP2 method is not prone to discontinuities in the potential energy surface, the rotation around a C−C bond in vitamin K_2 (menachinon-4), we computed the THC-ω-RI-CDD-SOS-MP2 energies for the rotation around the C−C bond between atoms 4 and 5 in steps of 10°. The same settings were used as in the main part of the paper, i.e., a natural blocking threshold of 10^{-6} and an attenuation strength $\omega = 0.1$. The corresponding energy diagram and additionally the structures at a rotation of -180° and 0° are shown in the figure below.

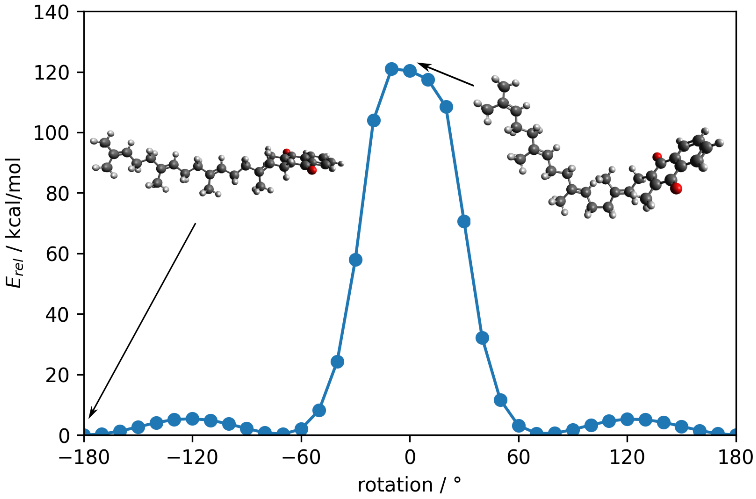


Figure S3: THC-ω-RI-CDD-SOS-MP2 energies relative to the geometry with a rotation angle of -180° of vitamin K_2. All calculations were performed with the cc-pVDZ basis set and the cc-pVDZ-RI fitting basis set.

The geometries of the Vitamin K_2 molecules are provided on our website.

References

(S1) Glasbrenner, M.; Graf, D.; Ochsenfeld, C. Efficient reduced-scaling second-order Møller-Plesset perturbation theory with Cholesky-decomposed densities and an attenuated Coulomb metric. *J. Chem. Theory Comput.* **2020**, *16*, 6856–6868.

(S2) 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.

(S3) Thompson, T. H.; Ochsenfeld, C. Integral partition bounds for fast and effective screening of general one-, two-, and many-electron integrals. *J. Chem. Phys.* **2019**, *150*, 044101.

(S4) 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.

(S5) Aquilante, F.; Pedersen, T. B. Quartic scaling evaluation of canonical scaled opposite spin second-order Møller-Plesset correlation energy using Cholesky decompositions. *Chem. Phys. Lett.* **2007**, *449*, 354–357.

(S6) Kokkila Schumacher, S. I.; Hohenstein, E. G.; Parrish, R. M.; Wang, L. P.; Martínez, T. J. Tensor Hypercontraction Second-Order Møller-Plesset Perturbation Theory: Grid Optimization and Reaction Energies. *J. Chem. Theory Comput.* **2015**, *11*, 3042–3052.

3.3 Efficient low-scaling computation of NMR shieldings at the second-order Møller–Plesset perturbation theory level with Cholesky-decomposed densities and an attenuated Coulomb metric

M. Glasbrenner, S. Vogler, C. Ochsenfeld

J. Chem. Phys. **155**, 224107 (2021).

Abstract: A method for the computation of nuclear magnetic resonance (NMR) shieldings with second-order Møller–Plesset perturbation theory (MP2) is presented which allows to efficiently compute the entire set of shieldings for a given molecular structure. The equations are derived using Laplace-transformed atomic orbital second-order Møller–Plesset perturbation theory as a starting point. The Z-vector approach is employed for minimizing the number of coupled-perturbed self-consistent-field equations that need to be solved. In addition, the method uses the resolution-of-the-identity approximation with an attenuated Coulomb metric and Cholesky decomposition of pseudo-density matrices. The sparsity in the three-center integrals is exploited with sparse linear algebra approaches, leading to reduced computational cost and memory demands. Test calculations show that the deviations from NMR shifts obtained with canonical MP2 are small if appropriate thresholds are used. The performance of the method is illustrated in calculations on DNA strands and on glycine chains with up to 283 atoms and 2864 basis functions.

Reproduced from:

M. Glasbrenner, S. Vogler, C. Ochsenfeld.
"Efficient low-scaling computation of NMR shieldings at the second-order Møller–Plesset perturbation theory level with Cholesky-decomposed densities and an attenuated Coulomb metric."
J. Chem. Phys. **155**, 224107 (2021).

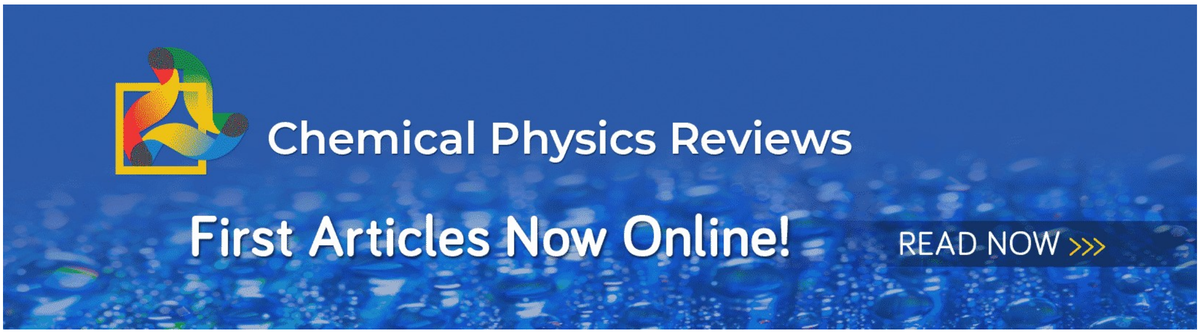
with the permission of AIP Publishing.
`https://aip.scitation.org/doi/full/10.1063/5.0069956`

The Journal of Chemical Physics

Efficient low-scaling computation of NMR shieldings at the second-order Møller–Plesset perturbation theory level with Cholesky-decomposed densities and an attenuated Coulomb metric

Cite as: J. Chem. Phys. **155**, 224107 (2021); https://doi.org/10.1063/5.0069956
Submitted: 02 September 2021 • Accepted: 08 November 2021 • Published Online: 10 December 2021

Michael Glasbrenner, Sigurd Vogler and Christian Ochsenfeld



J. Chem. Phys. **155**, 224107 (2021); https://doi.org/10.1063/5.0069956 **155**, 224107

The Journal of Chemical Physics ARTICLE scitation.org/journal/jcp

Efficient low-scaling computation of NMR shieldings at the second-order Møller–Plesset perturbation theory level with Cholesky-decomposed densities and an attenuated Coulomb metric

Cite as: J. Chem. Phys. 155, 224107 (2021); doi: 10.1063/5.0069956
Submitted: 2 September 2021 • Accepted: 8 November 2021 •
Published Online: 10 December 2021

Michael Glasbrenner,[1] Sigurd Vogler,[1] and Christian Ochsenfeld[1,2,a)]

AFFILIATIONS

[1] Chair of Theoretical Chemistry, Department of Chemistry, University of Munich (LMU), Butenandtstr. 7, D-81377 Munich, Germany
[2] Max Planck Institute for Solid State Research, Heisenbergstr. 1, D-70569 Stuttgart, Germany

[a)] **Author to whom correspondence should be addressed:** christian.ochsenfeld@uni-muenchen.de

ABSTRACT

A method for the computation of nuclear magnetic resonance (NMR) shieldings with second-order Møller–Plesset perturbation theory (MP2) is presented which allows to efficiently compute the entire set of shieldings for a given molecular structure. The equations are derived using Laplace-transformed atomic orbital second-order Møller–Plesset perturbation theory as a starting point. The Z-vector approach is employed for minimizing the number of coupled-perturbed self-consistent-field equations that need to be solved. In addition, the method uses the resolution-of-the-identity approximation with an attenuated Coulomb metric and Cholesky decomposition of pseudo-density matrices. The sparsity in the three-center integrals is exploited with sparse linear algebra approaches, leading to reduced computational cost and memory demands. Test calculations show that the deviations from NMR shifts obtained with canonical MP2 are small if appropriate thresholds are used. The performance of the method is illustrated in calculations on DNA strands and on glycine chains with up to 283 atoms and 2864 basis functions.

I. INTRODUCTION

Nuclear magnetic resonance (NMR) is a widely used spectroscopic method in the field of chemistry. NMR spectra are highly sensitive to the molecular geometry and contain a wealth of structural information. However, especially for large molecules, their interpretation can be difficult, and it can be challenging to unambiguously assign a molecular structure to a given spectrum. In such situations, the comparison with theoretically computed spectra can be very helpful. For this reason, much work has been done on the development of quantum-chemical methods for the simulation of NMR shieldings (for reviews, see, e.g., Refs. 1–3).

Methods for computing NMR shieldings with Hartree–Fock (HF) theory[4–8] or density functional theory (DFT)[9–11] can often provide useful shifts at moderate computational cost. For these theories, linear and sublinear scaling implementations have been developed which allow to apply them to large molecules with 1000 atoms or more.[12–14] If higher accuracy is desired, wave-function based correlation methods can be used. This includes second-order Møller–Plesset perturbation theory (MP2),[15,16] the multiconfigurational self-consistent field method (MCSCF),[17,18] or approaches based on coupled-cluster theory.[19–21] Among the mentioned correlated methods, MP2 is particularly attractive because it has lower formal scaling and computational cost than high-level coupled cluster methods but has been shown to yield typically NMR shifts of significantly higher quality than Hartree–Fock theory or DFT.[16,22] Therefore, a significant amount of work has been done on MP2-based methods for computing NMR shieldings. In 1992, Gauss presented the first method for computing NMR shifts with MP2[15,16] and a Z-vector approach.[23,24] In later years, a more efficient integral-direct implementation by Kollwitz and Gauss[25] and implementations using local-correlation approximations[26–28] were developed. Maurer and Ochsenfeld[29] demonstrated that a linear and sublinear scaling computation of NMR shieldings is possible with atomic-orbital MP2 (AO-MP2)[30–35] and integral screening approaches. Apart from pure MP2 theory, closely related methods, such as double-hybrid density functional theory[36,37] or spin-component-scaled[38,39] and scaled-opposite-spin MP2,[39,40] showed good accuracy in benchmark studies and proved to be promising for efficient calculations on large molecules considering that the algorithmic developments for MP2 can easily be transferred to these methods.

In this work, we present a new method for efficiently computing NMR shieldings at the MP2 level of theory. Like the related approach from the study by Maurer and Ochsenfeld,[29] the method presented in this work is based on Laplace-transformed AO-MP2 but differs in several aspects. First of all, our new method employs an *all-nuclei* formulation, which allows us to efficiently compute the entire set of NMR shieldings for all nuclei in a given molecule. In contrast, Maurer and Ochsenfeld[29] used a *selected-nuclei* formulation, which was first introduced by Beer *et al.*[14] at the SCF level of theory. The selected-nuclei formulation allows to exploit locality and accelerate the computation of the shieldings of a small subset of all nuclei. However, if all shieldings need to be computed, this formulation is inferior due to the large number of coupled-perturbed self-consistent field (CPSCF) equations that need to be solved. The working equations differ significantly between the all-nuclei and selected nuclei ansatz because certain terms involving derivatives of gauge-including atomic orbitals (GIAOs)[4,41,42] are present in one formulation and absent in the other one. GIAOs are used in our method in order to ensure gauge-origin independence. Compared to the method reported by Maurer and Ochsenfeld,[29] our new implementation avoids four-center electron repulsion integrals (ERIs) with the help of the resolution-of-the-identity (RI)[43,44] approximation. Like in our previous work on the efficient computation of MP2 energies,[45] an attenuated Coulomb metric[46–48] is used in order to further increase the available sparsity in the three-center integrals. In this context, we would like to mention an efficient MP2 shielding method using a Cholesky decomposition (CD) of ERIs that was very recently introduced by Burger *et al.*[49] As RI, the CD approach allows to avoid four-center integrals. Compared to RI, it has certain advantages, such as rigorous control of the numerical error of the decomposition and independence of pre-defined auxiliary basis sets.[50,51] However, the Cholesky factors have lower sparsity than three-center RI integrals with an attenuated Coulomb metric for large systems.

Here we use a pivoted Cholesky decomposition[52–54] of pseudo-density matrices which allows us to utilize the reduced rank of these matrices (especially for the occupied pseudo-densities) and further lowers computational cost and memory demands. The sparsity of the three-center integrals is exploited using sparse linear algebra approaches. The transformations of the three-center integrals in the atomic orbital basis with the Cholesky molecular orbitals (Cholesky-MOs) and with pseudo-density and perturbed pseudo-density matrices are accelerated using block-sparse matrix multiplications. For the subsequent steps involving three-center integrals, we use the natural blocking format,[45,47,55] which allows us to reduce the computational effort for contractions of the three-center integrals and the I/O overhead. We demonstrate in test calculations that the developed method provides accurate NMR shieldings if the same thresholds are used that also proved to be appropriate for achieving chemically accurate energies in our recent work on MP2 correlation energies.[45] In calculations on glycine chains and DNA strands, we analyze the efficiency of the method, the scaling, and the requirements on memory and disk space. We also compare calculations with standard MP2 to shielding calculations with scaled-opposite spin MP2 (SOS-MP2)[39,40] and show that the SOS-approximation leads to further substantial gains in efficiency.

II. THEORY

In Sec. II A, a brief review of the AO-MP2 method for energies is provided, which provides the starting point for the derivation. Section II B contains a discussion of AO-MP2 gradients with emphasis on the gradient with respect to the nuclear magnetic moment. In Sec. II C, the mixed second derivative of the energy with respect to nuclear magnetic moment and magnetic field is derived, which allows to compute NMR shieldings. In Sec. II D, a nested Z-vector approach is described, which allows to circumvent the explicit computation of the nuclear magnetic moment derivative of the density matrix and the second derivative of the density matrix and thus minimizes the number of CPSCF equations that need to be solved. Then, in Sec. II E, we discuss strategies for accelerating the time-dominating steps of the computation.

A. Review of atomic-orbital MP2

In general, the elements of the NMR shielding tensor $\boldsymbol{\sigma}^A$ of a nucleus A can be computed by taking a mixed second derivative of the electronic energy E. One derivative needs to be taken with respect to one of the components of the magnetic field B_s with $s \in \{x, y, z\}$, while the other derivative is taken with respect to a component of the nuclear magnetic moment vector m_r^A,

$$\sigma_{rs}^A = \frac{\partial^2}{\partial m_r^A \partial B_s} E\Big|_{\mathbf{m}^A=0,\mathbf{B}=0}. \tag{1}$$

Different quantum-chemical methods can be used for computing E; the method presented in this work is based on atomic-orbital MP2 (AO-MP2).[30,32,34,35] Starting from the closed-shell MP2 energy expression in terms of canonical, doubly occupied Hartree–Fock orbitals,

$$E_{\mathrm{MP2}} = -\sum_{ijab} \frac{(ia|jb)[2(ia|jb) - (ib|ja)]}{\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j}, \tag{2}$$

AO-MP2 can be derived by expressing the orbital energy denominator from Eq. (2) using a Laplace transformation, which can be approximated using numerical integration[30] as follows:

$$\frac{1}{\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j} = \int_0^\infty dt\, e^{-t(\epsilon_a+\epsilon_b-\epsilon_i-\epsilon_j)} \approx \sum_\alpha \omega_\alpha\, e^{-t_\alpha(\epsilon_a+\epsilon_b-\epsilon_i-\epsilon_j)}, \tag{3}$$

where t_α and ω_α are the exponent and weight of Laplace point α, respectively. After inserting the Laplace-transformed denominator into Eq. (2), insertion of the basis set expansion of the MOs, and subsequent summation over all MO-indices, the energy expression of AO-MP2 is obtained as follows:

$$E_{\mathrm{MP2}}^{\mathrm{AO}} = -\sum_{\alpha}\sum_{\mu\nu\lambda\sigma}\left(\underline{\mu}\overline{\nu}|\underline{\lambda}\overline{\sigma}\right)[2(\mu\nu|\lambda\sigma) - (\mu\sigma|\lambda\nu)], \tag{4}$$

where

$$\left(\underline{\mu}\overline{\nu}|\underline{\lambda}\overline{\sigma}\right) = \sum_{\mu'\nu'\lambda'\sigma'} \underline{P}_{\mu\mu'}\overline{P}_{\nu\nu'}\underline{P}_{\lambda\lambda'}\overline{P}_{\sigma\sigma'}\left(\mu'\nu'|\lambda'\sigma'\right) \tag{5}$$

and

$$\underline{P}_{\mu\mu'} = \sqrt[4]{\omega_\alpha}\sum_{i} C_{\mu i}e^{+\epsilon_i t_\alpha}C_{\mu' i}, \tag{6}$$

$$\overline{P}_{\nu\nu'} = \sqrt[4]{\omega_\alpha}\sum_{a} C_{\nu a}e^{-\epsilon_a t_\alpha}C_{\nu' a}. \tag{7}$$

The matrices $\underline{\mathbf{P}}$ and $\overline{\mathbf{P}}$ defined in Eqs. (6) and (7) are called the occupied and virtual pseudo-densities, respectively. Note that the pseudo-density matrices depend on the Laplace point α; for simplicity, we omit this dependence in all formulas shown below.

B. Gradient with respect to the nuclear magnetic moment

In Ref. 56, Schweizer *et al.* derived the gradient of AO-MP2. One possible expression for the gradient of AO-MP2 with respect to a general perturbation ξ reads as follows:

$$\begin{aligned}\frac{\partial}{\partial\xi}E_{\mathrm{MP2}}^{\mathrm{AO}} = &-2\sum_{\alpha}\left\{\mathrm{Tr}\left[\underline{\mathbf{R}}\overline{\mathbf{P}}^{\xi}\right] + \mathrm{Tr}\left[\overline{\mathbf{R}}\underline{\mathbf{P}}^{\xi}\right]\right\}\\ &-2\sum_{\alpha}\sum_{\mu\nu\lambda\sigma}\left(\underline{\mu}\overline{\nu}|\underline{\lambda}\overline{\sigma}\right)\left[2(\mu\nu|\lambda\sigma)^{\xi} - (\mu\nu|\lambda\sigma)^{\xi}\right],\end{aligned} \tag{8}$$

where

$$\overline{R}_{\mu\mu'} = \sum_{\nu\lambda\sigma}\left(\mu'\overline{\nu}|\underline{\lambda}\overline{\sigma}\right)[2(\mu\nu|\lambda\sigma) - (\mu\sigma|\lambda\nu)], \tag{9}$$

$$\underline{R}_{\nu\nu'} = \sum_{\mu\lambda\sigma}\left(\underline{\mu}\nu'|\underline{\lambda}\overline{\sigma}\right)[2(\mu\nu|\lambda\sigma) - (\mu\sigma|\lambda\nu)]. \tag{10}$$

Due to the pseudo-densities appearing in Eqs. (9) and (10), the matrices $\underline{\mathbf{R}}$ and $\overline{\mathbf{R}}$ are also dependent on the Laplace point. For the derivation of the working equations of our new MP2-NMR method, we will take the gradient with respect to the nuclear magnetic moment $\mathbf{m}^A$ as a starting point,

$$\frac{\partial}{\partial\mathbf{m}}E_{\mathrm{MP2}}^{\mathrm{AO}} = -2\sum_{\alpha}\left\{\mathrm{Tr}\left[\underline{\mathbf{R}}\overline{\mathbf{P}}^{\mathbf{m}}\right] + \mathrm{Tr}\left[\overline{\mathbf{R}}\underline{\mathbf{P}}^{\mathbf{m}}\right]\right\}. \tag{11}$$

Here and in the following, $\mathbf{m}^A$ shall be abbreviated as $\mathbf{m}$ in order to simplify the notation. Note that for this particular perturbation, the last term from Eq. (8) is zero because the two-electron integrals do not depend on $\mathbf{m}$. Equation (11) requires perturbed pseudo-density matrices. Explicit expressions for the unperturbed pseudo-density matrices in terms of MO coefficients and orbital energies are shown in Eqs. (6) and (7); these matrices can also be computed only from quantities in the AO basis[33,56,57] as follows:

$$\underline{\mathbf{P}} = \sqrt[4]{\omega_\alpha}e^{t\mathbf{PF}}\mathbf{P}, \tag{12}$$

$$\overline{\mathbf{P}} = \sqrt[4]{\omega_\alpha}e^{-t\mathbf{P}_{\mathrm{virt}}\mathbf{F}}\mathbf{P}_{\mathrm{virt}}, \tag{13}$$

where $\mathbf{P} = \sum_i C_{\mu i}^{*}C_{\nu i}$ is the HF density matrix and $\mathbf{P}_{\mathrm{virt}} = \sum_a C_{\mu a}^{*}C_{\nu a}$ is the virtual density matrix. After inserting the expressions from Eqs. (12) and (13) into Eq. (11) and differentiating using the product rule, one obtains

$$\begin{aligned}\frac{\partial}{\partial\mathbf{m}}E_{\mathrm{MP2}}^{\mathrm{AO}} = &-2\sum_{\alpha}\sqrt[4]{\omega_\alpha}\left\{\mathrm{Tr}\left[\overline{\mathbf{R}}\left(e^{t\mathbf{PF}}\right)^{\mathbf{m}}\mathbf{P}\right] + \mathrm{Tr}\left[\overline{\mathbf{R}}e^{t\mathbf{PF}}\mathbf{P}^{\mathbf{m}}\right]\right.\\ &\left. + \mathrm{Tr}\left[\underline{\mathbf{R}}\left(e^{-t\mathbf{P}_{\mathrm{virt}}\mathbf{F}}\right)^{\mathbf{m}}\mathbf{P}_{\mathrm{virt}}\right] + \mathrm{Tr}\left[\underline{\mathbf{R}}e^{-t\mathbf{P}_{\mathrm{virt}}\mathbf{F}}\mathbf{P}_{\mathrm{virt}}^{\mathbf{m}}\right]\right\}.\end{aligned} \tag{14}$$

By applying cyclic permutations under the trace, Eq. (14) becomes

$$\begin{aligned}\frac{\partial}{\partial\mathbf{m}}E_{\mathrm{MP2}}^{\mathrm{AO}} = &-2\sum_{\alpha}\sqrt[4]{\omega_\alpha}\left\{\mathrm{Tr}\left[\overline{\mathbf{R}}e^{t\mathbf{PF}}\mathbf{P}^{\mathbf{m}}\right] + \mathrm{Tr}\left[\mathbf{P}\overline{\mathbf{R}}\left(e^{t\mathbf{PF}}\right)^{\mathbf{m}}\right]\right.\\ &\left. + \mathrm{Tr}\left[\underline{\mathbf{R}}e^{-t\mathbf{P}_{\mathrm{virt}}\mathbf{F}}\mathbf{P}_{\mathrm{virt}}^{\mathbf{m}}\right] + \mathrm{Tr}\left[\mathbf{P}_{\mathrm{virt}}\underline{\mathbf{R}}\left(e^{-t\mathbf{P}_{\mathrm{virt}}\mathbf{F}}\right)^{\mathbf{m}}\right]\right\}.\end{aligned} \tag{15}$$

In general, the perturbed virtual density matrix can be computed as follows:[56]

$$\mathbf{P}_{\mathrm{virt}}^{\xi} = -\mathbf{P}^{\xi} - (\mathbf{P} + \mathbf{P}_{\mathrm{virt}})\mathbf{S}^{\xi}\mathbf{S}^{-1}. \tag{16}$$

For the nuclear magnetic moment perturbation, the derivative of $\mathbf{S}$ is zero, and Eq. (16) therefore becomes $\mathbf{P}_{\mathrm{virt}}^{\mathbf{m}} = -\mathbf{P}^{\mathbf{m}}$. The terms involving the perturbed density matrices can be efficiently treated with a Z-vector approach, which will be described in detail in Sec. II D. Here, we will focus on the terms containing perturbed matrix exponentials, which are of the general form $\mathrm{Tr}\left[\mathbf{B}\left(e^{\mathbf{A}}\right)^{\xi}\right]$. As shown in Appendix A, such a term can be rearranged as follows:

$$\mathrm{Tr}\left[\mathbf{B}\left(e^{\mathbf{A}}\right)^{\xi}\right] = \mathrm{Tr}\left[\mathbf{Y}\mathbf{A}^{\xi}\right], \tag{17}$$

where

$$\mathbf{Y} = \sum_{k=1}^{\infty}\sum_{i=0}^{k-1}\frac{1}{k!}\mathbf{A}^{i}\mathbf{B}\mathbf{A}^{k-1-i}. \tag{18}$$

$\mathbf{Y}$ can be computed efficiently by a recursive scheme, which is also shown in Appendix A. In order to rearrange the terms containing perturbed matrix exponentials in Eq. (15), we define two intermediates $\underline{\mathbf{Y}}$ and $\overline{\mathbf{Y}}$. $\overline{\mathbf{Y}}$ is computed by evaluating Eq. (18) with $\mathbf{A} = (t_\alpha\mathbf{PF})$ and $\mathbf{B} = \left(\mathbf{P}\overline{\mathbf{R}}\right)$. $\underline{\mathbf{Y}}$ is calculated using the same formula and setting $\mathbf{A} = (-t_\alpha\mathbf{P}_{\mathrm{virt}}\mathbf{F})$ and $\mathbf{B} = \left(\mathbf{P}_{\mathrm{virt}}\underline{\mathbf{R}}\right)$. Using $\underline{\mathbf{Y}}$ and $\overline{\mathbf{Y}}$, the intermediates $\underline{\mathbf{Y}}_1$, $\underline{\mathbf{Y}}_2$, $\overline{\mathbf{Y}}_1$, and $\overline{\mathbf{Y}}_2$ defined in Ref. 56 can be computed by a small number of additional matrix multiplications as follows:

$$\underline{\mathbf{Y}}_1 = -t_\alpha\mathbf{F}\underline{\mathbf{Y}}, \tag{19}$$

$$\underline{\mathbf{Y}}_2 = -t_\alpha\underline{\mathbf{Y}}\mathbf{P}_{\mathrm{virt}}, \tag{20}$$

$$\overline{\mathbf{Y}}_1 = t_\alpha\mathbf{F}\overline{\mathbf{Y}}, \tag{21}$$

$$\overline{\mathbf{Y}}_2 = t_\alpha\overline{\mathbf{Y}}\mathbf{P}. \tag{22}$$

Instead of evaluating four different recursion formulas for $\underline{\mathbf{Y}}_1$, $\underline{\mathbf{Y}}_2$, $\overline{\mathbf{Y}}_1$, and $\overline{\mathbf{Y}}_2$ as in Ref. 56, only two recursions for $\underline{\mathbf{Y}}$ and $\overline{\mathbf{Y}}$ need to be evaluated if the new recursion formulas presented in this work are used. This reduces the number of necessary matrix multiplications by roughly a factor of two. Note that also all "**Y**"-matrices from Eqs. (19)–(22) implicitly depend on the Laplace point.

Using the intermediates defined so far, we are able to write the formula for the gradient with respect to **m** in the following way:

$$\frac{\partial}{\partial \mathbf{m}} E_{\mathrm{MP2}}^{\mathrm{AO}} = -2\sum_{\alpha}\sqrt[4]{\omega_\alpha}\Big[\mathrm{Tr}[\boldsymbol{\mathcal{F}}^{(\alpha)}\mathbf{h}^{\mathbf{m}}] + \mathrm{Tr}[\boldsymbol{\mathcal{P}}^{(\alpha)}\mathbf{P}^{\mathbf{m}}]\Big], \tag{23}$$

$$\boldsymbol{\mathcal{F}}^{(\alpha)} = \underline{\mathbf{Y}}_2 + \overline{\mathbf{Y}}_2, \tag{24}$$

$$\boldsymbol{\mathcal{P}}^{(\alpha)} = \overline{\mathbf{Y}}_1 - \underline{\mathbf{Y}}_1 + \mathbf{G}\Big[\boldsymbol{\mathcal{F}}^{(\alpha)}\Big] + \overline{\mathbf{R}}e^{t_\alpha \mathbf{PF}} - \underline{\mathbf{R}}e^{-t_\alpha \mathbf{P}_{\mathrm{virt}}\mathbf{F}}. \tag{25}$$

$\mathbf{G}\Big[\boldsymbol{\mathcal{F}}^{(\alpha)}\Big]$ is computed analogously to the two-electron part of the Fock matrix with $\boldsymbol{\mathcal{F}}^{(\alpha)}$ taking the role of the density matrix,

$$\mathbf{G}_{\mu\nu}\Big[\boldsymbol{\mathcal{F}}^{(\alpha)}\Big] = \sum_{\lambda\sigma}[2(\mu\nu|\lambda\sigma) - (\mu\sigma|\lambda\nu)]\mathcal{F}_{\lambda\sigma}^{(\alpha)}. \tag{26}$$

Noticing that only $\boldsymbol{\mathcal{F}}^{(\alpha)}$ and $\boldsymbol{\mathcal{P}}^{(\alpha)}$ depend on the Laplace point in Eq. (23), we define

$$\boldsymbol{\mathcal{F}} \equiv \sum_{\alpha}\sqrt[4]{\omega_\alpha}\boldsymbol{\mathcal{F}}^{(\alpha)}, \tag{27}$$

$$\boldsymbol{\mathcal{P}} \equiv \sum_{\alpha}\sqrt[4]{\omega_\alpha}\boldsymbol{\mathcal{P}}^{(\alpha)}. \tag{28}$$

Using $\boldsymbol{\mathcal{F}}$ and $\boldsymbol{\mathcal{P}}$, Eq. (23) becomes

$$\frac{\partial}{\partial \mathbf{m}} E_{\mathrm{MP2}}^{\mathrm{AO}} = -2\big[\mathrm{Tr}[\boldsymbol{\mathcal{F}}\mathbf{h}^{\mathbf{m}}] + \mathrm{Tr}[\boldsymbol{\mathcal{P}}\mathbf{P}^{\mathbf{m}}]\big]. \tag{29}$$

C. Second derivative with respect to the magnetic field

As shown in Eq. (1), NMR shieldings can be computed as the mixed second derivative of the energy with respect to the nuclear magnetic moment and the magnetic field. The expression from Eq. (29) thus needs to be differentiated with respect to the magnetic field **B** as follows:

$$\begin{aligned}\frac{\partial^2}{\partial \mathbf{B}\partial \mathbf{m}} E_{\mathrm{MP2}}^{\mathrm{AO}} = &-2\big[\mathrm{Tr}[\boldsymbol{\mathcal{F}}\mathbf{h}^{\mathbf{Bm}}] + \mathrm{Tr}[\boldsymbol{\mathcal{F}}^{\mathbf{B}}\mathbf{h}^{\mathbf{m}}]\\ &+ \mathrm{Tr}[\boldsymbol{\mathcal{P}}\mathbf{P}^{\mathbf{Bm}}] + \mathrm{Tr}[\boldsymbol{\mathcal{P}}^{\mathbf{B}}\mathbf{P}^{\mathbf{m}}]\big].\end{aligned} \tag{30}$$

Equation (30) contains the derivative of $\boldsymbol{\mathcal{F}}$ and $\boldsymbol{\mathcal{P}}$, which can be computed as follows:

$$\boldsymbol{\mathcal{F}}^{\mathbf{B}} = \sum_{\alpha}\sqrt[4]{\omega_\alpha}(\boldsymbol{\mathcal{F}}^{(\alpha)})^{\mathbf{B}} = \sum_{\alpha}\sqrt[4]{\omega_\alpha}\Big\{\underline{\mathbf{Y}}_2^{\mathbf{B}} + \overline{\mathbf{Y}}_2^{\mathbf{B}}\Big\}, \tag{31}$$

$$\begin{aligned}\boldsymbol{\mathcal{P}}^{\mathbf{B}} = &\sum_{\alpha}\sqrt[4]{\omega_\alpha}\Big\{\overline{\mathbf{Y}}_1^{\mathbf{B}} - \underline{\mathbf{Y}}_1^{\mathbf{B}} + \overline{\mathbf{R}}^{\mathbf{B}}e^{t_\alpha \mathbf{PF}} + \overline{\mathbf{R}}\big(e^{t_\alpha \mathbf{PF}}\big)^{\mathbf{B}}\\ &- \underline{\mathbf{R}}^{\mathbf{B}}e^{-t_\alpha \mathbf{P}_{\mathrm{virt}}\mathbf{F}} - \underline{\mathbf{R}}\big(e^{-t_\alpha \mathbf{P}_{\mathrm{virt}}\mathbf{F}}\big)^{\mathbf{B}}\Big\}\\ &+ \mathbf{G}^{\mathbf{B}}[\boldsymbol{\mathcal{F}}] + \mathbf{G}\big[\boldsymbol{\mathcal{F}}^{\mathbf{B}}\big].\end{aligned} \tag{32}$$

Equations (31) and (32) require the magnetic field derivative of the "**Y**"-matrices from Eqs. (19)–(22),

$$\overline{\mathbf{Y}}_1^{\mathbf{B}} = t_\alpha\big(\mathbf{F}^{\mathbf{B}}\overline{\mathbf{Y}} + \mathbf{F}\overline{\mathbf{Y}}^{\mathbf{B}}\big), \tag{33}$$

$$\overline{\mathbf{Y}}_2^{\mathbf{B}} = t_\alpha\big(\overline{\mathbf{Y}}^{\mathbf{B}}\mathbf{P} + \overline{\mathbf{Y}}\mathbf{P}^{\mathbf{B}}\big), \tag{34}$$

$$\underline{\mathbf{Y}}_1^{\mathbf{B}} = -t_\alpha\big(\mathbf{F}^{\mathbf{B}}\underline{\mathbf{Y}} + \mathbf{F}\underline{\mathbf{Y}}^{\mathbf{B}}\big), \tag{35}$$

$$\underline{\mathbf{Y}}_2^{\mathbf{B}} = -t_\alpha\big(\underline{\mathbf{Y}}^{\mathbf{B}}\mathbf{P}_{\mathrm{virt}} + \underline{\mathbf{Y}}\mathbf{P}_{\mathrm{virt}}^{\mathbf{B}}\big). \tag{36}$$

Here, $\underline{\mathbf{Y}}^{\mathbf{B}}$ and $\overline{\mathbf{Y}}^{\mathbf{B}}$ can be computed recursively. The corresponding recursion formulas are derived in Appendix B by differentiating the recursion formulas for $\underline{\mathbf{Y}}$ and $\overline{\mathbf{Y}}$. For all recursion formulas, asymptotic linear scaling can be achieved by employing block-sparse matrices. The matrix $\mathbf{G}^{\mathbf{B}}[\boldsymbol{\mathcal{F}}]$ is computed in a similar way to $\mathbf{G}[\boldsymbol{\mathcal{F}}]$ in Eq. (26) with the difference that the magnetic field derivatives of the ERIs are used instead of the standard unperturbed ERIs,

$$\mathbf{G}_{\mu\nu}^{\mathbf{B}}[\boldsymbol{\mathcal{F}}] = \sum_{\lambda\sigma}\Big[2\,(\mu\nu|\lambda\sigma)^{(\mathbf{B})} - (\mu\sigma|\lambda\nu)^{(\mathbf{B})}\Big]\mathcal{F}_{\lambda\sigma}. \tag{37}$$

For an efficient calculation of $\mathbf{G}^{\mathbf{B}}[\boldsymbol{\mathcal{F}}]$ and $\mathbf{G}[\boldsymbol{\mathcal{F}}^{\mathbf{B}}]$ in Eq. (32) and of $\mathbf{G}[\boldsymbol{\mathcal{F}}]$ in Eq. (26), any approach for efficient Fock matrix construction in Hartree–Fock theory can be used, such as the Continuous Fast Multipole Method (CFMM)[58,59] or the RI-J method[60] for the Coulomb matrix and LinK screening[61,62] or semi-numerical exchange approaches[63] for the exchange matrix.

Equations (31) and (32) also require the magnetic field derivative of $\underline{\mathbf{R}}$ and $\overline{\mathbf{R}}$, which can be calculated as follows:

$$\begin{aligned}\underline{R}_{\nu\nu'}^{\mathbf{B}} = \sum_{\mu\lambda\sigma}\Big\{&\Big[\big(\underline{\mu}^{\mathbf{B}}\nu'|\underline{\lambda}\overline{\sigma}\big) + \big(\underline{\mu}\nu'|\underline{\lambda}^{\mathbf{B}}\overline{\sigma}\big) + \big(\underline{\mu}\nu'|\underline{\lambda}\overline{\sigma}^{\mathbf{B}}\big)\Big]\\ &\times[2(\mu\nu|\lambda\sigma) - (\mu\sigma|\lambda\nu)] + \big(\underline{\mu}\nu'|\underline{\lambda}\overline{\sigma}\big)^{(\mathbf{B})}[2(\mu\nu|\lambda\sigma) - (\mu\sigma|\lambda\nu)]\\ &+ \big(\underline{\mu}\nu'|\underline{\lambda}\overline{\sigma}\big)\Big[2(\mu\nu|\lambda\sigma)^{(\mathbf{B})} - (\mu\sigma|\lambda\nu)^{(\mathbf{B})}\Big]\Big\},\end{aligned} \tag{38}$$

$$\begin{aligned}\overline{R}_{\mu\mu'}^{\mathbf{B}} = \sum_{\nu\lambda\sigma}\Big\{&\Big[\big(\mu'\overline{\nu}^{\mathbf{B}}|\underline{\lambda}\overline{\sigma}\big) + \big(\mu'\overline{\nu}|\underline{\lambda}^{\mathbf{B}}\overline{\sigma}\big) + \big(\mu'\overline{\nu}|\underline{\lambda}\overline{\sigma}^{\mathbf{B}}\big)\Big]\\ &\times[2(\mu\nu|\lambda\sigma) - (\mu\sigma|\lambda\nu)] + \big(\mu'\overline{\nu}|\underline{\lambda}\overline{\sigma}\big)^{(\mathbf{B})}[2(\mu\nu|\lambda\sigma) - (\mu\sigma|\lambda\nu)]\\ &+ \big(\mu'\overline{\nu}|\underline{\lambda}\overline{\sigma}\big)\Big[2(\mu\nu|\lambda\sigma)^{(\mathbf{B})} - (\mu\sigma|\lambda\nu)^{(\mathbf{B})}\Big]\Big\}.\end{aligned} \tag{39}$$

Both Eqs. (38) and (39) contain two types of terms: terms involving magnetic field derivatives of ERIs and terms in which the unperturbed ERIs are transformed both with perturbed and unperturbed pseudo-densities. The efficient computation of these terms is discussed in Sec. II E. The notation with $(\mathbf{B})$ as in $\big(\mu'\overline{\nu}|\underline{\lambda}\overline{\sigma}\big)^{(\mathbf{B})}$ means that only the integrals are differentiated, not the pseudo-density or density matrices.

D. Nested Z-vector approach for avoiding $\mathbf{P}^{\mathbf{m}}$ and $\mathbf{P}^{\mathbf{Bm}}$

While Eq. (30) would allow us to compute the NMR shieldings of all nuclei in a molecule, these equations still contain $\mathbf{P}^{\mathbf{m}}$ and $\mathbf{P}^{\mathbf{Bm}}$. The explicit computation of these matrices should be avoided in order to reduce the number of CPSCF equations that need to be solved. This can be achieved using a nested Z-vector approach similar to the one presented by Maurer and Ochsenfeld.[29] In this section, we will present the theory behind this approach in a general way, which can be applied to any formulation of CPSCF. In our implementation, we employ the density matrix based CPSCF (D-CPSCF) method introduced by Ochsenfeld and Head-Gordon;[64] detailed working equations for this formulation are derived in Appendix D.

In general, a CPSCF equation for $\mathbf{P}^{\mathbf{m}}$ has the following structure:

$$\mathbf{A}\mathbf{P}^{\mathbf{m}} = \mathbf{b}^{\mathbf{m}}, \tag{40}$$

where $\mathbf{b}^{\mathbf{m}}$ is the right-hand side of the CPSCF equation and $\mathbf{A}$ is the Hartree–Fock Hessian matrix. Formally, this equation can be solved by multiplying with the inverse of the Hessian from the left as follows:

$$\mathbf{P}^{\mathbf{m}} = \mathbf{A}^{-1}\mathbf{b}^{\mathbf{m}}. \tag{41}$$

Usually, an explicit inversion of $\mathbf{A}$ is circumvented in favor of an iterative solution. In order to compute all matrices $\mathbf{P}^{\mathbf{m}}$ for a given molecule, $3 \times N_{\text{at}}$ CPSCF equations would need to be solved where N_{at} denotes the number of atoms. In a term of the form $\text{Tr}[\mathbf{X}\mathbf{P}^{\mathbf{m}}]$, however, this can avoided by applying a Z-vector approach,[23,24]

$$\text{Tr}[\mathbf{X}\mathbf{P}^{\mathbf{m}}] = \text{Tr}[\mathbf{X}\mathbf{A}^{-1}\mathbf{b}^{\mathbf{m}}] = \text{Tr}[\mathbf{Z}_{X}\mathbf{b}^{\mathbf{m}}], \tag{42}$$

where $\mathbf{Z}_X$ is obtained by solving a CPSCF-like equation with $\mathbf{X}$ as the right-hand side as follows:

$$\mathbf{A}\mathbf{Z}_X = \mathbf{X}. \tag{43}$$

The advantage here is that only a single CPSCF equation needs to be solved. The outlined strategy could be directly applied to the term $\text{Tr}[\boldsymbol{\mathcal{P}}^{\mathbf{B}}\mathbf{P}^{\mathbf{m}}]$; however, as we will discuss in the following, a contribution from the term $\text{Tr}[\boldsymbol{\mathcal{P}}\mathbf{P}^{\mathbf{Bm}}]$ is first added to $\boldsymbol{\mathcal{P}}^{\mathbf{B}}$ in our implementation.

The term $\text{Tr}[\boldsymbol{\mathcal{P}}\mathbf{P}^{\mathbf{Bm}}]$ contains the mixed second derivative of the density matrix. An equation for explicitly computing $\mathbf{P}^{\mathbf{Bm}}$ can be derived from the CPSCF equation $\mathbf{A}\mathbf{P}^{\mathbf{B}} = \mathbf{b}^{\mathbf{B}}$ by differentiating it with respect to $\mathbf{m}$ as follows:

$$\mathbf{A}^{\mathbf{m}}\mathbf{P}^{\mathbf{B}} + \mathbf{A}\mathbf{P}^{\mathbf{Bm}} = \mathbf{b}^{\mathbf{Bm}}. \tag{44}$$

$\mathbf{P}^{\mathbf{Bm}}$ can thus be obtained by solving a CPSCF equation with $(\mathbf{b}^{\mathbf{Bm}} - \mathbf{A}^{\mathbf{m}}\mathbf{P}^{\mathbf{B}})$ as a right-hand side as follows:

$$\mathbf{P}^{\mathbf{Bm}} = \mathbf{A}^{-1}\left(\mathbf{b}^{\mathbf{Bm}} - \mathbf{A}^{\mathbf{m}}\mathbf{P}^{\mathbf{B}}\right). \tag{45}$$

Next, we insert Eq. (45) into the term $\text{Tr}\left[\boldsymbol{\mathcal{P}}\mathbf{P}^{\mathbf{Bm}}\right]$ as follows:

$$\begin{aligned}\text{Tr}\left[\boldsymbol{\mathcal{P}}\mathbf{P}^{\mathbf{Bm}}\right] &= \text{Tr}\left[\boldsymbol{\mathcal{P}}\left(\mathbf{A}^{-1}\left(\mathbf{b}^{\mathbf{Bm}} - \mathbf{A}^{\mathbf{m}}\left[\mathbf{P}^{\mathbf{B}}\right]\right)\right)\right] \\ &= \text{Tr}\left[\mathbf{Z}_{\mathcal{P}}\left(\mathbf{b}^{\mathbf{Bm}} - \mathbf{A}^{\mathbf{m}}\left[\mathbf{P}^{\mathbf{B}}\right]\right)\right].\end{aligned} \tag{46}$$

$\mathbf{Z}_{\mathcal{P}}$ is computed by solving a CPSCF equation with $\boldsymbol{\mathcal{P}}$ as the right-hand side as follows:

$$\mathbf{A}\mathbf{Z}_{\mathcal{P}} = \boldsymbol{\mathcal{P}}. \tag{47}$$

All terms in $\text{Tr}\left[\mathbf{Z}_{\mathcal{P}}\left(\mathbf{b}^{\mathbf{Bm}} - \mathbf{A}^{\mathbf{m}}\left[\mathbf{P}^{\mathbf{B}}\right]\right)\right]$ that depend on $\mathbf{P}^{m}$ need to be rearranged in the form $\text{Tr}[\ldots\mathbf{P}^{\mathbf{m}}]$ such that another Z-vector approach can be applied. A detailed derivation is provided in Appendix D. It is shown that the sum of the result from Eq. (46) and the term $\text{Tr}[\boldsymbol{\mathcal{P}}^{\mathbf{B}}\mathbf{P}^{\mathbf{m}}]$ can be rearranged in the following form:

$$\begin{aligned}&\text{Tr}\left[\mathbf{Z}_{\mathcal{P}}\left(\mathbf{b}^{\mathbf{Bm}} - \mathbf{A}^{\mathbf{m}}\left[\mathbf{P}^{\mathbf{B}}\right]\right)\right] + \text{Tr}\left[\boldsymbol{\mathcal{P}}^{\mathbf{B}}\mathbf{P}^{\mathbf{m}}\right] \\ &\quad = \text{Tr}\left[\mathbf{O}_{\mathbf{F}^{\mathbf{m}}}\mathbf{h}^{\mathbf{m}}\right] + \text{Tr}\left[\mathbf{O}_{\mathbf{Y}^{\mathbf{m}}}\mathbf{h}^{\mathbf{Bm}}\right] + \text{Tr}\left[\mathbf{O}\mathbf{P}^{\mathbf{m}}\right],\end{aligned} \tag{48}$$

where $\mathbf{O}_{\mathbf{F}^{\mathbf{m}}}$, $\mathbf{O}_{\mathbf{Y}^{\mathbf{m}}}$, and $\mathbf{O}$ are defined in Appendix D. The explicit computation of $\mathbf{P}^{\mathbf{m}}$ can be avoided by using a Z-vector $\mathbf{Z}_{\mathbf{O}}$ as follows:

$$\text{Tr}[\mathbf{O}\mathbf{P}^{\mathbf{m}}] = \text{Tr}[\mathbf{Z}_{\mathbf{O}}\mathbf{b}^{\mathbf{m}}], \tag{49}$$

which is obtained by solving a CPSCF equation with $\mathbf{O}$ as the right-hand side as follows:

$$\mathbf{A}\mathbf{Z}_{\mathbf{O}} = \mathbf{O}. \tag{50}$$

With this, we arrive at the final form of the working equation for computing NMR shieldings as follows:

$$\begin{aligned}\boldsymbol{\sigma}^{A} = &-2\,\text{Tr}[(\boldsymbol{\mathcal{F}} + \mathbf{O}_{\mathbf{Y}^{\mathbf{m}}})\mathbf{h}^{\mathbf{Bm}}] - 2\,\text{Tr}[(\boldsymbol{\mathcal{F}}^{\mathbf{B}} + \mathbf{O}_{\mathbf{F}^{\mathbf{m}}})\mathbf{h}^{\mathbf{m}}] \\ &- 2\,\text{Tr}[\mathbf{Z}_{\mathbf{O}}\mathbf{b}^{\mathbf{m}}].\end{aligned} \tag{51}$$

Figure 1 summarizes the described algorithm for computing NMR shieldings and shows in which order the steps of the calculation are carried out in our implementation. With the described nested Z-vector approach, seven CPSCF equations need to be solved in total irrespective of the size of the molecule (one equation for $\mathbf{Z}_{\mathcal{P}}$ and three equations for the magnetic field components of $\mathbf{P}^{\mathbf{B}}$ and $\mathbf{Z}_{\mathbf{O}}$, respectively).

```
compute 2-center and 3-center RI integrals
compute $\mathbf{h}^{\mathbf{m}}$, $\mathbf{h}^{\mathbf{Bm}}$, $\mathbf{S}^{\mathbf{B}}$ and $\mathbf{F}^{(\mathbf{B})}$
solve CPSCF for $\mathbf{P}^{\mathbf{B}}$
compute $\mathbf{F}^{\mathbf{B}}$

for all Laplace-points $\alpha$ do
    compute $\overline{\mathbf{P}}$, $\underline{\mathbf{P}}$, $\overline{\mathbf{P}}^{\mathbf{B}}$ and $\underline{\mathbf{P}}^{\mathbf{B}}$
    compute $\overline{\mathbf{R}}$, $\underline{\mathbf{R}}$, $\overline{\mathbf{R}}^{\mathbf{B}}$ and $\underline{\mathbf{R}}^{\mathbf{B}}$
    recursions for $\underline{\mathbf{Y}}_1$, $\underline{\mathbf{Y}}_2$, $\overline{\mathbf{Y}}_1$, $\underline{\mathbf{Y}}_2$
    recursions for $\underline{\mathbf{Y}}_1^{\mathbf{B}}$, $\underline{\mathbf{Y}}_2^{\mathbf{B}}$, $\overline{\mathbf{Y}}_1^{\mathbf{B}}$, $\underline{\mathbf{Y}}_2^{\mathbf{B}}$
    sum up contributions to $\boldsymbol{\mathcal{F}}$, $\boldsymbol{\mathcal{P}}$, $\boldsymbol{\mathcal{F}}^{\mathbf{B}}$ and $\boldsymbol{\mathcal{P}}^{\mathbf{B}}$
end for

solve Z-vector equation $\mathbf{Z}_{\mathcal{P}} = \mathbf{A}^{-1}\boldsymbol{\mathcal{P}}$
compute $\mathbf{O}$, $\mathbf{O}_{\mathbf{F}^{\mathbf{m}}}$, and $\mathbf{O}_{\mathbf{Y}^{\mathbf{m}}}$
solve Z-vector equation $\mathbf{Z}_{\mathbf{O}} = \mathbf{A}^{-1}\mathbf{O}$

$\boldsymbol{\sigma}^{A} = -2\left\{\text{Tr}[(\boldsymbol{\mathcal{F}} + \mathbf{O}_{\mathbf{Y}^{\mathbf{m}}})\mathbf{h}^{\mathbf{Bm}}] + \text{Tr}[(\boldsymbol{\mathcal{F}}^{\mathbf{B}} + \mathbf{O}_{\mathbf{F}^{\mathbf{m}}})\mathbf{h}^{\mathbf{m}}] + \text{Tr}[\mathbf{Z}_{\mathbf{O}}\mathbf{b}^{\mathbf{m}}]\right\}$
```

FIG. 1. Algorithm for ω-RI-CDD-MP2-NMR.

E. Efficient computation of $\underline{\mathbf{R}}$, $\overline{\mathbf{R}}$, $\underline{\mathbf{R}}^{\mathbf{B}}$, and $\overline{\mathbf{R}}^{\mathbf{B}}$

The computation of the matrices $\underline{\mathbf{R}}$ and $\overline{\mathbf{R}}$ and their magnetic field derivatives (including the computation of the required three-center integrals) is the most time-consuming step of the shielding calculation with our new method. In this section, we describe how these steps can be accelerated using several approximations.

In order to exploit their rank deficiency, we subject the pseudo-density matrices a pivoted Cholesky decomposition,[52–54]

$$\underline{P}_{\mu\nu} = \sum_{\underline{i}} \underline{L}_{\mu\underline{i}} \underline{L}_{\nu\underline{i}}, \tag{52}$$

$$\overline{P}_{\mu\nu} = \sum_{\overline{a}} \overline{L}_{\mu\overline{a}} \overline{L}_{\nu\overline{a}}. \tag{53}$$

This decomposition has been applied before in several methods for MP2[45,65,66] and direct RPA,[67] and it was shown to preserve sparsity. It is especially useful for the occupied pseudo-density as the number of obtained Cholesky vectors is less than or equal to the number of occupied orbitals and therefore much smaller than the number of basis functions (assuming reasonably accurate basis sets). From the Cholesky factors of Eqs. (52) and (53), local pseudo-MOs $\phi_{\underline{i}}$ and $\phi_{\overline{a}}$ can be computed as follows:

$$\phi_{\underline{i}}(\mathbf{r}) = \sum_{\mu} \chi_\mu(\mathbf{r}) \underline{L}_{\mu\underline{i}}, \tag{54}$$

$$\phi_{\overline{a}}(\mathbf{r}) = \sum_{\nu} \chi_\nu(\mathbf{r}) \overline{L}_{\nu\overline{a}}. \tag{55}$$

The pivoted Cholesky decomposition is not applicable to the magnetic field perturbed pseudo-densities because they are not positive semi-definite. For this reason, we leave the perturbed pseudo-densities undecomposed. After inserting the Cholesky-decomposed pseudo-densities, Eqs. (9), (10), (38), and (39) become

$$\underline{R}_{\nu\nu'} = \sum_{\underline{i}\underline{j}\overline{b}} \left(\underline{i}\nu'|\underline{j}\overline{b}\right)\left[2\left(\underline{i}\nu|\underline{j}\overline{b}\right) - \left(\underline{i}\overline{b}|\underline{j}\nu\right)\right], \tag{56}$$

$$\overline{R}_{\mu\mu'} = \sum_{\overline{a}\underline{j}\overline{b}} \left(\mu'\overline{a}|\underline{j}\overline{b}\right)\left[2\left(\mu\overline{a}|\underline{j}\overline{b}\right) - \left(\mu\overline{b}|\underline{j}\overline{a}\right)\right], \tag{57}$$

$$\begin{aligned} \underline{R}^{\mathbf{B}}_{\nu\nu'} =& \sum_{\mu\lambda\sigma}\left[\left(\underline{\mu}^{\mathbf{B}}\nu'|\underline{\lambda}\overline{\sigma}\right) + \left(\underline{\mu}\nu'|\underline{\lambda}^{\mathbf{B}}\overline{\sigma}\right) + \left(\underline{\mu}\nu'|\underline{\lambda}\overline{\sigma}^{\mathbf{B}}\right)\right] \\ &\times [2(\mu\nu|\lambda\sigma) - (\mu\sigma|\lambda\nu)] \\ &+ \sum_{\underline{i}\underline{j}\overline{b}}\left\{\left(\underline{i}\nu'|\underline{j}\overline{b}\right)^{(\mathbf{B})}\left[2\left(\underline{i}\nu|\underline{j}\overline{b}\right) - \left(\underline{i}\overline{b}|\underline{j}\nu\right)\right]\right. \\ &+ \left.\left(\underline{i}\nu'|\underline{j}\overline{b}\right)\left[2\left(\underline{i}\nu|\underline{j}\overline{b}\right)^{(\mathbf{B})} - \left(\underline{i}\overline{b}|\underline{j}\nu\right)^{(\mathbf{B})}\right]\right\}, \end{aligned} \tag{58}$$

$$\begin{aligned} \overline{R}^{\mathbf{B}}_{\mu\mu'} =& \sum_{\nu\lambda\sigma}\left[\left(\mu'\overline{\nu}^{\mathbf{B}}|\underline{\lambda}\overline{\sigma}\right) + \left(\mu'\overline{\nu}|\underline{\lambda}^{\mathbf{B}}\overline{\sigma}\right) + \left(\mu'\overline{\nu}|\underline{\lambda}\overline{\sigma}^{\mathbf{B}}\right)\right] \\ &\times [2(\mu\nu|\lambda\sigma) - (\mu\sigma|\lambda\nu)] \\ &+ \sum_{\overline{a}\underline{j}\overline{b}}\left\{\left(\mu\overline{a}|\underline{j}\overline{b}\right)^{(\mathbf{B})}\left[2\left(\mu\overline{a}|\underline{j}\overline{b}\right) - \left(\mu\overline{b}|\underline{j}\overline{a}\right)\right]\right. \\ &+ \left.\left(\mu\overline{a}|\underline{j}\overline{b}\right)\left[2\left(\mu\overline{a}|\underline{j}\overline{b}\right)^{(\mathbf{B})} - \left(\mu\overline{b}|\underline{j}\overline{a}\right)^{(\mathbf{B})}\right]\right\}. \end{aligned} \tag{59}$$

We furthermore use the RI approximation in order to avoid four-center integrals. As a metric for the RI, we employ an attenuated Coulomb metric[46–48] as in our recent work on MP2 energies.[45] With this metric, four-center ERIs can be approximated as follows:

$$(\mu\nu|\lambda\sigma) \approx \sum_{PS} (\mu\nu{\mid}P)\, \tilde{C}_{PS}\, (S{\mid}\lambda\sigma), \tag{60}$$

where

$$\tilde{C}_{PS} = \sum_{QR} (P{\mid}Q)^{-1} (Q|R) (R{\mid}S)^{-1}, \tag{61}$$

$$(\mu\nu{\mid}P) = \iint \chi^*_\mu(\mathbf{r}_1)\chi_\nu(\mathbf{r}_1)\frac{\mathrm{erfc}(\omega r_{12})}{r_{12}}\chi_P(\mathbf{r}_2)d\mathbf{r}_1 d\mathbf{r}_2, \tag{62}$$

and

$$(P{\mid}Q) = \iint \chi_P(\mathbf{r}_1)\frac{\mathrm{erfc}(\omega r_{12})}{r_{12}}\chi_Q(\mathbf{r}_2)d\mathbf{r}_1 d\mathbf{r}_2. \tag{63}$$

The attenuated Coulomb metric depends on the parameter ω, which determines accuracy and sparsity in the three-center integrals. For ω equal to zero, the attenuated Coulomb metric reduces to the standard Coulomb metric, which is highly accurate but has long-ranged coupling between the bra and ket of the three-center integrals. On the other hand, in the limit $\omega \to \infty$, the rather inaccurate overlap metric is obtained, which leads to a high degree of sparsity in the three-center integrals. It has been shown for MP2 and direct RPA that $\omega = 0.1$ provides a good compromise and gives significant sparsity in the integrals in combination with only small errors compared to the Coulomb metric.[45,67] Therefore, we also use $\omega = 0.1$ for all timings shown in this work.

Using the RI approximation allows us to eliminate all transformed four-center integrals from Eqs. (56)–(59) by substituting them with three- and two-center RI-integrals. In consequence, several different types of transformed three-center integrals are needed. This includes integrals that are transformed with Cholesky-MO coefficients or with perturbed pseudo-densities and transformed three-center GIAO integrals, such as $(\underline{i}\overline{a}|P)^{(\mathbf{B})}$. In order to compute all of these transformed integrals, the three-center integrals in the atomic orbital basis are computed once and written to disk and then read into memory for each Laplace point. The AO three-center integrals for any auxiliary basis function are copied to a block-sparse matrix; likewise, the Cholesky-MOs and pseudo-density matrices are kept in a block-sparse matrix format. Then, the necessary integral transformations can be carried out efficiently using block-sparse matrix multiplications.

Next, the matrices $\underline{\mathbf{R}}$ and $\overline{\mathbf{R}}$, as well as their B-field derivatives, are computed from the three-center integrals. Figure 2 displays the algorithm used for efficiently computing the Coulomb-type contributions to the matrices $\underline{\mathbf{R}}$ and $\overline{\mathbf{R}}$. For this, the transformed integrals are copied into a natural blocking[47,55] data format, which has also been described in Ref. 45 and is used for speeding up the subsequent contractions. As an example, let us discuss the following contraction, which is one of the formally $\mathcal{O}(N^4)$-scaling steps in the algorithm from Fig. 2,

$$A_{PQ} = \sum_{\underline{i}\overline{a}} (\underline{i}\overline{a}{\mid}P)\,(\underline{i}\overline{a}{\mid}Q). \tag{64}$$

```
$\overline{\mathbf{R}} = \underline{\mathbf{R}} = \mathbf{A} = 0$

for all $\underline{i}$ do
    for all $P \in \{P\}_{\underline{i}}^{\overline{a}}$ and $Q \in \{Q\}_{\underline{i}}^{\overline{a}}$ do
        $A_{PQ} \mathrel{+}= \sum_{\overline{a} \in \{\overline{a}\}_{\underline{i}}^{P}} (\underline{i}\overline{a}|P)\,(\underline{i}\overline{a}|Q)$
    end for
end for

$\mathbf{B} = \tilde{\mathbf{C}}\mathbf{A}\tilde{\mathbf{C}}$

for all $\underline{i}$ do
    $\mathbf{D} = 0$
    for all $\nu' \in \{\nu'\}_{\underline{i}}^{P}$ and $Q$ do
        $D_{\nu' Q} = \sum_{P \in \{P\}_{\underline{i}}^{\nu'}} (\underline{i}\nu'|P)\,B_{PQ}$
    end for
    for all $\nu \in \{\nu\}_{\underline{i}}^{Q}$ and $\nu' \in \{\nu'\}_{\underline{i}}^{P}$ do
        $\underline{R}_{\nu\nu'} \mathrel{+}= 2 \sum_{Q \in \{Q\}_{\underline{i}}^{\nu}} (\underline{i}\nu|Q)\,D_{\nu' Q}$
    end for
end for

for all $\overline{a}$ do
    $\mathbf{E} = 0$
    for all $\mu' \in \{\mu'\}_{\overline{a}}^{P}$ and $Q$ do
        $E_{\mu' Q} = \sum_{P \in \{P\}_{\overline{a}}^{\mu'}} (\mu'\overline{a}|P)\,B_{PQ}$
    end for
    for all $\mu \in \{\mu\}_{\overline{a}}^{Q}$ and $\mu' \in \{\mu'\}_{\overline{a}}^{P}$ do
        $\overline{R}_{\mu\mu'} \mathrel{+}= 2 \sum_{Q \in \{Q\}_{\overline{a}}^{\mu}} (\mu\overline{a}|Q)\,E_{\mu' Q}$
    end for
end for
```

FIG. 2. Algorithm for the computation of the Coulomb contributions to $\underline{\mathbf{R}}$ and $\overline{\mathbf{R}}$.

The integrals used in this contraction are arranged such that there is one matrix for each $\underline{i}$ with initial dimensions of $n_{\text{virt}} \times n_{\text{aux}}$ (where n_{virt} is the number of virtual Cholesky MOs and n_{aux} is the number of auxiliary basis functions). Before carrying out any contraction, rows and columns in these matrices get deleted if they only contain elements that are below the chosen natural blocking threshold. By this, the dimensions of these matrices are reduced and the integral contractions are accelerated.

For a given $\underline{i}$, the indices of the significant rows and columns are collected in sparse lists $\{\overline{a}\}_{\underline{i}}^{P}$ and $\{P\}_{\underline{i}}^{\overline{a}}$. The list $\{\overline{a}\}_{\underline{i}}^{P}$, e.g., contains all virtual pseudo-MOs $\phi_{\overline{a}}$ that are significant for the orbital $\phi_{\underline{i}}$. The superscript index P shall denote the third index of the three-center integrals, from which the sparse list is derived (in this case, $(\underline{i}\overline{a}|P)$). An additional superscript index $\mathbf{B}$ is used to denote lists derived from GIAO integrals. The list $\{\overline{a}\}_{\underline{i}}^{P,\mathbf{B}}$, e.g., is derived from the $(\underline{i}\overline{a}|P)^{(\mathbf{B})}$ integrals. Note also that two sparse lists are identical if the contained indices refer to the same type of orbital, e.g., $\{\overline{a}\}_{\underline{i}}^{P} = \{\overline{b}\}_{\underline{j}}^{Q}$. Such sparse lists, which have also been employed in Ref. 45, are used extensively for exploiting sparsity in the algorithms from Figs. 2–5.

For the contraction in Eq. (64), the scaling can be reduced from formally $\mathcal{O}(N^4)$ to asymptotically linear because for large systems, only a constant number of virtual pseudo-MOs and auxiliary functions are expected to be significant for a given $\phi_{\underline{i}}$.

The matrix $\mathbf{A}$ from Eq. (64) is then multiplied from both sides with $\tilde{\mathbf{C}}$; this step scales cubically but has a very small prefactor. The resulting matrix $\mathbf{B}$ in Fig. 2 and the intermediates $\mathbf{D}$ and $\mathbf{E}$ are all dense matrices. As these matrices are dense, the scaling of the remaining steps in the algorithm shown in Fig. 2 cannot be reduced to asymptotically linear but only to asymptotically quadratic; in these steps, natural blocking is employed again in order to exploit the sparsity in the three-center integrals.

For MP2 shielding calculations, not only the computational cost but also the memory and disk space requirements can be prohibitive. Compared to the implementation of Maurer and Ochsenfeld,[29] the memory requirements are substantially reduced by the use of the RI approximation, which allows us to avoid four-center integrals. The employed attenuated Coulomb metric increases the degree of sparsity in the three-center integrals. As in our implementation, the three-center integrals are always kept either in the natural blocking format or a block-sparse data format, the sparsity can be efficiently exploited, and the memory requirements be reduced significantly. In order to further increase the size of systems that can be treated with a given amount of memory, the three-center integrals are stored on disk and only read into memory when needed during the integral transformations and contractions, which are all done in batches. The size of the batches is chosen in accordance with the available memory. Therefore, the available disk space determines the maximum system size that can be treated. In Sec. IV B, we show data for the required disk space in illustrative calculations on glycine chains and DNA strands.

Figures 3 and 4 show the algorithm used for computing the Coulomb-type contributions to the $\underline{\mathbf{R}}^{\mathbf{B}}$ and $\overline{\mathbf{R}}^{\mathbf{B}}$ matrices. It is more involved than the corresponding algorithm for the contributions to the unperturbed $\underline{\mathbf{R}}$ and $\overline{\mathbf{R}}$ matrices from Fig. 2 but uses similar intermediates and computation steps. Formally, the algorithm leads to an $\mathcal{O}(N^4)$ scaling. Again, natural blocking is used for exploiting the sparsity in the three-center integrals. The asymptotic scaling of all steps except for the matrix multiplication $\tilde{\mathbf{C}}\mathbf{A}^{\mathbf{B}}\tilde{\mathbf{C}}$ can be reduced to either linear or quadratic in this way.

Finally, the algorithm for computing the exchange contributions to $\underline{\mathbf{R}}$ and $\overline{\mathbf{R}}$ and their magnetic field derivatives is shown in

```
$\mathbf{A}^{\mathbf{B}} = 0$

for all $\underline{i}$ do
    for all $P \in \{P\}_{\underline{i}}^{\overline{a}}$ and $Q \in \{Q\}_{\underline{i}}^{\overline{a},\mathbf{B}}$ do
        $A_{PQ}^{\mathbf{B}} \mathrel{+}= \sum_{\overline{a} \in [\{\overline{a}\}_{\underline{i}}^{P} \cap \{\overline{a}\}_{\underline{i}}^{Q,\mathbf{B}}]} (\underline{i}\overline{a}|P)\,(\underline{i}\overline{a}|Q)^{(\mathbf{B})}$ + transpose
    end for
    for all $P \in \{P\}_{\underline{i}}^{\nu}$ and $Q \in \{Q\}_{\underline{i}}^{\overline{\nu}^{\mathbf{B}}}$ do
        $A_{PQ}^{\mathbf{B}} \mathrel{+}= \sum_{\nu \in [\{\nu\}_{\underline{i}}^{P} \cap \{\overline{\nu}^{\mathbf{B}}\}_{\underline{i}}^{Q}]} (\underline{i}\nu|P)\,(\underline{i}\overline{\nu}^{\mathbf{B}}|Q)$
    end for
end for

for all $\overline{a}$ do
    for all $P \in \{P\}_{\overline{a}}^{\mu}$ and $Q \in \{Q\}_{\overline{a}}^{\underline{\mu}^{\mathbf{B}}}$ do
        $A_{PQ}^{\mathbf{B}} \mathrel{+}= \sum_{\mu \in [\{\mu\}_{\overline{a}}^{P} \cap \{\underline{\mu}^{\mathbf{B}}\}_{\overline{a}}^{Q}]} (\mu\overline{a}|P)\,(\underline{\mu}^{\mathbf{B}}\overline{a}|Q)$
    end for
end for

$\mathbf{B}^{\mathbf{B}} = \tilde{\mathbf{C}}\mathbf{A}^{\mathbf{B}}\tilde{\mathbf{C}}$
```

FIG. 3. Algorithm for the computation of $\mathbf{B}^{\mathbf{B}}$. For the definition of $\tilde{\mathbf{C}}$, see Fig. 2. In several steps, intersections of two sparse lists are employed.

```
$\overline{\mathbf{R}}^{\mathbf{B}} = \underline{\mathbf{R}}^{\mathbf{B}} = 0$

for all $\mu$ do
    for all $\nu \in \{\nu\}_{\mu}^{P}$ and $Q$ do
        $H_{\nu Q} = \sum_{P \in \{P\}_{\mu}^{\nu}} (\mu\nu \vdots P)\, B_{PQ}$
    end for
    for all $\nu' \in \{\nu'\}_{\mu}^{P}$ and $\nu \in \{\nu\}_{\underline{\mu}^{\mathbf{B}}}^{Q}$ do
        $\underline{R}_{\nu\nu'}^{\mathbf{B}} \mathrel{+}= 2 \sum_{Q \in \{Q\}_{\underline{\mu}^{\mathbf{B}}}^{\nu}} (\underline{\mu}^{\mathbf{B}}\nu \vdots Q)\, H_{\nu' Q}$
    end for
end for

for all $\nu$ do
    for all $\mu' \in \{\mu'\}_{\nu}^{P}$ and $\mu \in \{\mu\}_{\overline{\nu}^{\mathbf{B}}}^{Q}$ do
        $\overline{R}_{\mu\mu'}^{\mathbf{B}} \mathrel{+}= 2 \sum_{Q \in \{Q\}_{\mu}^{\underline{\nu}^{\mathbf{B}}}} (\mu\underline{\nu}^{\mathbf{B}} \vdots Q)\, H_{\mu' Q}$
    end for
end for

for all $\underline{i}$ do
    for all $\nu \in \{\nu\}_{\underline{i}}^{P}$ and $Q$ do
        $I_{\nu Q} = \sum_{P \in \{P\}_{\underline{i}}^{\nu}} (\underline{i}\nu | P)\, B_{PQ}$
    end for
    for all $\nu \in \{\nu\}_{\underline{i}}^{P}$ and $\nu' \in \{\nu'\}_{\underline{i}}^{Q,\mathbf{B}}$ do
        $\underline{R}_{\nu\nu'}^{\mathbf{B}} \mathrel{+}= 2 \sum_{Q \in \{Q\}_{\underline{i}}^{\nu',\mathbf{B}}} (\underline{i}\nu' \vdots Q)^{(\mathbf{B})}\, I_{\nu Q} + \text{transpose}$
    end for
end for

for all $\overline{a}$ do
    for all $\mu \in \{\mu\}_{\overline{a}}^{P}$ and $Q$ do
        $J_{\mu Q} = \sum_{P \in \{P\}_{\overline{a}}^{\mu}} (\mu\overline{a} | P)\, B_{PQ}$
    end for
    for all $\mu \in \{\mu\}_{\overline{a}}^{P}$ and $\mu' \in \{\mu'\}_{\overline{a}}^{Q,\mathbf{B}}$ do
        $\overline{R}_{\mu\mu'}^{\mathbf{B}} \mathrel{+}= 2 \sum_{Q \in \{Q\}_{\overline{a}}^{\mu',\mathbf{B}}} (\mu'\overline{a} \vdots Q)^{(\mathbf{B})}\, J_{\mu Q} + \text{transpose}$
    end for
end for

for all $\underline{i}$ do
    for all $\nu \in \{\nu\}_{\underline{i}}^{P}$ and $Q$ do
        $L_{\nu Q} = \sum_{P \in \{P\}_{\underline{i}}^{\nu}} (\underline{i}\nu \vdots P)\, B_{PQ}^{\mathbf{B}}$
    end for
    for all $\nu' \in \{\nu'\}_{\underline{i}}^{Q}$ and $\nu \in \{\nu\}_{\underline{i}}^{P}$ do
        $\underline{R}_{\nu\nu'}^{\mathbf{B}} \mathrel{+}= 2 \sum_{Q \in \{Q\}_{\underline{i}}^{\nu}} (\underline{i}\nu' \vdots Q)\, L_{\nu Q}$
    end for
end for

for all $\overline{a}$ do
    for all $\mu \in \{\mu\}_{\overline{a}}^{P}$ and $Q$ do
        $M_{\mu Q} = \sum_{P \in \{P\}_{\overline{a}}^{\mu}} (\mu\overline{a} \vdots P)\, B_{PQ}^{\mathbf{B}}$
    end for
    for all $\mu \in \{\mu\}_{\overline{a}}^{P}$ and $\mu' \in \{\mu'\}_{\overline{a}}^{Q}$ do
        $\overline{R}_{\mu\mu'}^{\mathbf{B}} \mathrel{+}= 2 \sum_{Q \in \{Q\}_{\overline{a}}^{\mu}} (\mu'\overline{a} \vdots Q)\, M_{\mu Q}$
    end for
end for
```

FIG. 4. Algorithm for the computation of the Coulomb contributions to $\underline{\mathbf{R}}^{\mathbf{B}}$ and $\overline{\mathbf{R}}^{\mathbf{B}}$. For the definitions of **B** and $\tilde{\mathbf{C}}$, see Fig. 2. $\mathbf{B}^{\mathbf{B}}$ is computed as shown in Fig. 3.

Fig. 5. In this algorithm, four-center integrals are built and then contracted as opposed to the algorithms used for the Coulomb contributions. As a first step, three-center integrals are multiplied with the matrix $\tilde{\mathbf{C}}$ from Eq. (60) as follows:

```
$\mathbf{R} = 0$
for all $w$ do
    for all $x \in \{x\}_{w}^{P}$ and $Q$ do
        $\tilde{B}_{wx}^{Q} = \sum_{P \in \{P\}_{w}^{x}} (wx \vdots P)\, \tilde{C}_{PQ}$
    end for
    for all $z \in \{z\}_{w}^{P}$ and $Q$ do
        $\bar{B}_{wz}^{Q} = \sum_{P \in \{P\}_{w}^{x}} (wz \vdots P)\, \tilde{C}_{PQ}$
    end for
    for all $y$ do
        for all $x \in \{x\}_{w}^{Q}$ and $z \in \{z\}_{y}^{Q}$ do
            $(wx|yz) = \sum_{Q \in \{Q\}_{y}^{z}} \tilde{B}_{wx}^{Q}\, (Q \vdots yz)$
        end for
        for all $x' \in \{x'\}_{y}^{Q}$ and $z \in \{z\}_{w}^{Q}$ do
            $(wz|yx') = \sum_{Q \in \{Q\}_{y}^{x'}} \bar{B}_{wz}^{Q}\, (Q \vdots yx')$
        end for
        for all $x \in \{x\}_{w}^{P}$ and $x' \in \{x'\}_{y}^{P}$ do
            $R_{xx'} \mathrel{+}= \sum_{z \in [\{z\}_{w}^{P} \cap \{z\}_{y}^{P}]} (wx|yz)\, (wz|yx')$
        end for
    end for
end for
```

FIG. 5. Algorithm for the computation of an exchange-type contribution to $\underline{\mathbf{R}}$, $\overline{\mathbf{R}}$, $\underline{\mathbf{R}}^{\mathbf{B}}$, or $\overline{\mathbf{R}}^{\mathbf{B}}$, which is of the general form $\sum_{w,y,z} (wx|yz)(wz|yx')$, where w, x, y, and z can stand either for Cholesky-MO-indices or for transformed or untransformed AO indices. **R** can stand for $\underline{\mathbf{R}}$, $\overline{\mathbf{R}}$, $\underline{\mathbf{R}}^{\mathbf{B}}$, or $\overline{\mathbf{R}}^{\mathbf{B}}$.

$$\tilde{B}_{wx}^{Q} = \sum_{P \in \{P\}_{w}^{x}} (wx \vdots P)\, \tilde{C}_{PQ}, \tag{65}$$

where w and x can denote either Cholesky-MO-indices or transformed or untransformed AO indices. This step scales asymptotically quadratic even if the short-range coupling of the indices in the three-center integrals is exploited because $\tilde{\mathbf{C}}$ is a dense matrix. Next, the four-center integrals are built by contracting the intermediate $\tilde{\mathbf{B}}$ with three-center integrals as follows:

$$(wx|yz) = \sum_{Q} \tilde{B}_{wx}^{Q}\, (Q \vdots yz)\,. \tag{66}$$

For any wy-pair, the computational effort is $\mathcal{O}(1)$ because only $\mathcal{O}(1)$ orbitals with indices x and z are significant due to exponential coupling between w and x and between y and z. In addition, the summation over Q can be restricted to $\mathcal{O}(1)$ elements by exploiting the short-range coupling between Q and y. Therefore, the asymptotic scaling of this step is quadratic. Likewise, also the remaining steps of the algorithm in Fig. 5 scale asymptotically quadratic.

III. COMPUTATIONAL DETAILS

The described method was implemented in a development version of the FermiONs++ program.[68–70] Unless stated differently, the def2-SVP basis set[71] was used in combination with the corresponding auxiliary basis set[72] in all calculations. Two different sets of parameters are used in the calculations shown in Sec. IV A, which we denote as "Loose" and "Tight." For the "Loose" settings, the attenuation parameter ω is set to 0.1, and 5 Laplace points are used in MP2 and 10 Laplace points in DL-CPSCF[73] iterations, which are used for computing $\mathbf{P}^{\mathrm{B}}$ and solving the Z-vector equations. Furthermore, 10^{-6} is used as the threshold for deleting rows and columns in the natural blocking format. These values have also been used in our recent work on efficient computation of MP2 energies with the ω-RI-CDD-MP2 method, where we demonstrated that chemically accurate energies can be obtained.[45] For the "Tight" settings, the natural blocking threshold and ω are set to zero (meaning that effectively a Coulomb metric is used); the number of Laplace points for both the MP2 shielding part and the DL-CPSCF is increased to 13 in both cases. All timings in Sec. IV B were carried out using the "Loose" settings.

In all calculations, shell-pairs with an overlap of less than 10^{-12} were neglected. Integral screening with a threshold of 10^{-10} is employed during the SCF. The Laplace points are determined using a minimax algorithm as described in Refs. 74 and 75; the number of integration points is reduced automatically in our implementation if the fitting interval is small and no improved accuracy can be obtained with more integration points. The pseudo-density matrices are orthogonalized prior to the pivoted Cholesky decomposition as described in Ref. 67; afterward, the orthogonalization is reverted. No frozen-core approximation is employed. For comparison, canonical MP2 shielding calculations were carried out with the Turbomole program.[25,76,77] All timings were performed using 64 threads on a node with AMD EPYC 7302 processors, 256 GB RAM, and a solid-state drive (SSD) with a capacity of 1.7 TB.

IV. RESULTS

A. Accuracy of the introduced approximations

In order to analyze the influence of the employed approximations on the accuracy, NMR shieldings were computed for all complexes/dimers from the S22 test set[78] and for all structures from the benchmark set used by Flaig *et al.*[22] In these calculations, both the "Loose" and "Tight" settings were employed. As a reference, NMR shielding calculations with canonical MP2 performed with the Turbomole program[25,76,77] are used. The statistics on the two benchmark sets are shown in Tables I and II. The "Loose" settings, which will be employed also for the timings in Sec. IV B, give satisfactory accuracy for the NMR shieldings, as the mean absolute deviations amount to only 0.024 ppm for the S22 test set and to 0.036 ppm for the benchmark set from the study by Flaig *et al.*[22] No clear improvement compared to the "Loose" settings is observed when using the "Tight" settings. This suggests that the dominant source of error compared to canonical MP2 is the RI approximation; the errors due to the other approximations appear to be negligible for these test sets. Indeed, for both test sets, the already small deviations to Turbomole are reduced further in calculations with "Tight" settings and a cc-pVQZ-RI auxiliary basis set[79,80] instead of a def2-SVP-RI basis set. We thus conclude that the RI approximation is the largest source of error, which does not preclude the computation of shieldings, in very good agreement with canonical MP2 shieldings.

TABLE I. Statistics for MP2 shielding calculations for all dimers from the S22 test set. "Loose" and "Tight" settings are defined in the text. "Tight*" denotes shieldings computed with "Tight" settings and a cc-pVQZ-RI auxiliary basis set instead of a def2-SVP-RI basis set. MP2 shieldings computed with Turbomole[25,76,77] are used as a reference. All values are given in ppm.

	Loose	Tight	Tight*
Mean error	0.016	0.024	0.001
Mean absolute error	0.024	0.026	0.011
Root mean squared error	0.044	0.047	0.021
Maximum error	0.229	0.271	0.151

TABLE II. Statistics for MP2 shielding calculations on the benchmark set from the study by Flaig *et al.*[22] "Loose" and "Tight" settings are defined in the text. "Tight*" denotes shieldings computed with "Tight" settings and a cc-pVQZ-RI auxiliary basis set instead of a def2-SVP-RI basis set. MP2 shieldings computed with Turbomole[25,76,77] are used as a reference. All values are given in ppm.

	Loose	Tight	Tight*
Mean error	0.028	0.033	−0.002
Mean absolute error	0.036	0.035	0.004
Root mean squared error	0.098	0.104	0.017
Maximum error	0.662	0.604	0.150

However, one needs to consider that the molecules in the S22 test set and the test set from the study by Flaig *et al.*[22] are relatively small and that the errors for several of the used approximations, such as the local RI-metric or the sparse linear algebra, might be larger for more extended systems. Therefore, we also carried out calculations on all monomers in the L7 benchmark set[81] and display the results in Table III. For the purpose of computational efficiency, we use the SOS-approximation in these calculations; this means that the exchange contributions are neglected, while the Coulomb contributions are scaled with 1.3 as suggested by Jung *et al.*[40] (note that alternatively, for application studies, the basis-set specific scaling parameters from the study by Maurer and Ochsenfeld[39] might also be useful). In addition, for the L7 test set, the differences between the "Loose" and "Tight" settings are small with a mean absolute deviation of 0.028 ppm. This suggests that the "Loose" settings are appropriate for computing reliable NMR shieldings also for larger systems; therefore, we use these settings also for the timings shown in Sec. IV B.

B. Scaling behavior and efficiency

In Table IV, timings with the new method in shielding calculations on glycine chains are shown. The scaling decreases for larger chain lengths and amounts to 2.74 between the two largest

TABLE III. Statistics for SOS-MP2 shielding calculations for all monomers from the L7 test set. SOS-MP2 shieldings computed with FermiONs++[68–70] and "Tight" settings are used as a reference. All values are given in ppm.

	Loose
Mean error	−0.013
Mean absolute error	0.028
Root mean square error	0.061
Maximum error	0.462

TABLE IV. Timings for MP2 shielding calculations (including exchange) on glycine chains. "nglyc" denotes the number of glycines, and "nbas" is the number of basis functions.

nglyc	nbas	Wall time (s)
1	95	415
2	166	2 051
4	308	14 063
6	450	34 861
10	734	133 282

systems, glycine-6 and glycine-10. With this, the observed scaling is well below the formal $\mathcal{O}(N^5)$ scaling of the MO-based method. This agrees well with the expectations, as the asymptotic scaling of all $\mathcal{O}(N^4)$ and $\mathcal{O}(N^5)$ scaling steps is reduced to linear or quadratic if sparsity is exploited (see Sec. II).

Figure 6 also displays timings on glycine chains; in contrast to the calculations from Table IV, the SOS-approximation is used in all calculations. The wall times needed for the SOS-MP2 shielding calculations are much lower than those for the full MP2 shieldings (e.g., by a factor of 32.9 for glycine-10). This shows that the exchange contributions are the major bottleneck in the MP2 shielding calculations from Fig. 4. Some acceleration for the exchange contributions might be possible by using integral screening, e.g., with Schwarz estimates,[82] in conjunction with the natural blocking, as done in our recent work on MP2 energies,[45] but we have not explored this so far. For SOS-MP2, the scaling is close to quadratic for the larger chain lengths and thus lower than the formal scaling of SOS-MP2 ($\mathcal{O}(N^4)$). Due to the significantly higher computational efficiency obtained with the SOS-approximation, it is likely preferable to use SOS-MP2 for studying very large systems. For this, we recommend to use the basis-set specific scaling factors from the study by Maurer and Ochsenfeld,[39] which were shown to provide similar or even better accuracy than canonical MP2.

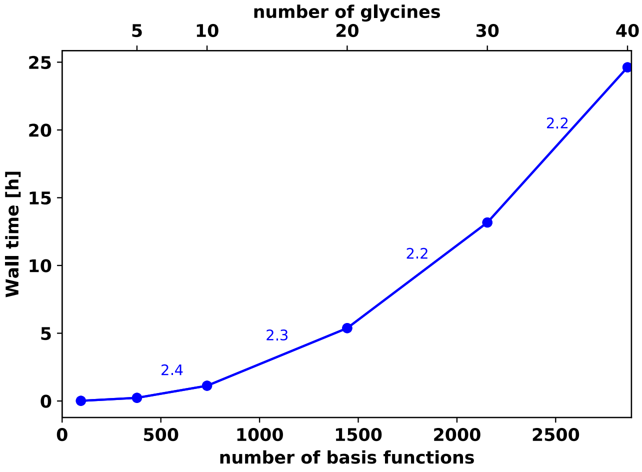


FIG. 6. Timings for SOS-MP2 shielding calculations on glycine chains. The colored numbers indicate the effective scaling between two data points.

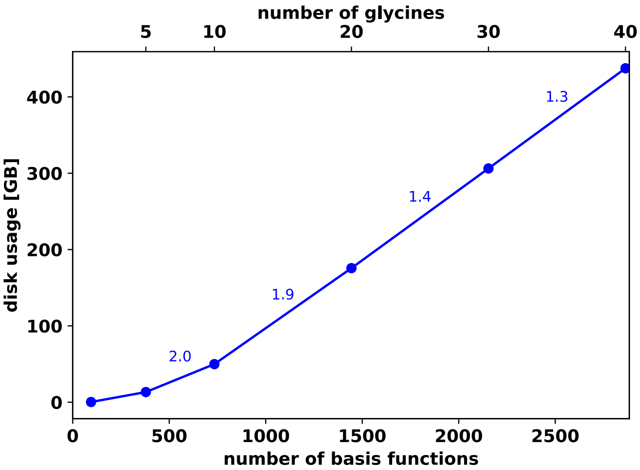


FIG. 7. Maximum disk usage for three-center integrals in SOS-MP2 shielding calculations on glycine chains. The colored numbers indicate the effective scaling between two data points.

As discussed, memory and/or disk space can also be a bottleneck in MP2 shielding calculations. Therefore, we show data for the used disk space during calculations on glycine chains in Fig. 7. For the largest system (glycine-40), roughly 437 GB of disk space were used. The scaling continuously decreases with the increasing chain length and amounts to 1.3 between glycine-30 and glycine-40. This is significantly lower than the formal cubic scaling for the memory demands of the three-center integrals and is enabled by the sparsity provided by the attenuated Coulomb metric and by the employed sparse data formats (block sparse matrices and natural blocking), which allow us to neglect many of the insignificant integrals.

In addition to the glycine chains, also timings on short DNA strands with one and two adenine–thymine base pairs were carried out (see Table V). These molecules are more representative of potential applications on large biomolecular systems. The calculation on the largest DNA strand, AT_2, with 128 atoms and 1332 basis functions took 11 h, which shows that the new method allows to treat systems in this size range with moderate computational effort. The scaling of the compute time between AT_1 and AT_2 amounts to 3.4. This is below the formal $\mathcal{O}(N^4)$ scaling of Laplace-transformed SOS-MP2, which indicates that there is already usable sparsity for systems in this size range. As expected, due to the less extended shape of the AT base pairs, the observed scaling is larger than for glycine chains with a similar number of basis

TABLE V. Wall times, scaling, and needed disk space in calculations on DNA strands with up to two adenine–thymine base pairs. "nbas" is the number of basis functions.

	AT_1	AT_2
nbas	625	1332
Wall time (h)	1.0	14.0
Scaling	⋯	3.44
Disk space (GB)	59.5	531.4

TABLE VI. Timings of individual steps during a SOS-MP2 shielding calculation on AT_2.

	Wall time (s)
Integral transformation	15 702
$\underline{\mathbf{R}}$ and $\overline{\mathbf{R}}$	1 205
$\underline{\mathbf{R}}^{\mathbf{B}}$ and $\overline{\mathbf{R}}^{\mathbf{B}}$	10 782
$\underline{\mathbf{Y}}^{\mathbf{B}}$ and $\overline{\mathbf{Y}}^{\mathbf{B}}$	2 795
Nested **Z**-vector	4 470
Total	39 619

functions. In Table VI, timings for several of the calculation steps are shown. Most of the time during the SOS-MP2 shielding calculations is used for the transformations of the three-center integrals followed by the computation of the perturbed $\underline{\mathbf{R}}^{\mathbf{B}}$ and $\overline{\mathbf{R}}^{\mathbf{B}}$ matrices. Significantly less compute time is needed for the unperturbed $\underline{\mathbf{R}}$ and $\overline{\mathbf{R}}$ matrices, the recursion formulas, and the nested Z-vector approach. In Table V, also the maximum amount of disk space used during the shielding calculations on the DNA strands is shown. Significantly more disk space than that for glycine chains with a similar number of basis functions is needed, which is caused by the lower sparsity. For AT_2, 531.4 GB disk space is needed for storing three-center integrals. Calculations on larger DNA strands were not possible with the available disk space on the employed nodes. In order to improve the applicability to larger systems, it is therefore important to circumvent the disk space bottleneck. A potential solution would be an integral-direct approach, by which we plan to address the disk space bottleneck in a future publication.

V. CONCLUSION

We presented an efficient method for computing NMR shieldings at the MP2 level of theory, which is based on Laplace-transformed AO-MP2. It employs many of the approximations that were also used in our recently published ω-RI-CDD-MP2 method[45] for correlation energies, including RI with an attenuated Coulomb metric and Cholesky decomposition of pseudo-density matrices. Furthermore, block sparse linear algebra and natural blocking are used in order to speed up the computations and reduce the memory needed for the three-center integrals. A nested Z-vector allows us to efficiently compute the entire set of NMR shieldings for all nuclei (as opposed to the selected-nuclei formulation of Maurer and Ochsenfeld[29]) and requires the solution of only seven CPSCF equations irrespective of the molecule size. Benchmark calculations indicated that the same thresholds and settings that were shown previously to give accurate correlation energies with the ω-RI-CDD-MP2 method also allow us to compute accurate NMR shieldings. Timings on glycine chains show close to quadratic scaling for larger systems. Particularly high efficiency is obtained in combination with an SOS-approximation. In addition, the scaling of the needed disk space decreases with molecule size and becomes sub-quadratic for larger chain lengths. Calculations on DNA strands with up to two base pairs illustrate the potential of the method for computing highly accurate NMR shieldings of biomolecular systems. The new method thus enables the computation of NMR shieldings of large molecules with significantly reduced computational cost and memory requirements.

SUPPLEMENTARY MATERIAL

See the supplementary material for all computed NMR shieldings from Sec. IV A and the employed xyz-structures for DNA and glycine chains.

ACKNOWLEDGMENTS

The authors acknowledge the financial support of the "Deutsche Forschungsgemeinschaft" (DFG) under Grant No. SFB 1309-32587107 and the cluster of excellence (Grant No. EXC2111-390814868) "Munich Center for Quantum Science and Technology" (MCQST). M.G. and S.V. acknowledge the "Studienstiftung des Deutschen Volkes" for graduate fellowships. C.O. acknowledges additional financial support as a Max-Planck-Fellow at MPI-FKF Stuttgart.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

DATA AVAILABILITY

The data that support the findings of this study are available within the article and its supplementary material.

APPENDIX A: RECURSION FORMULAS FOR Y

In this section, the recursion formulas for a matrix **Y** are derived such that $\mathrm{Tr}[\mathbf{Y}\mathbf{A}^{\xi}] = \mathrm{Tr}\left[\mathbf{B}\left(e^{\mathbf{A}}\right)^{\xi}\right]$ holds. First, we insert the Taylor series expansion of the matrix exponential

$$\mathrm{Tr}\left[\mathbf{B}\left(e^{\mathbf{A}}\right)^{\xi}\right] = \mathrm{Tr}\left[\mathbf{B}\left(1 + \mathbf{A} + \frac{1}{2!}\mathbf{A}\mathbf{A} + \frac{1}{3!}\mathbf{A}\mathbf{A}\mathbf{A} + \cdots\right)^{\xi}\right] \tag{A1}$$

and subsequently differentiate each term using the product rule

$$\mathrm{Tr}\left[\mathbf{B}\left(e^{\mathbf{A}}\right)^{\xi}\right] = \mathrm{Tr}\Bigg[\mathbf{B}\bigg(0 + \mathbf{A}^{\xi} + \frac{1}{2!}\left(\mathbf{A}^{\xi}\mathbf{A} + \mathbf{A}\mathbf{A}^{\xi}\right) + \frac{1}{3!}\left(\mathbf{A}^{\xi}\mathbf{A}\mathbf{A} + \mathbf{A}\mathbf{A}^{\xi}\mathbf{A} + \mathbf{A}\mathbf{A}\mathbf{A}^{\xi}\right) + \cdots\bigg)\Bigg]. \tag{A2}$$

In each term, cyclic permutations under the trace are used to move $\mathbf{A}^{\xi}$ to the far right,

$$\mathrm{Tr}\left[\mathbf{B}\left(e^{\mathbf{A}}\right)^{\xi}\right] = \mathrm{Tr}\Bigg[\bigg(\mathbf{B} + \frac{1}{2!}(\mathbf{A}\mathbf{B} + \mathbf{B}\mathbf{A}) + \frac{1}{3!}(\mathbf{A}\mathbf{A}\mathbf{B} + \mathbf{A}\mathbf{B}\mathbf{A} + \mathbf{B}\mathbf{A}\mathbf{A}) + \cdots\bigg)\mathbf{A}^{\xi}\Bigg]. \tag{A3}$$

Upon comparing with Eq. (17), the term in round brackets can then be identified with **Y**,

$$\mathbf{Y} = \mathbf{B} + \frac{1}{2!}(\mathbf{AB} + \mathbf{BA}) + \frac{1}{3!}(\mathbf{AAB} + \mathbf{ABA} + \mathbf{BAA}) + \cdots$$
$$= \sum_{k=1}^{\infty}\sum_{i=0}^{k-1}\frac{1}{k!}\mathbf{A}^{i}\mathbf{B}\mathbf{A}^{k-1-i}. \tag{A4}$$

$\mathbf{Y}$ can be efficiently computed using recursion as shown in the following equations:

$$\mathbf{Y} = \sum_{k=0}^{k_{\max}}\mathbf{Y}_k, \tag{A5}$$

$$\mathbf{Y}_k = \frac{1}{k}[\mathbf{A}\mathbf{Y}_{k-1} + \mathbf{B}\mathbf{A}_{k-1}] \quad \forall k > 0, \tag{A6}$$

$$\mathbf{Y}_0 = \mathbf{0}, \tag{A7}$$

$$\mathbf{A}_k \equiv \frac{1}{k!}\mathbf{A}^k. \tag{A8}$$

In order to ensure a fast convergence and avoid numerical difficulties, we employ a modified version of the "scaling and squaring" approach,[83,84] which is commonly used for computing matrix exponentials. Instead of using $\mathbf{A}$ and $\mathbf{B}$ in Eqs. (A5)–(A8), two matrices $\mathbf{A}'$ and $\mathbf{B}'$ are employed as follows:

$$\mathbf{A}' = \frac{1}{n}\mathbf{A}, \tag{A9}$$

where n is an integer number chosen such that $|\mathbf{A}'| < 0.5$, and

$$\mathbf{B}' = \sum_{i=0}^{n-1}\exp(\mathbf{A}')^{i}\,\mathbf{B}\exp(\mathbf{A}')^{n-1-i}. \tag{A10}$$

From the result $\mathbf{Y}'$, one can then compute the matrix $\mathbf{Y}$ as follows:

$$\mathbf{Y} = \frac{1}{n}\mathbf{Y}'. \tag{A11}$$

APPENDIX B: DIFFERENTIATED RECURSION FORMULAS FOR OBTAINING $\mathbf{Y}^{\xi}$

For computing MP2 shieldings, also derivatives of the $\mathbf{Y}$ matrices from Appendix A are needed. These perturbed $\mathbf{Y}$ matrices can also be computed using recursions, which are obtained by differentiating the recursion from Eqs. (A5)–(A8) as follows:

$$\mathbf{Y}^{\xi} = \sum_{k=0}^{k_{\max}}\mathbf{Y}_k^{\xi}, \tag{B1}$$

with

$$\mathbf{Y}_k^{\xi} = \frac{1}{k}\left[\mathbf{A}^{\xi}\mathbf{Y}_{k-1} + \mathbf{A}\mathbf{Y}_{k-1}^{\xi} + \mathbf{B}^{\xi}\mathbf{A}_{k-1} + \mathbf{B}(\mathbf{A}_{k-1})^{\xi}\right] \quad \forall k > 0, \tag{B2}$$

$$\mathbf{Y}_0^{\xi} = 0. \tag{B3}$$

$\mathbf{A}_{k-1}^{\xi}$ in Eq. (B2) can also be computed recursively as follows:

$$\mathbf{A}_k^{\xi} = \frac{1}{k}\left[\mathbf{A}\mathbf{A}_{k-1}^{\xi} + \mathbf{A}^{\xi}\mathbf{A}_{k-1}\right] \quad \forall k > 0, \tag{B4}$$

$$\mathbf{A}_0^{\xi} = 0. \tag{B5}$$

Faster convergence can be obtained using a modified scaling and squaring approach. For this, $\mathbf{A}' = \mathbf{A}/n$ and $\mathbf{A}'^{\xi} = \mathbf{A}^{\xi}/n$ should be used instead of $\mathbf{A}$ and $\mathbf{A}^{\xi}$ in the recursion from Eqs. (B2) and (B4). Furthermore, $\mathbf{B}$ needs to be replaced with $\mathbf{B}'$ from Eq. (A10). The result $\mathbf{Y}'^{\xi}$, which is computed by carrying out the recursion from Eq. (B2) with $\mathbf{A}'$, $\mathbf{A}'^{\xi}$, $\mathbf{B}'$, and $\mathbf{B}'^{\xi}$ instead of their respective unprimed counterparts, finally needs to be multiplied with $\frac{1}{n}$ in order to obtain $\mathbf{Y}^{\xi}$.

APPENDIX C: EFFICIENT COMPUTATION OF MATRIX EXPONENTIALS AND PERTURBED MATRIX EXPONENTIALS

The Taylor series expansion of a matrix exponential $e^{\mathbf{A}}$ is given by

$$e^{\mathbf{A}} = \mathbf{1} + \mathbf{A} + \frac{1}{2!}\mathbf{AA} + \frac{1}{3!}\mathbf{AAA} + \cdots = \sum_{k=0}^{\infty}\frac{1}{k!}\mathbf{A}^k. \tag{C1}$$

The matrix exponential can be evaluated efficiently using recursion as follows:

$$e^{\mathbf{A}} = \sum_{k=0}^{\infty} e_k, \tag{C2}$$

$$e_k = \frac{1}{k}\, e_{k-1}\ \mathbf{A} \quad \forall k > 0, \tag{C3}$$

$$e_0 = \mathbf{1}. \tag{C4}$$

The recursion from Eq. (C2) can only be expected to converge quickly if the norm of $\mathbf{A}$ is small. If this is not the case, "scaling and squaring"[83,84] can be applied. For this, an integer number n is chosen such that $|\mathbf{A}/n| < 0.5$. Then, the matrix exponential $e^{\mathbf{A}/n}$ is computed using the recursion from Eq. (C2). The exponential of $\mathbf{A}$ is obtained by taking $e^{\mathbf{A}/n}$ to the power n as follows:

$$e^{\mathbf{A}} = \left[e^{\mathbf{A}/n}\right]^{n}. \tag{C5}$$

For the presented MP2-NMR method, also perturbed matrix exponentials of the form $(e^{\mathbf{A}})^{\xi}$ need to be computed. A formula for the recursive computation of the perturbed matrix exponential can be derived by differentiating Eq. (C2) as follows:

$$(e^{\mathbf{A}})^{\xi} = \sum_{k=0}^{\infty} e_k{}^{\xi}, \tag{C6}$$

$$e_k{}^{\xi} = \frac{1}{k}\left(e_{k-1}{}^{\xi}\ \mathbf{A} + e_{k-1}\ \mathbf{A}^{\xi}\right) \ \forall k > 0, \tag{C7}$$

$$e_0{}^{\xi} = 0. \tag{C8}$$

If scaling and squaring is used, computing $\left(e^{\mathbf{A}}\right)^{\xi}$ requires the recursive computation of $\left(e^{\mathbf{A}/n}\right)^{\xi}$ and $e^{\mathbf{A}/n}$ as follows:

$$(e^{\mathbf{A}})^{\xi} = (e^{\mathbf{A}/n})^{\xi}(e^{\mathbf{A}/n})^{n-1} + (e^{\mathbf{A}/n})^{1}(e^{\mathbf{A}/n})^{\xi}(e^{\mathbf{A}/n})^{n-2} + \cdots$$
$$+ (e^{\mathbf{A}/n})^{n-1}(e^{\mathbf{A}/n})^{\xi} = \sum_{i=0}^{n-1}(e^{\mathbf{A}/n})^{i}(e^{\mathbf{A}/n})^{\xi}(e^{\mathbf{A}/n})^{n-1-i}. \tag{C9}$$

An efficient evaluation of this formula is possible via another recursion,

$$(e^{\mathbf{A}})^{\xi} = E_n, \tag{C10}$$

$$E_n = (e^{\mathbf{A}/n})E_{n-1} + (e^{\mathbf{A}/n})^{\xi}(e^{\mathbf{A}/n})^{n-1} \ \forall n > 1, \tag{C11}$$

$$E_1 = (e^{\mathbf{A}/n})^{\xi}. \tag{C12}$$

APPENDIX D: NESTED Z-VECTOR APPROACH WITH D-CPSCF

There are different formulations of CPSCF, which lead to different equations for $\mathbf{A}$ and $\mathbf{b}^{\mathbf{B}}$ in Eq. (40). Here, we use the D-CPSCF formulation from Ref. 64, which results in the following expressions for $\mathbf{A}$ and $\mathbf{b}^{\mathbf{B}}$:

$$\begin{aligned}\mathbf{A}\left[\mathbf{P}^{\mathbf{B}}\right] = {}& 3\,\mathbf{FP}^{\mathbf{B}}\mathbf{S} + 3\,\mathbf{SP}^{\mathbf{B}}\mathbf{F} - 2\,\mathbf{FP}^{\mathbf{B}}\mathbf{SPS} \\ & - 2\,\mathbf{SPSP}^{\mathbf{B}}\mathbf{F} - 4\,\mathbf{FPSP}^{\mathbf{B}}\mathbf{S} - 4\,\mathbf{SP}^{\mathbf{B}}\mathbf{SPF} \\ & + \mathbf{G}[\mathbf{X}]\mathbf{PS} + \mathbf{SPG}[\mathbf{X}] - 2\,\mathbf{SPG}[\mathbf{X}]\mathbf{PS},\end{aligned} \tag{D1}$$

where

$$\mathbf{X} = \mathbf{P}^{\mathbf{B}}\mathbf{SP} + \mathbf{PSP}^{\mathbf{B}} - 2\,\mathbf{PSP}^{\mathbf{B}}\mathbf{SP} \tag{D2}$$

and

$$\begin{aligned}\mathbf{b}^{\mathbf{B}} = {}& \mathbf{FPS}^{\mathbf{B}} + \mathbf{S}^{\mathbf{B}}\mathbf{PF} + 2\,\mathbf{FPS}^{\mathbf{B}}\mathbf{PS} + 2\,\mathbf{SPS}^{\mathbf{B}}\mathbf{PF} \\ & - \mathbf{YPS} - \mathbf{SPY} + 2\,\mathbf{SPYPS},\end{aligned} \tag{D3}$$

with

$$\mathbf{Y} = \mathbf{F}^{(\mathbf{B})} - \mathbf{G}\left[\mathbf{PS}^{\mathbf{B}}\mathbf{P}\right]. \tag{D4}$$

The difference between $\mathbf{F}^{\mathbf{B}}$ and $\mathbf{F}^{(\mathbf{B})}$ is that the latter does by definition not include contributions containing the perturbed density $\mathbf{P}^{\mathbf{B}}$. For computing the second derivative of the density matrix $\mathbf{P}^{\mathbf{Bm}}$, the nuclear magnetic moment derivatives $\mathbf{A}^{\mathbf{m}}\left[\mathbf{P}^{\mathbf{B}}\right]$ and $\mathbf{b}^{\mathbf{Bm}}$ are also needed as shown in Eq. (45),

$$\begin{aligned}\mathbf{A}^{\mathbf{m}}\left[\mathbf{P}^{\mathbf{B}}\right] = {}& 3\,\mathbf{F}^{\mathbf{m}}\mathbf{P}^{\mathbf{B}}\mathbf{S} + 3\,\mathbf{SP}^{\mathbf{B}}\mathbf{F}^{\mathbf{m}} - 2\,\mathbf{F}^{\mathbf{m}}\mathbf{P}^{\mathbf{B}}\mathbf{SPS} - 2\,\mathbf{FP}^{\mathbf{B}}\mathbf{SP}^{\mathbf{m}}\mathbf{S} \\ & - 2\,\mathbf{SP}^{\mathbf{m}}\mathbf{SP}^{\mathbf{B}}\mathbf{F} - 2\,\mathbf{SPSP}^{\mathbf{B}}\mathbf{F}^{\mathbf{m}} - 4\,\mathbf{F}^{\mathbf{m}}\mathbf{PSP}^{\mathbf{B}}\mathbf{S} \\ & - 4\,\mathbf{FP}^{\mathbf{m}}\mathbf{SP}^{\mathbf{B}}\mathbf{S} - 4\,\mathbf{SP}^{\mathbf{B}}\mathbf{SP}^{\mathbf{m}}\mathbf{F} - 4\,\mathbf{SP}^{\mathbf{B}}\mathbf{SPF}^{\mathbf{m}} + \mathbf{G}\left[\mathbf{X}^{\mathbf{m}}\right]\mathbf{PS} \\ & + \mathbf{G}[\mathbf{X}]\mathbf{P}^{\mathbf{m}}\mathbf{S} + \mathbf{SP}^{\mathbf{m}}\mathbf{G}[\mathbf{X}] + \mathbf{SPG}\left[\mathbf{X}^{\mathbf{m}}\right] - 2\,\mathbf{SP}^{\mathbf{m}}\mathbf{G}[\mathbf{X}]\mathbf{PS} \\ & - 2\,\mathbf{SPG}\left[\mathbf{X}^{\mathbf{m}}\right]\mathbf{PS} - 2\,\mathbf{SPG}[\mathbf{X}]\mathbf{P}^{\mathbf{m}}\mathbf{S},\end{aligned} \tag{D5}$$

where

$$\mathbf{X}^{\mathbf{m}} = \mathbf{P}^{\mathbf{B}}\mathbf{SP}^{\mathbf{m}} + \mathbf{P}^{\mathbf{m}}\mathbf{SP}^{\mathbf{B}} - 2\,\mathbf{P}^{\mathbf{m}}\mathbf{SP}^{\mathbf{B}}\mathbf{SP} - 2\,\mathbf{PSP}^{\mathbf{B}}\mathbf{SP}^{\mathbf{m}} \tag{D6}$$

and

$$\begin{aligned}\mathbf{b}^{\mathbf{Bm}} = {}& \mathbf{F}^{\mathbf{m}}\mathbf{PS}^{\mathbf{B}} + \mathbf{FP}^{\mathbf{m}}\mathbf{S}^{\mathbf{B}} + \mathbf{S}^{\mathbf{B}}\mathbf{P}^{\mathbf{m}}\mathbf{F} + \mathbf{S}^{\mathbf{B}}\mathbf{PF}^{\mathbf{m}} + 2\,\mathbf{F}^{\mathbf{m}}\mathbf{PS}^{\mathbf{B}}\mathbf{PS} \\ & + 2\,\mathbf{FP}^{\mathbf{m}}\mathbf{S}^{\mathbf{B}}\mathbf{PS} + 2\,\mathbf{FPS}^{\mathbf{B}}\mathbf{P}^{\mathbf{m}}\mathbf{S} + 2\,\mathbf{SP}^{\mathbf{m}}\mathbf{S}^{\mathbf{B}}\mathbf{PF} \\ & + 2\,\mathbf{SPS}^{\mathbf{B}}\mathbf{P}^{\mathbf{m}}\mathbf{F} + 2\,\mathbf{SPS}^{\mathbf{B}}\mathbf{PF}^{\mathbf{m}} - \mathbf{Y}^{\mathbf{m}}\mathbf{PS} - \mathbf{YP}^{\mathbf{m}}\mathbf{S} \\ & - \mathbf{SP}^{\mathbf{m}}\mathbf{Y} - \mathbf{SPY}^{\mathbf{m}} + 2\,\mathbf{SP}^{\mathbf{m}}\mathbf{YPS} + 2\,\mathbf{SPY}^{\mathbf{m}}\mathbf{PS} + 2\,\mathbf{SPYP}^{\mathbf{m}}\mathbf{S},\end{aligned} \tag{D7}$$

where

$$\mathbf{Y}^{\mathbf{m}} = \mathbf{h}^{\mathbf{Bm}} + \mathbf{G}^{\mathbf{B}}\left[\mathbf{P}^{\mathbf{m}}\right] - \mathbf{G}\left[\mathbf{P}^{\mathbf{m}}\mathbf{S}^{\mathbf{B}}\mathbf{P} + \mathbf{PS}^{\mathbf{B}}\mathbf{P}^{\mathbf{m}}\right]. \tag{D8}$$

As it is commonly done in implementations of NMR shieldings, we do not use complex-valued matrices explicitly, but instead we treat purely imaginary matrices using skew-symmetric, real-valued matrices. In this case, one needs to be careful to use the correct sign for each term; whenever a term contains two imaginary matrices, like the term $\mathbf{F}^{\mathbf{m}}\mathbf{PS}^{\mathbf{B}}$ with $\mathbf{F}^{\mathbf{m}}$ and $\mathbf{S}^{\mathbf{B}}$, an additional sign change needs to be applied because both imaginary matrices carry a factor i and $i^2 = -1$.

Instead of iteratively solving for the full $\mathbf{P}^{\mathbf{Bm}}$ as suggested by Eq. (45), we only solve for the occupied-virtual (ov) and virtual-occupied (vo) subspace projections. For the occupied–occupied (oo) and virtual–virtual (vv) subspace projections, explicit expressions can be derived by differentiating the idempotency condition $\mathbf{PSP} = \mathbf{P}$ twice and projecting onto the corresponding subspaces as follows:[29]

$$\mathrm{Tr}[\mathcal{P}\mathbf{P}^{\mathbf{Bm}}] = \mathrm{Tr}[\mathcal{P}(\mathbf{P}_{ov}^{\mathbf{Bm}} + \mathbf{P}_{vo}^{\mathbf{Bm}})] + \mathrm{Tr}[\mathcal{P}(\mathbf{P}_{oo}^{\mathbf{Bm}} + \mathbf{P}_{vv}^{\mathbf{Bm}})], \tag{D9}$$

with

$$\begin{aligned}\mathbf{P}_{oo}^{\mathbf{Bm}} + \mathbf{P}_{vv}^{\mathbf{Bm}} = {}& -2\big(\mathbf{PSP}^{\mathbf{B}}\mathbf{SP}^{\mathbf{m}}\mathbf{SP} + \mathbf{PSP}^{\mathbf{m}}\mathbf{S}^{\mathbf{B}}\mathbf{P} \\ & + \mathbf{PS}^{\mathbf{B}}\mathbf{P}^{\mathbf{m}}\mathbf{SP} + \mathbf{PSP}^{\mathbf{m}}\mathbf{SP}^{\mathbf{B}}\mathbf{SP}\big) \\ & + \mathbf{P}^{\mathbf{B}}\mathbf{SP}^{\mathbf{m}} + \mathbf{P}^{\mathbf{m}}\mathbf{SP}^{\mathbf{B}} + \mathbf{P}^{\mathbf{m}}\mathbf{S}^{\mathbf{B}}\mathbf{P} + \mathbf{PS}^{\mathbf{B}}\mathbf{P}^{\mathbf{m}}.\end{aligned} \tag{D10}$$

The computation of the ov- and vo-projections of $\mathbf{P}^{\mathbf{Bm}}$ in the first term on the right-hand side of Eq. (D9) can be avoided using a Z-vector approach,

$$\mathrm{Tr}[\mathcal{P}(\mathbf{P}_{ov}^{\mathbf{Bm}} + \mathbf{P}_{vo}^{\mathbf{Bm}})] = \mathrm{Tr}\Big[(\mathbf{Z}_{ov} + \mathbf{Z}_{vo})\big(\mathbf{b}^{\mathbf{Bm}} - \mathbf{A}^{\mathbf{m}}\left[\mathbf{P}^{\mathbf{B}}\right]\big)\Big]. \tag{D11}$$

Here, $\mathbf{Z}_{ov}$ and $\mathbf{Z}_{vo}$ are the ov- and vo-projections of the Z-vector $\mathbf{Z}_{\mathcal{P}} = \mathbf{A}^{-1}\mathcal{P}$.

After eliminating $\mathbf{P}^{\mathbf{Bm}}$ from the equations, $\mathbf{P}^{\mathbf{m}}$ remains to be eliminated. For this, we will show how terms involving $\mathbf{P}^{\mathbf{m}}$ can be rearranged in the form $\mathrm{Tr}[\ldots\mathbf{P}^{\mathbf{m}}]$, which then allows us to apply another Z-vector approach. The terms involving $\mathbf{P}_{oo}^{\mathbf{Bm}} + \mathbf{P}_{vv}^{\mathbf{Bm}}$ can be brought into this form by applying cyclic permutations under the trace,

$$\mathrm{Tr}\left[\mathcal{P}\left(\mathbf{P}_{oo}^{\mathbf{Bm}} + \mathbf{P}_{vv}^{\mathbf{Bm}}\right)\right] = \mathrm{Tr}\left[\mathbf{O}_{[\mathbf{P}_{oo}^{\mathbf{Bm}} + \mathbf{P}_{vv}^{\mathbf{Bm}}]}\mathbf{P}^{\mathbf{m}}\right], \tag{D12}$$

where

$$\begin{aligned}\mathbf{O}_{[\mathbf{P}_{oo}^{\mathbf{Bm}} + \mathbf{P}_{vv}^{\mathbf{Bm}}]} = {}& -2\big(\mathbf{SP}\mathcal{P}\mathbf{PSP}^{\mathbf{B}}\mathbf{S} + \mathbf{S}^{\mathbf{B}}\mathbf{P}\mathcal{P}\mathbf{PS} + \mathbf{SP}\mathcal{P}\mathbf{PS}^{\mathbf{B}} + \mathbf{SP}^{\mathbf{B}}\mathbf{SP}\mathcal{P}\mathbf{PS}\big) \\ & + \big(\mathcal{P}\mathbf{P}^{\mathbf{B}}\mathbf{S} + \mathbf{SP}^{\mathbf{B}}\mathcal{P} + \mathbf{S}^{\mathbf{B}}\mathbf{P}\mathcal{P} + \mathcal{P}\mathbf{PS}^{\mathbf{B}}\big).\end{aligned} \tag{D13}$$

Next, we focus on the expression $\mathrm{Tr}[(\mathbf{Z}_{ov} + \mathbf{Z}_{vo})\big(\mathbf{b}^{\mathbf{Bm}} - \mathbf{A}^{\mathbf{m}}\left[\mathbf{P}^{\mathbf{B}}\right]\big)]$. First, we insert Eqs. (D5) and (D7). After applying cyclic permutations, we arrive at

$$\begin{aligned}&\mathrm{Tr}\big[(\mathbf{Z}_{ov}+\mathbf{Z}_{vo})\big(\mathbf{b}^{\mathbf{Bm}}-\mathbf{A}^{\mathbf{m}}\big[\mathbf{P}^{\mathbf{B}}\big]\big)\big]\\&\quad=\mathrm{Tr}\Big[\mathbf{O}_{\mathbf{P}^{\mathrm{m}}}\mathbf{P}^{\mathbf{m}}+\mathbf{O}_{\mathbf{F}^{\mathrm{m}}}\mathbf{F}^{\mathbf{m}}+\mathbf{O}_{\mathbf{Y}^{\mathrm{m}}}\big(\mathbf{G}\big[\mathbf{X}^{\mathbf{m}}\big]+\mathbf{Y}^{\mathbf{m}}\big)\Big],\end{aligned}\tag{D14}$$

where

$$\begin{aligned}\mathbf{O}_{\mathbf{P}^{\mathrm{m}}}=\;&\mathbf{S}^{\mathbf{B}}\bar{\mathbf{Z}}\mathbf{F}+\mathbf{F}\bar{\mathbf{Z}}\mathbf{S}^{\mathbf{B}}+2\,\mathbf{S}^{\mathbf{B}}\mathbf{P}\mathbf{S}\bar{\mathbf{Z}}\mathbf{F}+2\,\mathbf{S}\bar{\mathbf{Z}}\mathbf{F}\mathbf{P}\mathbf{S}^{\mathbf{B}}+2\,\mathbf{S}^{\mathbf{B}}\mathbf{P}\mathbf{F}\bar{\mathbf{Z}}\mathbf{S}\\&+2\,\mathbf{F}\bar{\mathbf{Z}}\mathbf{S}\mathbf{P}\mathbf{S}^{\mathbf{B}}-\mathbf{S}\bar{\mathbf{Z}}\mathbf{Y}-\mathbf{Y}\bar{\mathbf{Z}}\mathbf{S}+2\,\mathbf{Y}\mathbf{P}\mathbf{S}\bar{\mathbf{Z}}\mathbf{S}+2\,\mathbf{S}\bar{\mathbf{Z}}\mathbf{S}\mathbf{P}\mathbf{Y}\\&+2\,\mathbf{S}\bar{\mathbf{Z}}\mathbf{F}\mathbf{P}^{\mathbf{B}}\mathbf{S}+2\,\mathbf{S}\mathbf{P}^{\mathbf{B}}\mathbf{F}\bar{\mathbf{Z}}\mathbf{S}+4\,\mathbf{S}\mathbf{P}^{\mathbf{B}}\mathbf{S}\bar{\mathbf{Z}}\mathbf{F}+4\,\mathbf{F}\bar{\mathbf{Z}}\mathbf{S}\mathbf{P}^{\mathbf{B}}\mathbf{S}\\&-\mathbf{S}\bar{\mathbf{Z}}\mathbf{G}[\mathbf{X}]-\mathbf{G}[\mathbf{X}]\bar{\mathbf{Z}}\mathbf{S}+2\,\mathbf{G}[\mathbf{X}]\mathbf{P}\mathbf{S}\bar{\mathbf{Z}}\mathbf{S}+2\,\mathbf{S}\bar{\mathbf{Z}}\mathbf{S}\mathbf{P}\mathbf{G}[\mathbf{X}],\end{aligned}\tag{D15}$$

$$\begin{aligned}\mathbf{O}_{\mathbf{F}^{\mathrm{m}}}=\;&\mathbf{P}\mathbf{S}^{\mathbf{B}}\bar{\mathbf{Z}}+\bar{\mathbf{Z}}\mathbf{S}^{\mathbf{B}}\mathbf{P}+2\,\mathbf{P}\mathbf{S}^{\mathbf{B}}\mathbf{P}\mathbf{S}\bar{\mathbf{Z}}+2\,\bar{\mathbf{Z}}\mathbf{S}\mathbf{P}\mathbf{S}^{\mathbf{B}}\mathbf{P}-3\,\mathbf{P}^{\mathbf{B}}\mathbf{S}\bar{\mathbf{Z}}-3\,\bar{\mathbf{Z}}\mathbf{S}\mathbf{P}^{\mathbf{B}}\\&+2\,\mathbf{P}^{\mathbf{B}}\mathbf{S}\mathbf{P}\mathbf{S}\bar{\mathbf{Z}}+2\,\bar{\mathbf{Z}}\mathbf{S}\mathbf{P}\mathbf{S}\mathbf{P}^{\mathbf{B}}+4\,\mathbf{P}\mathbf{S}\mathbf{P}^{\mathbf{B}}\mathbf{S}\bar{\mathbf{Z}}+4\,\bar{\mathbf{Z}}\mathbf{S}\mathbf{P}^{\mathbf{B}}\mathbf{S}\mathbf{P},\end{aligned}\tag{D16}$$

and

$$\mathbf{O}_{\mathbf{Y}^{\mathrm{m}}}=-\mathbf{P}\mathbf{S}\bar{\mathbf{Z}}-\bar{\mathbf{Z}}\mathbf{S}\mathbf{P}+2\,\mathbf{P}\mathbf{S}\bar{\mathbf{Z}}\mathbf{S}\mathbf{P},\tag{D17}$$

where $\bar{\mathbf{Z}}\equiv\mathbf{Z}_{ov}+\mathbf{Z}_{vo}$.

Equation (D14) contains the Fock matrix derivative $\mathbf{F}^{\mathbf{m}}$, which also needs to be circumvented in the all-nuclei formulation. For this, we use the following identities:

$$\mathrm{Tr}\big[\mathbf{W}\mathbf{F}^{\mathbf{m}}\big]=\mathrm{Tr}\big[\mathbf{W}\mathbf{h}^{\mathbf{m}}\big]+\mathrm{Tr}\big[\mathbf{W}\mathbf{G}\big[\mathbf{P}^{\mathbf{m}}\big]\big],\tag{D18}$$

$$\mathrm{Tr}\big[\mathbf{W}\mathbf{G}\big[\mathbf{P}^{\mathbf{m}}\big]\big]=\mathrm{Tr}\big[\mathbf{G}[\mathbf{W}]\mathbf{P}^{\mathbf{m}}\big],\tag{D19}$$

where $\mathbf{W}$ is some general matrix and $\mathbf{G}[\mathbf{W}]$ is

$$G_{\mu\nu}[\mathbf{W}]\equiv\sum_{\lambda\sigma}W_{\lambda\sigma}[2(\mu\nu|\lambda\sigma)-(\mu\sigma|\lambda\nu)].\tag{D20}$$

Using Eqs. (D18) and (D19), the contribution from $\mathbf{F}^{\mathbf{m}}$ in Eq. (D14) can be rearranged as follows:

$$\mathrm{Tr}\big[\mathbf{O}_{\mathbf{F}^{\mathrm{m}}}\mathbf{F}^{\mathbf{m}}\big]=\mathrm{Tr}\big[\mathbf{O}_{\mathbf{F}^{\mathrm{m}}}\mathbf{h}^{\mathbf{m}}\big]+\mathrm{Tr}\big[\mathbf{G}\big[\mathbf{O}_{\mathbf{F}^{\mathrm{m}}}\big]\mathbf{P}^{\mathbf{m}}\big].\tag{D21}$$

Similarly, the terms involving $\mathbf{G}[\mathbf{X}^{\mathbf{m}}]$ and $\mathbf{Y}^{\mathbf{m}}$ in Eq. (D14) can be rearranged by using Eq. (D19) and then applying cyclic permutations as follows:

$$\begin{aligned}\mathrm{Tr}\big[\mathbf{O}_{\mathbf{Y}^{\mathrm{m}}}\mathbf{G}\big[\mathbf{X}^{\mathbf{m}}\big]\big]&=\mathrm{Tr}\big[\mathbf{G}\big[\mathbf{O}_{\mathbf{Y}^{\mathrm{m}}}\big]\mathbf{X}^{\mathbf{m}}\big]\\&=\mathrm{Tr}\big[\big\{\mathbf{G}\big[\mathbf{O}_{\mathbf{Y}^{\mathrm{m}}}\big]\mathbf{P}^{\mathbf{B}}\mathbf{S}+\mathbf{S}\mathbf{P}^{\mathbf{B}}\mathbf{G}\big[\mathbf{O}_{\mathbf{Y}^{\mathrm{m}}}\big]\\&\quad-2\mathbf{S}\mathbf{P}^{\mathbf{B}}\mathbf{S}\mathbf{P}\,\mathbf{G}\big[\mathbf{O}_{\mathbf{Y}^{\mathrm{m}}}\big]\;-2\mathbf{G}\big[\mathbf{O}_{\mathbf{Y}^{\mathrm{m}}}\big]\mathbf{P}\mathbf{S}\mathbf{P}^{\mathbf{B}}\mathbf{S}\big\}\mathbf{P}^{\mathbf{m}}\big],\end{aligned}\tag{D22}$$

$$\begin{aligned}\mathrm{Tr}\big[\mathbf{O}_{\mathbf{Y}^{\mathrm{m}}}\mathbf{Y}^{\mathbf{m}}\big]&=\mathrm{Tr}\big[\mathbf{O}_{\mathbf{Y}^{\mathrm{m}}}\mathbf{h}^{\mathbf{Bm}}\big]+\mathrm{Tr}\big[\mathbf{G}^{\mathbf{B}}\big[\mathbf{O}_{\mathbf{Y}^{\mathrm{m}}}\big]\mathbf{P}^{\mathbf{m}}\big]\\&\quad-\mathrm{Tr}\big[\mathbf{G}\big[\mathbf{O}_{\mathbf{Y}^{\mathrm{m}}}\big]\big(\mathbf{P}^{\mathbf{m}}\mathbf{S}^{\mathbf{B}}\mathbf{P}+\mathbf{P}\mathbf{S}^{\mathbf{B}}\mathbf{P}^{\mathbf{m}}\big)\big]\\&=\mathrm{Tr}\big[\mathbf{O}_{\mathbf{Y}^{\mathrm{m}}}\mathbf{h}^{\mathbf{Bm}}\big]+\mathrm{Tr}\big[\mathbf{G}^{\mathbf{B}}\big[\mathbf{O}_{\mathbf{Y}^{\mathrm{m}}}\big]\mathbf{P}^{\mathbf{m}}\big]\\&\quad-\mathrm{Tr}\big[\big(\mathbf{S}^{\mathbf{B}}\mathbf{P}\mathbf{G}\big[\mathbf{O}_{\mathbf{Y}^{\mathrm{m}}}\big]+\mathbf{G}\big[\mathbf{O}_{\mathbf{Y}^{\mathrm{m}}}\big]\mathbf{P}\mathbf{S}^{\mathbf{B}}\big)\mathbf{P}^{\mathbf{m}}\big].\end{aligned}\tag{D23}$$

Using the results from Eqs. (D18)–(D23) allows us to rewrite the expression $\mathrm{Tr}[(\mathbf{Z}_{ov}+\mathbf{Z}_{vo})(\mathbf{b}^{\mathbf{Bm}}-\mathbf{A}^{\mathbf{m}}[\mathbf{P}^{\mathbf{B}}])]$ in the following form:

$$\begin{aligned}&\mathrm{Tr}\big[(\mathbf{Z}_{ov}+\mathbf{Z}_{vo})\big(\mathbf{b}^{\mathbf{Bm}}-\mathbf{A}^{\mathbf{m}}\big[\mathbf{P}^{\mathbf{B}}\big]\big)\big]\\&\quad=\mathrm{Tr}\big[\mathbf{O}_{\mathbf{F}^{\mathrm{m}}}\mathbf{h}^{\mathbf{m}}\big]+\mathrm{Tr}\big[\mathbf{O}_{\mathbf{Y}^{\mathrm{m}}}\mathbf{h}^{\mathbf{Bm}}\big]+\mathrm{Tr}\big[\tilde{\mathbf{O}}\mathbf{P}^{\mathbf{m}}\big],\end{aligned}\tag{D24}$$

where

$$\begin{aligned}\tilde{\mathbf{O}}=\;&\mathbf{O}_{\mathbf{P}^{\mathrm{m}}}+\mathbf{G}\big[\mathbf{O}_{\mathbf{F}^{\mathrm{m}}}\big]+\mathbf{G}^{\mathbf{B}}\big[\mathbf{O}_{\mathbf{Y}^{\mathrm{m}}}\big]+\mathbf{G}\big[\mathbf{O}_{\mathbf{Y}^{\mathrm{m}}}\big]\mathbf{P}^{\mathbf{B}}\mathbf{S}\\&+\mathbf{S}\mathbf{P}^{\mathbf{B}}\mathbf{G}\big[\mathbf{O}_{\mathbf{Y}^{\mathrm{m}}}\big]-2\,\mathbf{S}\mathbf{P}^{\mathbf{B}}\mathbf{S}\mathbf{P}\mathbf{G}\big[\mathbf{O}_{\mathbf{Y}^{\mathrm{m}}}\big]-2\,\mathbf{G}\big[\mathbf{O}_{\mathbf{Y}^{\mathrm{m}}}\big]\mathbf{P}\mathbf{S}\mathbf{P}^{\mathbf{B}}\mathbf{S}\\&-\mathbf{S}^{\mathbf{B}}\mathbf{P}\mathbf{G}\big[\mathbf{O}_{\mathbf{Y}^{\mathrm{m}}}\big]-\mathbf{G}\big[\mathbf{O}_{\mathbf{Y}^{\mathrm{m}}}\big]\mathbf{P}\mathbf{S}^{\mathbf{B}}.\end{aligned}\tag{D25}$$

Combining $\tilde{\mathbf{O}}$, $\mathbf{O}_{[\mathbf{P}_{oo}^{\mathbf{Bm}}+\mathbf{P}_{vv}^{\mathbf{Bm}}]}$, and $\mathcal{P}^{\mathbf{B}}$ from the term $\mathrm{Tr}[\mathcal{P}^{\mathbf{B}}\mathbf{P}^{\mathbf{m}}]$ allows us to compactly express all remaining terms involving $\mathbf{P}^{\mathbf{m}}$ in the form $\mathrm{Tr}[\mathbf{O}\mathbf{P}^{\mathbf{m}}]$,

$$\mathbf{O}=\tilde{\mathbf{O}}+\mathbf{O}_{\left[\mathbf{P}_{oo}^{\mathbf{Bm}}+\mathbf{P}_{vv}^{\mathbf{Bm}}\right]}+\mathcal{P}^{\mathbf{B}}.\tag{D26}$$

This term now has the proper form for applying a Z-vector approach, which allows us to circumvent $\mathbf{P}^{\mathbf{m}}$,

$$\mathrm{Tr}\big[\mathbf{O}\mathbf{P}^{\mathbf{m}}\big]=\mathrm{Tr}\big[\mathbf{Z}_{\mathbf{O}}\mathbf{b}^{\mathbf{m}}\big],\tag{D27}$$

where

$$\mathbf{Z}_{\mathbf{O}}=\mathbf{A}^{-1}\mathbf{O}.\tag{D28}$$

In our implementation, the Z-vector equation for $\mathbf{Z}_{\mathbf{O}}$ is solved iteratively using the DL-CPSCF approach from the study by Beer and Ochsenfeld.[73]

REFERENCES

[1] J. Gauss, J. Chem. Phys. **99**, 3629 (1993).

[2] T. Helgaker, M. Jaszuński, and K. Ruud, Chem. Rev. **99**, 293 (1999).

[3] L. B. Casabianca and A. C. De Dios, J. Chem. Phys. **128**, 052201 (2008).

[4] R. Ditchfield, Mol. Phys. **27**, 789 (1974).

[5] K. Wolinski, J. F. Hinton, and P. Pulay, J. Am. Chem. Soc. **112**, 8251 (1990).

[6] M. Häser, R. Ahlrichs, H. Baron, P. Weis, and H. Horn, Theor. Chim. Acta **83**, 455 (1992).

[7] R. M. Stevens, R. M. Pitzer, and W. N. Lipscomb, J. Chem. Phys. **38**, 550 (1963).

[8] H. Larsen, T. Helgaker, J. Olsen, and P. Jørgensen, J. Chem. Phys. **115**, 10344 (2001).

[9] G. Schreckenbach and T. Ziegler, J. Phys. Chem. **99**, 606 (1995).

[10] J. R. Cheeseman, G. W. Trucks, T. A. Keith, and M. J. Frisch, J. Chem. Phys. **104**, 5497 (1996).

[11] G. Rauhut, S. Puyear, K. Wolinski, and P. Pulay, J. Phys. Chem. **100**, 6310 (1996).

[12] C. Ochsenfeld, J. Kussmann, and F. Koziol, Angew. Chem. **116**, 4585 (2004).

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

[14] M. Beer, J. Kussmann, and C. Ochsenfeld, J. Chem. Phys. **134**, 074102 (2011).

[15] J. Gauss, Chem. Phys. Lett. **191**, 614 (1992).

[16] J. Gauss and J. F. Stanton, Adv. Chem. Phys. **123**, 355 (2002).

[17] M. R. Hoffmann, D. J. Fox, J. F. Gaw, Y. Osamura, Y. Yamaguchi, R. S. Grev, G. Fitzgerald, H. F. Schaefer III, P. J. Knowles, and N. C. Handy, J. Chem. Phys. **80**, 2660 (1984).

[18] K. Ruud, T. Helgaker, R. Kobayashi, P. Jørgensen, K. L. Bak, and H. J. A. Jensen, J. Chem. Phys. **100**, 8178 (1994).

[19] J. Gauss and J. F. Stanton, J. Chem. Phys. **102**, 251 (1995).

[20] J. Gauss and J. F. Stanton, J. Chem. Phys. **104**, 2574 (1996).

[21] M. Kállay and J. Gauss, J. Chem. Phys. **120**, 6841 (2004).

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

[23] N. C. Handy and H. F. Schaefer III, J. Chem. Phys. **81**, 5031 (1984).
[24] Y. Yamaguchi, H. F. Schaefer III, Y. Osamura, J. D. Goddard *et al.*, *A New Dimension to Quantum Chemistry: Analytic Derivative Methods in Ab Initio Molecular Electronic Structure Theory* (Oxford University Press, 1994).
[25] M. Kollwitz and J. Gauss, Chem. Phys. Lett. **260**, 639 (1996).
[26] J. Gauss and H.-J. Werner, Phys. Chem. Chem. Phys. **2**, 2083 (2000).
[27] S. Loibl and M. Schütz, J. Chem. Phys. **137**, 084107 (2012).
[28] G. L. Stoychev, A. A. Auer, J. Gauss, and F. Neese, J. Chem. Phys. **154**, 164110 (2021).
[29] M. Maurer and C. Ochsenfeld, J. Chem. Phys. **138**, 174104 (2013).
[30] J. Almlöf, Chem. Phys. Lett. **181**, 319 (1991).
[31] M. Häser and J. Almlöf, J. Chem. Phys. **96**, 489 (1992).
[32] M. Häser, Theor. Chim. Acta **87**, 147 (1993).
[33] P. Y. Ayala and G. E. Scuseria, J. Chem. Phys. **110**, 3660 (1999).
[34] S. A. Maurer, D. S. Lambrecht, D. Flaig, and C. Ochsenfeld, J. Chem. Phys. **136**, 144107 (2012).
[35] S. A. Maurer, D. S. Lambrecht, J. Kussmann, and C. Ochsenfeld, J. Chem. Phys. **138**, 014101 (2013).
[36] S. Grimme, J. Chem. Phys. **124**, 034108 (2006).
[37] G. L. Stoychev, A. A. Auer, and F. Neese, J. Chem. Theory Comput. **14**, 4756 (2018).
[38] S. Grimme, J. Chem. Phys. **118**, 9095 (2003).
[39] M. Maurer and C. Ochsenfeld, J. Chem. Theory Comput. **11**, 37 (2015).
[40] Y. Jung, R. C. Lochan, A. D. Dutoi, and M. Head-Gordon, J. Chem. Phys. **121**, 9793 (2004).
[41] F. London, J. Phys. Radium **8**, 397 (1937).
[42] W. Kutzelnigg, Isr. J. Chem. **19**, 193 (1980).
[43] M. Feyereisen, G. Fitzgerald, and A. Komornicki, Chem. Phys. Lett. **208**, 359 (1993).
[44] J. L. Whitten, J. Chem. Phys. **58**, 4496 (1973).
[45] M. Glasbrenner, D. Graf, and C. Ochsenfeld, J. Chem. Theory Comput. **16**, 6856 (2020).
[46] Y. Jung, A. Sodt, P. M. W. Gill, and M. Head-Gordon, Proc. Natl. Acad. Sci. U. S. A. **102**, 6692 (2005).
[47] Y. Jung, Y. Shao, and M. Head-Gordon, J. Comput. Chem. **28**, 1953 (2007).
[48] S. Reine, E. Tellgren, A. Krapp, T. Kjærgaard, T. Helgaker, B. Jansik, S. Høst, and P. Salek, J. Chem. Phys. **129**, 104101 (2008).
[49] S. Burger, F. Lipparini, J. Gauss, and S. Stopkowicz, J. Chem. Phys. **155**, 074105 (2021).
[50] N. H. F. Beebe and J. Linderberg, Int. J. Quantum Chem. **12**, 683 (1977).
[51] H. Koch, A. Sánchez de Merás, and T. B. Pedersen, J. Chem. Phys. **118**, 9481 (2003).
[52] F. Aquilante, T. B. Pedersen, A. Sánchez de Merás, and H. Koch, J. Chem. Phys. **125**, 174101 (2006).
[53] N. J. Higham, Wiley Interdiscip. Rev.: Comput. Stat. **1**, 251 (2009).
[54] H. Harbrecht, M. Peters, and R. Schneider, Appl. Numer. Math. **62**, 428 (2012).
[55] Y. Jung and M. Head-Gordon, Phys. Chem. Chem. Phys. **8**, 2831 (2006).
[56] S. Schweizer, B. Doser, and C. Ochsenfeld, J. Chem. Phys. **128**, 154101 (2008).
[57] P. R. Surján, Chem. Phys. Lett. **406**, 318 (2005).
[58] C. A. White, B. G. Johnson, P. M. W. Gill, and M. Head-Gordon, Chem. Phys. Lett. **230**, 8 (1994).
[59] C. A. White, B. G. Johnson, P. M. W. Gill, and M. Head-Gordon, Chem. Phys. Lett. **253**, 268 (1996).
[60] F. Weigend, Phys. Chem. Chem. Phys. **4**, 4285 (2002).
[61] C. Ochsenfeld, C. A. White, and M. Head-Gordon, J. Chem. Phys. **109**, 1663 (1998).
[62] C. Ochsenfeld, Chem. Phys. Lett. **327**, 216 (2000).
[63] H. Laqua, T. H. Thompson, J. Kussmann, and C. Ochsenfeld, J. Chem. Theory Comput. **16**, 1456 (2020).
[64] C. Ochsenfeld and M. Head-Gordon, Chem. Phys. Lett. **270**, 399 (1997).
[65] J. Zienau, L. Clin, B. Doser, and C. Ochsenfeld, J. Chem. Phys. **130**, 204112 (2009).
[66] S. A. Maurer, L. Clin, and C. Ochsenfeld, J. Chem. Phys. **140**, 224112 (2014).
[67] A. Luenser, H. F. Schurkus, and C. Ochsenfeld, J. Chem. Theory Comput. **13**, 1647 (2017).
[68] J. Kussmann and C. Ochsenfeld, J. Chem. Phys. **138**, 134114 (2013).
[69] J. Kussmann and C. Ochsenfeld, J. Chem. Theory Comput. **11**, 918 (2015).
[70] J. Kussmann and C. Ochsenfeld, J. Chem. Theory Comput. **13**, 3153 (2017).
[71] F. Weigend and R. Ahlrichs, Phys. Chem. Chem. Phys. **7**, 3297 (2005).
[72] F. Weigend, M. Häser, H. Patzelt, and R. Ahlrichs, Chem. Phys. Lett. **294**, 143 (1998).
[73] M. Beer and C. Ochsenfeld, J. Chem. Phys. **128**, 221102 (2008).
[74] M. Kaltak, J. Klimeš, and G. Kresse, J. Chem. Theory Comput. **10**, 2498 (2014).
[75] A. Takatsuka, S. Ten-No, and W. Hackbusch, J. Chem. Phys. **129**, 044112 (2008).
[76] TURBOMOLE V7.4 2019, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from http://www.turbomole.com.
[77] M. Kollwitz, M. Häser, and J. Gauss, J. Chem. Phys. **108**, 8295 (1998).
[78] P. Jurečka, J. Šponer, J. Černý, and P. Hobza, Phys. Chem. Chem. Phys. **8**, 1985 (2006).
[79] C. Hättig, Phys. Chem. Chem. Phys. **7**, 59 (2005).
[80] F. Weigend, A. Köhn, and C. Hättig, J. Chem. Phys. **116**, 3175 (2002).
[81] R. Sedlak, T. Janowski, M. Pitonak, J. Rezac, P. Pulay, and P. Hobza, J. Chem. Theory Comput. **9**, 3364 (2013).
[82] M. Häser and R. Ahlrichs, J. Comput. Chem. **10**, 104 (1989).
[83] C. Moler and C. Van Loan, SIAM Rev. **45**, 3 (2003).
[84] T. Fung, Int. J. Numer. Methods Eng. **59**, 1273 (2004).

Supplementary Material for:
Efficient low-scaling computation of NMR shieldings at the second-order Møller-Plesset perturbation theory level with Cholesky-decomposed densities and an attenuated Coulomb-metric

Michael Glasbrenner[1], Sigurd Vogler[1], and Christian Ochsenfeld[1,2]

[1]Chair of Theoretical Chemistry, Department of Chemistry, University of Munich (LMU), Butenandtstr. 7, D-81377 Munich, Germany
[2]Max Planck Institute for Solid State Research, Heisenbergstr. 1, D-70569 Stuttgart, Germany

August 27, 2021

List of Tables

Table 1: MP2 shielding calculations for all dimers from the S22 test set.[1] "Loose" and "Tight" settings are defined in the text. "Tight*" denotes shieldings computed with "Tight" settings and a cc-pVQZ-RI auxiliary basis set instead of a def2-SVP-RI basis set. MP2 shieldings computed with Turbomole are used as a reference. All values are given in ppm.

Molecule	Atom	Loose	Tight	Tight*	Ref.
ammonia dimer	N	280.264	280.271	280.257	280.258
ammonia dimer	H	31.915	31.915	31.914	31.914
ammonia dimer	H	31.915	31.915	31.914	31.914
ammonia dimer	H	31.011	31.011	31.010	31.010
ammonia dimer	N	280.255	280.271	280.257	280.258
ammonia dimer	H	31.915	31.915	31.914	31.914
ammonia dimer	H	31.011	31.011	31.010	31.010
ammonia dimer	H	31.915	31.915	31.914	31.914
water dimer	O	365.630	365.648	365.630	365.630
water dimer	H	32.021	32.022	32.020	32.020
water dimer	H	28.707	28.708	28.705	28.705
water dimer	O	358.178	358.195	358.176	358.178
water dimer	H	30.865	30.866	30.864	30.864
water dimer	H	30.865	30.866	30.864	30.864
benzene–methane	C	83.554	83.558	83.588	83.584
benzene–methane	C	83.512	83.508	83.539	83.534
benzene–methane	C	83.562	83.558	83.588	83.585
benzene–methane	C	83.513	83.508	83.539	83.533
benzene–methane	C	83.562	83.558	83.588	83.584
benzene–methane	C	83.512	83.508	83.539	83.534
benzene–methane	H	24.453	24.453	24.454	24.455
benzene–methane	H	24.452	24.452	24.453	24.453
benzene–methane	H	24.454	24.453	24.454	24.455
benzene–methane	H	24.453	24.452	24.453	24.454
benzene–methane	H	24.454	24.453	24.454	24.455
benzene–methane	H	24.452	24.452	24.453	24.454
benzene–methane	C	208.639	208.639	208.634	208.636
benzene–methane	H	32.494	32.493	32.493	32.492
benzene–methane	H	32.494	32.493	32.493	32.492
benzene–methane	H	34.041	34.040	34.040	34.039
benzene–methane	H	32.494	32.493	32.493	32.492
formic acid dimer	C	48.646	48.639	48.662	48.663
formic acid dimer	O	163.655	163.645	163.750	163.746
formic acid dimer	O	23.671	23.658	23.767	23.769
formic acid dimer	H	24.548	24.549	24.547	24.547
formic acid dimer	H	19.312	19.312	19.311	19.313
formic acid dimer	C	48.645	48.639	48.662	48.663
formic acid dimer	O	163.656	163.645	163.750	163.746
formic acid dimer	O	23.650	23.658	23.767	23.769
formic acid dimer	H	24.548	24.549	24.547	24.547
formic acid dimer	H	19.312	19.312	19.311	19.313
formamide dimer	C	54.007	54.004	54.031	54.034
formamide dimer	O	32.213	32.192	32.317	32.329
formamide dimer	N	173.556	173.553	173.611	173.605
formamide dimer	H	27.619	27.619	27.620	27.620
formamide dimer	H	22.559	22.558	22.557	22.558
formamide dimer	H	24.326	24.327	24.327	24.326
formamide dimer	C	54.008	54.004	54.031	54.034
formamide dimer	O	32.215	32.192	32.317	32.329
formamide dimer	N	173.557	173.553	173.611	173.605
formamide dimer	H	27.619	27.619	27.620	27.620
formamide dimer	H	22.559	22.558	22.557	22.558
formamide dimer	H	24.326	24.327	24.327	24.326
uracil H-bonded dimer	O	114.183	114.168	114.292	114.340
uracil H-bonded dimer	C	57.549	57.533	57.547	57.548
uracil H-bonded dimer	N	152.959	152.929	152.991	152.948
uracil H-bonded dimer	C	74.251	74.250	74.298	74.301
uracil H-bonded dimer	C	107.664	107.646	107.688	107.719
uracil H-bonded dimer	C	53.935	53.923	53.944	53.945
uracil H-bonded dimer	N	126.305	126.267	126.361	126.327
uracil H-bonded dimer	O	-17.750	-17.822	-17.699	-17.654
uracil H-bonded dimer	H	26.687	26.691	26.692	26.685
uracil H-bonded dimer	H	20.529	20.526	20.526	20.538
uracil H-bonded dimer	H	25.406	25.402	25.402	25.397
uracil H-bonded dimer	H	25.148	25.150	25.151	25.154
uracil H-bonded dimer	O	114.179	114.168	114.292	114.340
uracil H-bonded dimer	C	57.548	57.533	57.547	57.548
uracil H-bonded dimer	N	152.959	152.929	152.991	152.948
uracil H-bonded dimer	C	74.250	74.250	74.298	74.301
uracil H-bonded dimer	C	107.664	107.646	107.688	107.719
uracil H-bonded dimer	C	53.934	53.923	53.944	53.945
uracil H-bonded dimer	N	126.305	126.267	126.361	126.327
uracil H-bonded dimer	O	-17.750	-17.822	-17.699	-17.654
uracil H-bonded dimer	H	26.687	26.691	26.692	26.685
uracil H-bonded dimer	H	20.529	20.526	20.526	20.538
uracil H-bonded dimer	H	25.406	25.402	25.402	25.397
uracil H-bonded dimer	H	25.148	25.150	25.151	25.154
2-pyridoxine–2-aminopyridine	O	67.683	67.651	67.740	67.740
2-pyridoxine–2-aminopyridine	N	102.458	102.356	102.416	102.460
2-pyridoxine–2-aminopyridine	C	77.077	77.121	77.172	77.181
2-pyridoxine–2-aminopyridine	C	106.148	106.093	106.115	106.097
2-pyridoxine–2-aminopyridine	C	80.473	80.500	80.545	80.557

Continued on next page...

Molecule	Atom	Loose	Tight	Tight*	Ref.
2-pyridoxine–2-aminopyridine	C	49.304	49.274	49.280	49.253
2-pyridoxine–2-aminopyridine	C	86.818	86.752	86.781	86.775
2-pyridoxine–2-aminopyridine	H	24.737	24.744	24.749	24.745
2-pyridoxine–2-aminopyridine	H	26.019	26.021	26.024	26.023
2-pyridoxine–2-aminopyridine	H	24.851	24.850	24.850	24.855
2-pyridoxine–2-aminopyridine	H	25.368	25.369	25.372	25.367
2-pyridoxine–2-aminopyridine	H	17.596	17.583	17.578	17.590
2-pyridoxine–2-aminopyridine	N	32.173	32.011	32.021	32.097
2-pyridoxine–2-aminopyridine	C	51.046	51.047	51.072	51.093
2-pyridoxine–2-aminopyridine	C	99.130	99.068	99.091	99.083
2-pyridoxine–2-aminopyridine	C	79.310	79.373	79.420	79.417
2-pyridoxine–2-aminopyridine	C	96.163	96.101	96.123	96.101
2-pyridoxine–2-aminopyridine	C	69.340	69.392	69.443	69.453
2-pyridoxine–2-aminopyridine	H	25.392	25.391	25.391	25.390
2-pyridoxine–2-aminopyridine	H	24.641	24.646	24.649	24.649
2-pyridoxine–2-aminopyridine	H	25.361	25.360	25.361	25.362
2-pyridoxine–2-aminopyridine	H	24.093	24.090	24.090	24.092
2-pyridoxine–2-aminopyridine	N	205.968	205.981	206.035	206.049
2-pyridoxine–2-aminopyridine	H	28.567	28.567	28.567	28.567
2-pyridoxine–2-aminopyridine	H	23.343	23.334	23.331	23.329
adenine–thymine WC complex	N	60.851	60.756	60.795	60.722
adenine–thymine WC complex	C	55.191	55.182	55.210	55.242
adenine–thymine WC complex	C	87.109	87.055	87.102	87.144
adenine–thymine WC complex	C	63.864	63.901	63.958	63.951
adenine–thymine WC complex	N	52.749	52.538	52.525	52.553
adenine–thymine WC complex	C	60.890	60.954	61.000	61.043
adenine–thymine WC complex	N	51.046	51.082	51.133	51.049
adenine–thymine WC complex	C	79.086	79.070	79.109	79.133
adenine–thymine WC complex	N	132.556	132.502	132.581	132.535
adenine–thymine WC complex	N	206.299	206.315	206.360	206.400
adenine–thymine WC complex	H	23.338	23.334	23.333	23.341
adenine–thymine WC complex	H	24.714	24.713	24.712	24.697
adenine–thymine WC complex	H	23.602	23.603	23.604	23.588
adenine–thymine WC complex	H	23.187	23.178	23.175	23.181
adenine–thymine WC complex	H	26.580	26.576	26.575	26.589
adenine–thymine WC complex	N	158.262	158.210	158.280	158.287
adenine–thymine WC complex	C	78.725	78.713	78.766	78.810
adenine–thymine WC complex	C	100.250	100.233	100.273	100.211
adenine–thymine WC complex	C	45.397	45.378	45.401	45.407
adenine–thymine WC complex	N	120.856	120.804	120.886	120.941
adenine–thymine WC complex	C	62.813	62.799	62.808	62.815
adenine–thymine WC complex	C	189.913	189.916	189.922	189.924
adenine–thymine WC complex	O	22.754	22.677	22.792	22.854
adenine–thymine WC complex	O	82.944	82.922	83.048	83.131
adenine–thymine WC complex	H	26.208	26.210	26.212	26.205
adenine–thymine WC complex	H	17.470	17.456	17.452	17.485
adenine–thymine WC complex	H	25.338	25.347	25.351	25.335
adenine–thymine WC complex	H	29.881	29.883	29.884	29.883
adenine–thymine WC complex	H	29.887	29.888	29.889	29.888
adenine–thymine WC complex	H	30.604	30.609	30.611	30.603
methane dimer	C	205.858	205.861	205.856	205.856
methane dimer	H	31.500	31.501	31.500	31.500
methane dimer	H	31.500	31.501	31.500	31.500
methane dimer	H	31.499	31.500	31.499	31.499
methane dimer	H	31.515	31.514	31.514	31.514
methane dimer	C	205.866	205.861	205.856	205.856
methane dimer	H	31.514	31.514	31.514	31.514
methane dimer	H	31.501	31.501	31.500	31.500
methane dimer	H	31.501	31.501	31.500	31.500
methane dimer	H	31.500	31.500	31.499	31.499
ethene dimer	C	90.520	90.517	90.560	90.560
ethene dimer	C	90.520	90.517	90.560	90.560
ethene dimer	H	26.269	26.268	26.269	26.270
ethene dimer	H	26.269	26.268	26.269	26.270
ethene dimer	H	26.482	26.482	26.484	26.483
ethene dimer	H	26.482	26.482	26.484	26.483
ethene dimer	C	90.520	90.517	90.560	90.560
ethene dimer	C	90.520	90.517	90.560	90.560
ethene dimer	H	26.269	26.268	26.269	26.270
ethene dimer	H	26.269	26.268	26.269	26.270
ethene dimer	H	26.482	26.482	26.484	26.483
ethene dimer	H	26.482	26.482	26.484	26.483
benzene stack	C	84.866	84.860	84.891	84.905
benzene stack	C	82.795	82.793	82.822	82.823
benzene stack	C	82.795	82.793	82.822	82.823
benzene stack	C	85.558	85.552	85.583	85.559
benzene stack	C	85.041	85.039	85.069	85.047
benzene stack	C	85.558	85.552	85.583	85.559
benzene stack	H	25.088	25.088	25.089	25.089
benzene stack	H	24.734	24.734	24.735	24.736
benzene stack	H	24.643	24.643	24.644	24.644
benzene stack	H	24.734	24.734	24.735	24.736
benzene stack	H	25.088	25.088	25.089	25.089
benzene stack	H	25.734	25.734	25.735	25.739
benzene stack	C	84.865	84.860	84.891	84.905
benzene stack	C	82.795	82.793	82.823	82.823
benzene stack	C	82.795	82.793	82.822	82.823
benzene stack	C	85.557	85.552	85.583	85.559
benzene stack	C	85.039	85.039	85.069	85.047

Continued on next page. . .

Molecule	Atom	Loose	Tight	Tight*	Ref.
benzene stack	C	85.557	85.552	85.583	85.559
benzene stack	H	25.735	25.734	25.735	25.739
benzene stack	H	25.088	25.088	25.089	25.089
benzene stack	H	24.734	24.734	24.735	24.736
benzene stack	H	24.643	24.643	24.644	24.644
benzene stack	H	24.734	24.734	24.735	24.736
benzene stack	H	25.088	25.088	25.089	25.089
pyrazine dimer	C	67.702	67.700	67.740	67.757
pyrazine dimer	C	67.703	67.700	67.740	67.757
pyrazine dimer	N	-29.730	-29.688	-29.637	-29.647
pyrazine dimer	C	66.678	66.680	66.721	66.722
pyrazine dimer	C	66.678	66.680	66.721	66.722
pyrazine dimer	N	-29.730	-29.688	-29.637	-29.647
pyrazine dimer	H	24.536	24.536	24.537	24.540
pyrazine dimer	H	24.536	24.536	24.537	24.540
pyrazine dimer	H	23.684	23.684	23.685	23.685
pyrazine dimer	H	23.684	23.684	23.685	23.685
pyrazine dimer	C	67.703	67.697	67.738	67.749
pyrazine dimer	C	65.706	65.708	65.748	65.758
pyrazine dimer	N	-26.983	-26.958	-26.905	-26.896
pyrazine dimer	C	65.706	65.708	65.748	65.758
pyrazine dimer	C	67.699	67.697	67.738	67.749
pyrazine dimer	N	-28.724	-28.668	-28.615	-28.629
pyrazine dimer	H	23.759	23.759	23.760	23.761
pyrazine dimer	H	24.281	24.281	24.282	24.284
pyrazine dimer	H	24.281	24.281	24.282	24.284
pyrazine dimer	H	23.759	23.759	23.760	23.761
uracil stack	N	155.157	155.124	155.191	155.159
uracil stack	C	75.409	75.409	75.458	75.485
uracil stack	H	25.513	25.515	25.517	25.516
uracil stack	C	108.019	107.998	108.039	108.056
uracil stack	H	27.069	27.071	27.072	27.070
uracil stack	C	52.119	52.104	52.125	52.098
uracil stack	O	-16.277	-16.398	-16.279	-16.179
uracil stack	N	125.310	125.264	125.358	125.329
uracil stack	H	25.397	25.389	25.388	25.385
uracil stack	C	63.574	63.558	63.567	63.556
uracil stack	O	74.539	74.525	74.655	74.710
uracil stack	H	26.221	26.220	26.221	26.218
uracil stack	N	155.157	155.124	155.191	155.159
uracil stack	C	75.409	75.409	75.458	75.485
uracil stack	H	25.513	25.515	25.517	25.516
uracil stack	C	108.019	107.998	108.039	108.056
uracil stack	H	27.069	27.071	27.072	27.070
uracil stack	C	52.119	52.104	52.125	52.098
uracil stack	O	-16.277	-16.398	-16.279	-16.179
uracil stack	N	125.310	125.264	125.358	125.329
uracil stack	H	25.397	25.389	25.388	25.385
uracil stack	C	63.574	63.558	63.567	63.556
uracil stack	O	74.539	74.525	74.655	74.710
uracil stack	H	26.221	26.220	26.221	26.218
indole–benzene stack	C	84.102	84.098	84.126	84.175
indole–benzene stack	C	86.526	86.518	86.548	86.554
indole–benzene stack	C	84.382	84.377	84.407	84.388
indole–benzene stack	C	86.806	86.800	86.831	86.790
indole–benzene stack	C	86.719	86.714	86.745	86.719
indole–benzene stack	C	87.804	87.797	87.827	87.838
indole–benzene stack	H	25.876	25.875	25.875	25.877
indole–benzene stack	H	25.738	25.736	25.737	25.735
indole–benzene stack	H	25.045	25.044	25.045	25.043
indole–benzene stack	H	24.876	24.875	24.876	24.875
indole–benzene stack	H	25.074	25.073	25.074	25.071
indole–benzene stack	H	25.458	25.457	25.458	25.459
indole–benzene stack	H	24.760	24.759	24.760	24.762
indole–benzene stack	C	91.248	91.227	91.254	91.222
indole–benzene stack	C	91.965	91.981	92.016	91.956
indole–benzene stack	H	24.219	24.218	24.218	24.219
indole–benzene stack	C	81.436	81.427	81.473	81.518
indole–benzene stack	C	108.265	108.262	108.304	108.246
indole–benzene stack	H	25.347	25.344	25.343	25.347
indole–benzene stack	C	91.851	91.858	91.906	91.891
indole–benzene stack	N	149.141	149.077	149.130	149.120
indole–benzene stack	C	78.267	78.281	78.328	78.344
indole–benzene stack	C	100.108	100.076	100.098	100.069
indole–benzene stack	H	25.384	25.383	25.384	25.384
indole–benzene stack	C	91.679	91.702	91.737	91.675
indole–benzene stack	H	25.018	25.020	25.021	25.022
indole–benzene stack	H	25.207	25.204	25.203	25.201
indole–benzene stack	H	25.152	25.151	25.151	25.152
adenine–thymine stack	N	129.349	129.300	129.382	129.372
adenine–thymine stack	C	76.816	76.799	76.836	76.834
adenine–thymine stack	H	24.540	24.535	24.534	24.530
adenine–thymine stack	N	49.033	49.056	49.110	49.117
adenine–thymine stack	C	85.573	85.520	85.566	85.530
adenine–thymine stack	C	54.545	54.555	54.591	54.557
adenine–thymine stack	N	209.663	209.676	209.725	209.751
adenine–thymine stack	H	27.672	27.671	27.670	27.672
adenine–thymine stack	H	26.871	26.864	26.863	26.868
adenine–thymine stack	N	38.818	38.693	38.719	38.712

Continued on next page...

Molecule	Atom	Loose	Tight	Tight*	Ref.
adenine–thymine stack	C	61.879	61.933	61.974	62.055
adenine–thymine stack	H	23.914	23.916	23.917	23.923
adenine–thymine stack	N	48.372	48.158	48.146	48.143
adenine–thymine stack	C	64.987	65.015	65.068	65.055
adenine–thymine stack	H	23.775	23.771	23.771	23.769
adenine–thymine stack	N	158.706	158.667	158.737	158.711
adenine–thymine stack	C	80.211	80.207	80.261	80.291
adenine–thymine stack	H	26.116	26.116	26.117	26.115
adenine–thymine stack	C	101.291	101.267	101.305	101.312
adenine–thymine stack	C	189.342	189.339	189.343	189.330
adenine–thymine stack	H	30.224	30.221	30.220	30.223
adenine–thymine stack	H	30.390	30.389	30.389	30.391
adenine–thymine stack	H	31.078	31.077	31.078	31.076
adenine–thymine stack	C	52.161	52.151	52.173	52.154
adenine–thymine stack	O	-5.732	-5.812	-5.692	-5.541
adenine–thymine stack	N	129.995	129.958	130.050	130.007
adenine–thymine stack	H	26.380	26.373	26.372	26.373
adenine–thymine stack	C	64.259	64.253	64.264	64.257
adenine–thymine stack	O	86.906	86.842	86.977	86.958
adenine–thymine stack	H	26.777	26.775	26.776	26.771
ethene–ethyne	C	89.433	89.431	89.474	89.473
ethene–ethyne	C	89.433	89.431	89.474	89.473
ethene–ethyne	H	26.468	26.467	26.469	26.469
ethene–ethyne	H	26.468	26.467	26.469	26.469
ethene–ethyne	H	26.468	26.467	26.469	26.469
ethene–ethyne	H	26.468	26.467	26.469	26.469
ethene–ethyne	C	141.330	141.327	141.337	141.341
ethene–ethyne	C	137.916	137.925	137.936	137.938
ethene–ethyne	H	29.948	29.948	29.946	29.946
ethene–ethyne	H	30.563	30.563	30.561	30.561
benzene–water	C	83.279	83.272	83.302	83.300
benzene–water	C	82.927	82.923	82.953	82.949
benzene–water	C	83.658	83.653	83.684	83.683
benzene–water	C	82.926	82.923	82.953	82.949
benzene–water	C	83.277	83.272	83.302	83.300
benzene–water	C	83.710	83.706	83.736	83.732
benzene–water	H	24.423	24.423	24.424	24.424
benzene–water	H	24.355	24.355	24.356	24.356
benzene–water	H	24.358	24.358	24.359	24.359
benzene–water	H	24.355	24.355	24.356	24.356
benzene–water	H	24.423	24.423	24.424	24.424
benzene–water	H	24.465	24.465	24.466	24.466
benzene–water	O	364.277	364.278	364.261	364.272
benzene–water	H	32.860	32.858	32.857	32.857
benzene–water	H	33.622	33.620	33.619	33.618
benzene–ammonia	C	83.336	83.330	83.360	83.359
benzene–ammonia	C	82.695	82.690	82.721	82.718
benzene–ammonia	C	83.335	83.330	83.360	83.359
benzene–ammonia	C	83.792	83.787	83.818	83.814
benzene–ammonia	C	83.868	83.863	83.893	83.888
benzene–ammonia	C	83.793	83.787	83.818	83.814
benzene–ammonia	H	24.412	24.412	24.413	24.413
benzene–ammonia	H	24.331	24.330	24.331	24.331
benzene–ammonia	H	24.413	24.412	24.413	24.413
benzene–ammonia	H	24.490	24.489	24.490	24.490
benzene–ammonia	H	24.509	24.509	24.510	24.510
benzene–ammonia	H	24.490	24.489	24.490	24.490
benzene–ammonia	N	287.339	287.345	287.328	287.321
benzene–ammonia	H	33.269	33.268	33.267	33.267
benzene–ammonia	H	33.269	33.268	33.267	33.267
benzene–ammonia	H	34.527	34.526	34.525	34.524
benzene–HCN	C	83.017	83.013	83.044	83.047
benzene–HCN	C	82.902	82.898	82.929	82.930
benzene–HCN	C	83.017	83.013	83.044	83.047
benzene–HCN	C	83.266	83.262	83.292	83.293
benzene–HCN	C	83.400	83.396	83.427	83.427
benzene–HCN	C	83.266	83.262	83.292	83.293
benzene–HCN	H	24.395	24.395	24.396	24.397
benzene–HCN	H	24.385	24.385	24.386	24.387
benzene–HCN	H	24.395	24.395	24.396	24.397
benzene–HCN	H	24.415	24.415	24.416	24.417
benzene–HCN	H	24.425	24.424	24.425	24.427
benzene–HCN	H	24.415	24.415	24.416	24.417
benzene–HCN	N	32.265	32.247	32.269	32.260
benzene–HCN	C	107.230	107.225	107.245	107.254
benzene–HCN	H	33.222	33.222	33.221	33.218
benzene dimer T-shaped	C	83.552	83.536	83.567	83.605
benzene dimer T-shaped	C	84.261	84.254	84.283	84.299
benzene dimer T-shaped	C	84.888	84.887	84.917	84.902
benzene dimer T-shaped	C	84.717	84.710	84.740	84.714
benzene dimer T-shaped	C	84.889	84.887	84.917	84.902
benzene dimer T-shaped	C	84.262	84.254	84.283	84.299
benzene dimer T-shaped	H	27.704	27.702	27.704	27.712
benzene dimer T-shaped	H	25.067	25.067	25.068	25.069
benzene dimer T-shaped	H	24.761	24.761	24.762	24.760
benzene dimer T-shaped	H	24.721	24.720	24.721	24.718
benzene dimer T-shaped	H	24.761	24.761	24.762	24.760
benzene dimer T-shaped	H	25.068	25.067	25.068	25.069
benzene dimer T-shaped	C	82.843	82.840	82.870	82.874

Continued on next page...

Molecule	Atom	Loose	Tight	Tight*	Ref.
benzene dimer T-shaped	C	83.364	83.361	83.392	83.387
benzene dimer T-shaped	C	83.363	83.361	83.392	83.387
benzene dimer T-shaped	C	82.835	82.840	82.870	82.874
benzene dimer T-shaped	C	83.362	83.361	83.392	83.387
benzene dimer T-shaped	C	83.364	83.361	83.392	83.387
benzene dimer T-shaped	H	24.336	24.336	24.336	24.337
benzene dimer T-shaped	H	24.255	24.255	24.256	24.257
benzene dimer T-shaped	H	24.255	24.255	24.256	24.257
benzene dimer T-shaped	H	24.335	24.336	24.336	24.337
benzene dimer T-shaped	H	24.255	24.255	24.256	24.257
benzene dimer T-shaped	H	24.255	24.255	24.256	24.257
indole–benzene T-shaped	C	82.457	82.455	82.486	82.502
indole–benzene T-shaped	C	82.525	82.521	82.552	82.555
indole–benzene T-shaped	C	82.478	82.472	82.503	82.497
indole–benzene T-shaped	C	83.464	83.458	83.489	83.489
indole–benzene T-shaped	C	82.475	82.472	82.503	82.497
indole–benzene T-shaped	C	82.525	82.521	82.552	82.555
indole–benzene T-shaped	H	23.875	23.874	23.875	23.875
indole–benzene T-shaped	H	24.129	24.127	24.128	24.130
indole–benzene T-shaped	H	24.187	24.186	24.187	24.188
indole–benzene T-shaped	H	24.076	24.075	24.076	24.075
indole–benzene T-shaped	H	24.187	24.186	24.187	24.188
indole–benzene T-shaped	H	24.129	24.127	24.128	24.130
indole–benzene T-shaped	H	27.986	27.979	27.979	27.991
indole–benzene T-shaped	N	149.334	149.253	149.304	149.323
indole–benzene T-shaped	C	91.346	91.356	91.405	91.373
indole–benzene T-shaped	C	109.523	109.522	109.564	109.514
indole–benzene T-shaped	C	81.850	81.842	81.890	81.919
indole–benzene T-shaped	C	77.477	77.492	77.539	77.589
indole–benzene T-shaped	C	100.888	100.850	100.871	100.878
indole–benzene T-shaped	C	91.948	91.972	92.008	91.985
indole–benzene T-shaped	C	91.691	91.663	91.689	91.655
indole–benzene T-shaped	C	91.969	91.989	92.024	91.968
indole–benzene T-shaped	H	25.574	25.572	25.572	25.574
indole–benzene T-shaped	H	25.577	25.575	25.575	25.575
indole–benzene T-shaped	H	24.882	24.879	24.879	24.882
indole–benzene T-shaped	H	24.896	24.897	24.899	24.898
indole–benzene T-shaped	H	24.846	24.845	24.846	24.848
indole–benzene T-shaped	H	24.301	24.301	24.302	24.302
phenol dimer	C	53.287	53.278	53.314	53.319
phenol dimer	O	260.431	260.412	260.489	260.499
phenol dimer	H	25.219	25.215	25.213	25.216
phenol dimer	C	96.214	96.184	96.217	96.240
phenol dimer	C	85.604	85.625	85.660	85.673
phenol dimer	C	91.914	91.889	91.914	91.897
phenol dimer	C	84.460	84.481	84.516	84.486
phenol dimer	C	93.967	93.940	93.972	93.972
phenol dimer	H	25.587	25.584	25.584	25.593
phenol dimer	H	25.179	25.179	25.179	25.182
phenol dimer	H	25.179	25.178	25.179	25.178
phenol dimer	H	24.704	24.706	24.707	24.706
phenol dimer	H	24.748	24.747	24.748	24.746
phenol dimer	O	266.156	266.093	266.169	266.205
phenol dimer	C	57.015	57.004	57.041	57.052
phenol dimer	H	27.838	27.836	27.835	27.835
phenol dimer	C	93.194	93.169	93.199	93.210
phenol dimer	C	82.780	82.798	82.833	82.841
phenol dimer	C	88.885	88.864	88.889	88.875
phenol dimer	C	84.516	84.532	84.567	84.549
phenol dimer	C	97.833	97.807	97.840	97.837
phenol dimer	H	24.701	24.698	24.698	24.706
phenol dimer	H	24.744	24.743	24.744	24.745
phenol dimer	H	24.875	24.874	24.874	24.873
phenol dimer	H	24.634	24.635	24.635	24.634
phenol dimer	H	25.130	25.129	25.129	25.127

Table 2: MP2 shielding calculations on the benchmark set from Flaig *et al.*[2] “Loose” and “Tight” settings are defined in the text. “Tight*” denotes shieldings computed with “Tight” settings and a cc-pVQZ-RI auxiliary basis set instead of a def2-SVP-RI basis set. MP2 shieldings computed with Turbomole[3–5] are used as a reference. All values are given in ppm.

Molecule	Atom	Loose	Tight	Tight*	Ref.
C_2H_2	H	30.597	30.597	30.595	30.595
C_2H_2	C	140.705	140.704	140.714	140.713
C_2H_2	C	140.705	140.704	140.714	140.713
C_2H_2	H	30.597	30.597	30.595	30.595
C_2H_4	H	26.631	26.630	26.631	26.632
C_2H_4	H	26.631	26.630	26.631	26.632
C_2H_4	C	91.113	91.110	91.153	91.153
C_2H_4	H	26.631	26.630	26.631	26.632
C_2H_4	C	91.111	91.109	91.152	91.153
C_2H_4	H	26.630	26.630	26.631	26.632
C_2H_6	H	31.094	31.094	31.093	31.093
C_2H_6	H	31.094	31.094	31.093	31.093
C_2H_6	H	31.094	31.094	31.093	31.093
C_2H_6	C	197.027	197.016	197.005	197.004
C_2H_6	H	31.093	31.093	31.093	31.093
C_2H_6	H	31.094	31.093	31.093	31.093
C_2H_6	C	197.023	197.014	197.003	197.003
C_2H_6	H	31.094	31.093	31.093	31.093
C_6H_6	H	24.236	24.235	24.236	24.236
C_6H_6	H	24.236	24.235	24.236	24.236
C_6H_6	C	81.848	81.843	81.874	81.872
C_6H_6	H	24.236	24.235	24.236	24.236
C_6H_6	C	81.848	81.843	81.874	81.872
C_6H_6	H	24.236	24.235	24.236	24.236
C_6H_6	C	81.848	81.843	81.874	81.872
C_6H_6	H	24.236	24.235	24.236	24.236
C_6H_6	C	81.850	81.843	81.874	81.872
C_6H_6	C	81.850	81.843	81.874	81.872
C_6H_6	C	81.861	81.843	81.874	81.872
C_6H_6	H	24.235	24.235	24.236	24.236
CCl_4	C	80.240	80.163	80.214	80.170
CCl_4	Cl	542.921	542.731	543.424	543.335
CCl_4	Cl	542.924	542.731	543.424	543.335
CCl_4	Cl	542.919	542.731	543.424	543.335
CCl_4	Cl	542.921	542.731	543.424	543.335
CF_4	F	300.550	300.555	300.684	300.687
CF_4	F	300.570	300.555	300.684	300.687
CF_4	F	300.549	300.555	300.684	300.687
CF_4	C	81.845	81.847	81.860	81.860
CF_4	F	300.548	300.555	300.684	300.687
CH_2CCH_2	H	27.398	27.398	27.398	27.398
CH_2CCH_2	H	27.397	27.397	27.398	27.398
CH_2CCH_2	C	135.781	135.778	135.794	135.795
CH_2CCH_2	C	0.135	0.150	0.205	0.206
CH_2CCH_2	H	27.398	27.397	27.398	27.398
CH_2CCH_2	C	135.778	135.775	135.791	135.793
CH_2CCH_2	H	27.397	27.397	27.398	27.398
CH_3CHO	O	-247.028	-246.908	-246.548	-246.528
CH_3CHO	H	22.885	22.884	22.886	22.886
CH_3CHO	C	28.440	28.447	28.522	28.520
CH_3CHO	H	30.023	30.023	30.023	30.023
CH_3CHO	H	30.023	30.023	30.023	30.023
CH_3CHO	C	175.874	175.873	175.896	175.899
CH_3CHO	H	30.304	30.304	30.303	30.303
CH_3Cl	C	179.327	179.314	179.321	179.320
CH_3Cl	Cl	1010.339	1010.283	1010.228	1010.236
CH_3Cl	H	29.337	29.338	29.339	29.339
CH_3Cl	H	29.337	29.338	29.339	29.339
CH_3Cl	H	29.337	29.338	29.339	29.339
CH_3CN	N	42.119	42.119	42.151	42.150
CH_3CN	C	99.424	99.405	99.418	99.416
CH_3CN	H	30.439	30.439	30.438	30.438
CH_3CN	H	30.436	30.436	30.436	30.436
CH_3CN	C	204.131	204.127	204.117	204.120
CH_3CN	H	30.437	30.437	30.437	30.437
CH_3COCH_3	H	30.393	30.393	30.393	30.392
CH_3COCH_3	H	29.964	29.964	29.964	29.964
CH_3COCH_3	H	29.964	29.964	29.964	29.964
CH_3COCH_3	C	176.847	176.849	176.872	176.872
CH_3COCH_3	O	-241.915	-241.875	-241.474	-241.466
CH_3COCH_3	C	20.053	20.057	20.144	20.141
CH_3COCH_3	H	29.964	29.964	29.964	29.964
CH_3COCH_3	H	29.964	29.964	29.964	29.964
CH_3COCH_3	C	176.847	176.850	176.872	176.873
CH_3COCH_3	H	30.393	30.393	30.393	30.393
CH_3F	F	488.310	488.316	488.335	488.339
CH_3F	H	27.914	27.914	27.914	27.914
CH_3F	H	27.914	27.914	27.914	27.914
CH_3F	C	137.177	137.189	137.206	137.204
CH_3F	H	27.914	27.914	27.914	27.914
CH_3NH_2	H	32.123	32.123	32.123	32.123
CH_3NH_2	H	32.124	32.123	32.123	32.123

Continued on next page...

Molecule	Atom	Loose	Tight	Tight*	Ref.
CH_3NH_2	N	272.114	272.102	272.087	272.085
CH_3NH_2	H	29.696	29.696	29.695	29.695
CH_3NH_2	H	29.718	29.718	29.717	29.717
CH_3NH_2	C	175.379	175.376	175.373	175.373
CH_3NH_2	H	29.718	29.718	29.717	29.717
CH_3OCH_3	H	28.417	28.417	28.417	28.417
CH_3OCH_3	H	29.111	29.111	29.111	29.111
CH_3OCH_3	H	29.111	29.111	29.111	29.111
CH_3OCH_3	C	145.941	145.933	145.946	145.949
CH_3OCH_3	O	367.591	367.566	367.578	367.583
CH_3OCH_3	H	29.111	29.111	29.111	29.111
CH_3OCH_3	H	29.111	29.111	29.111	29.111
CH_3OCH_3	C	145.945	145.933	145.946	145.949
CH_3OCH_3	H	28.418	28.417	28.417	28.417
CH_3OH	H	32.553	32.554	32.552	32.552
CH_3OH	O	363.971	363.957	363.952	363.949
CH_3OH	H	28.705	28.705	28.705	28.705
CH_3OH	H	28.705	28.705	28.705	28.705
CH_3OH	C	155.117	155.105	155.113	155.113
CH_3OH	H	28.590	28.590	28.590	28.590
CH_3PH_2	C	207.616	207.611	207.602	207.601
CH_3PH_2	P	599.208	599.252	599.353	599.357
CH_3PH_2	H	29.258	29.259	29.258	29.258
CH_3PH_2	H	29.258	29.259	29.258	29.258
CH_3PH_2	H	31.447	31.449	31.449	31.449
CH_3PH_2	H	31.447	31.449	31.449	31.449
CH_3PH_2	H	30.621	30.622	30.623	30.623
CH_3SH	C	196.239	196.225	196.216	196.217
CH_3SH	S	748.256	748.374	748.507	748.509
CH_3SH	H	31.123	31.121	31.120	31.120
CH_3SH	H	30.644	30.647	30.647	30.647
CH_3SH	H	30.151	30.152	30.153	30.153
CH_3SH	H	30.151	30.152	30.153	30.153
CH_4	H	31.583	31.583	31.582	31.582
CH_4	H	31.583	31.583	31.582	31.582
CH_4	H	31.583	31.583	31.582	31.582
CH_4	C	207.146	207.139	207.135	207.135
CH_4	H	31.583	31.583	31.582	31.582
CO_2	O	264.065	264.056	264.058	264.061
CO_2	C	87.547	87.541	87.535	87.534
CO_2	O	264.063	264.056	264.058	264.061
CO	O	-3.069	-3.073	-3.046	-3.039
CO	C	38.931	38.937	38.960	38.961
furan	H	25.858	25.857	25.857	25.857
furan	H	25.086	25.085	25.085	25.085
furan	C	105.058	105.054	105.089	105.088
furan	H	25.858	25.857	25.857	25.857
furan	C	75.879	75.873	75.916	75.919
furan	C	105.058	105.054	105.089	105.089
furan	O	98.706	98.617	98.786	98.789
furan	C	75.879	75.874	75.916	75.918
furan	H	25.086	25.085	25.085	25.085
glycine	H	30.709	30.709	30.709	30.709
glycine	H	31.192	31.192	31.191	31.191
glycine	N	262.070	262.061	262.064	262.069
glycine	H	29.178	29.178	29.177	29.177
glycine	H	29.275	29.275	29.274	29.275
glycine	C	163.456	163.450	163.460	163.462
glycine	O	2.625	2.612	2.733	2.759
glycine	C	48.754	48.744	48.771	48.769
glycine	O	192.706	192.689	192.808	192.804
glycine	H	21.774	21.773	21.774	21.774
H_2CO	O	-293.516	-293.301	-292.894	-292.855
H_2CO	H	23.096	23.094	23.096	23.096
H_2CO	C	34.365	34.371	34.443	34.442
H_2CO	H	23.096	23.094	23.096	23.096
HCN	H	29.687	29.687	29.685	29.685
HCN	C	109.975	109.971	109.990	109.990
HCN	N	31.431	31.422	31.444	31.447
$HCONH_2$	H	27.971	27.971	27.971	27.971
$HCONH_2$	H	27.994	27.995	27.996	27.996
$HCONH_2$	N	189.562	189.562	189.624	189.620
$HCONH_2$	H	24.559	24.559	24.559	24.559
$HCONH_2$	C	61.449	61.449	61.473	61.473
$HCONH_2$	O	-8.028	-8.029	-7.901	-7.868
HCOOH	H	26.643	26.643	26.644	26.644
HCOOH	O	184.003	183.992	184.100	184.102
HCOOH	H	24.892	24.892	24.891	24.891
HCOOH	C	59.210	59.205	59.227	59.225
HCOOH	O	-14.651	-14.639	-14.530	-14.507
imidazole	H	23.863	23.861	23.862	23.861
imidazole	H	25.069	25.069	25.068	25.068
imidazole	N	128.828	128.799	128.873	128.877
imidazole	H	25.304	25.303	25.303	25.303
imidazole	C	83.090	83.100	83.139	83.139
imidazole	C	98.344	98.332	98.368	98.373
imidazole	N	25.424	25.388	25.430	25.428
imidazole	C	83.241	83.237	83.283	83.285
imidazole	H	25.016	25.015	25.015	25.016

Continued on next page...

Molecule	Atom	Loose	Tight	Tight*	Ref.
pyridine	H	23.449	23.448	23.449	23.449
pyridine	N	-31.689	-31.783	-31.770	-31.773
pyridine	H	23.449	23.448	23.449	23.449
pyridine	C	64.908	64.931	64.976	64.977
pyridine	H	24.719	24.718	24.718	24.718
pyridine	C	64.908	64.931	64.976	64.977
pyridine	H	24.719	24.718	24.718	24.718
pyridine	C	86.844	86.815	86.842	86.844
pyridine	C	86.844	86.815	86.842	86.844
pyridine	C	80.863	80.890	80.932	80.932
pyridine	H	24.446	24.447	24.448	24.449
pyrimidine	H	23.487	23.486	23.487	23.487
pyrimidine	N	-16.845	-17.002	-17.009	-17.004
pyrimidine	H	22.865	22.864	22.864	22.864
pyrimidine	C	60.606	60.633	60.683	60.685
pyrimidine	H	24.804	24.803	24.803	24.804
pyrimidine	C	54.877	54.898	54.933	54.936
pyrimidine	C	87.218	87.181	87.213	87.218
pyrimidine	N	-16.850	-17.008	-17.015	-17.009
pyrimidine	C	60.592	60.620	60.670	60.673
pyrimidine	H	23.481	23.481	23.481	23.482
TMS	H	31.905	31.906	31.906	31.905
TMS	H	31.905	31.906	31.906	31.905
TMS	H	31.905	31.906	31.906	31.905
TMS	H	31.905	31.906	31.906	31.905
TMS	H	31.905	31.906	31.906	31.905
TMS	H	31.905	31.906	31.906	31.906
TMS	H	31.905	31.906	31.906	31.905
TMS	H	31.905	31.906	31.906	31.905
TMS	H	31.905	31.906	31.906	31.905
TMS	C	205.816	205.809	205.815	205.815
TMS	C	205.816	205.809	205.815	205.817
TMS	C	205.816	205.809	205.815	205.817
TMS	Si	439.889	439.859	439.981	439.831
TMS	H	31.905	31.906	31.906	31.905
TMS	H	31.905	31.906	31.906	31.906
TMS	C	205.816	205.809	205.815	205.817
TMS	H	31.905	31.906	31.906	31.905

Table 3: SOS-MP2 shielding calculations for all monomers from the L7 test set.[6] SOS-MP2 shieldings computed with FermiONs++[7–9] and “Tight” settings are used as a reference. All values are given in ppm.

Molecule	Atom	Loose	Tight
C2C2PD monomer A	C	97.161	97.140
C2C2PD monomer A	C	97.149	97.136
C2C2PD monomer A	C	97.144	97.134
C2C2PD monomer A	C	97.149	97.136
C2C2PD monomer A	C	97.160	97.140
C2C2PD monomer A	C	97.168	97.142
C2C2PD monomer A	C	93.503	93.512
C2C2PD monomer A	C	100.579	100.589
C2C2PD monomer A	C	100.578	100.588
C2C2PD monomer A	C	93.513	93.515
C2C2PD monomer A	C	93.519	93.518
C2C2PD monomer A	C	100.583	100.589
C2C2PD monomer A	C	100.584	100.590
C2C2PD monomer A	C	100.583	100.589
C2C2PD monomer A	C	100.583	100.590
C2C2PD monomer A	C	93.511	93.515
C2C2PD monomer A	C	100.578	100.588
C2C2PD monomer A	C	100.580	100.589
C2C2PD monomer A	C	93.503	93.512
C2C2PD monomer A	C	100.574	100.587
C2C2PD monomer A	C	100.575	100.587
C2C2PD monomer A	C	93.498	93.510
C2C2PD monomer A	C	100.579	100.587
C2C2PD monomer A	C	100.578	100.587
C2C2PD monomer A	H	23.539	23.543
C2C2PD monomer A	H	23.536	23.542
C2C2PD monomer A	H	23.535	23.541
C2C2PD monomer A	H	23.533	23.540
C2C2PD monomer A	H	23.533	23.540
C2C2PD monomer A	H	23.535	23.541
C2C2PD monomer A	H	23.537	23.542
C2C2PD monomer A	H	23.539	23.543
C2C2PD monomer A	H	23.541	23.544
C2C2PD monomer A	H	23.542	23.544
C2C2PD monomer A	H	23.542	23.544
C2C2PD monomer A	H	23.541	23.544
C2C2PD monomer B	C	97.148	97.136
C2C2PD monomer B	C	97.162	97.140
C2C2PD monomer B	C	97.167	97.142
C2C2PD monomer B	C	97.162	97.140
C2C2PD monomer B	C	97.147	97.136
C2C2PD monomer B	C	97.143	97.134
C2C2PD monomer B	C	93.513	93.515
C2C2PD monomer B	C	100.577	100.588
C2C2PD monomer B	C	100.580	100.589
C2C2PD monomer B	C	93.502	93.512
C2C2PD monomer B	C	93.496	93.510
C2C2PD monomer B	C	100.576	100.587
C2C2PD monomer B	C	100.574	100.587
C2C2PD monomer B	C	100.575	100.587
C2C2PD monomer B	C	100.574	100.587
C2C2PD monomer B	C	93.500	93.512
C2C2PD monomer B	C	100.580	100.589
C2C2PD monomer B	C	100.579	100.588
C2C2PD monomer B	C	93.514	93.515
C2C2PD monomer B	C	100.583	100.590
C2C2PD monomer B	C	100.582	100.589
C2C2PD monomer B	C	93.521	93.518
C2C2PD monomer B	C	100.587	100.589
C2C2PD monomer B	C	100.588	100.590
C2C2PD monomer B	H	23.537	23.542
C2C2PD monomer B	H	23.539	23.543
C2C2PD monomer B	H	23.541	23.544
C2C2PD monomer B	H	23.542	23.544
C2C2PD monomer B	H	23.542	23.544
C2C2PD monomer B	H	23.541	23.544
C2C2PD monomer B	H	23.539	23.543
C2C2PD monomer B	H	23.537	23.542
C2C2PD monomer B	H	23.534	23.541
C2C2PD monomer B	H	23.533	23.540
C2C2PD monomer B	H	23.533	23.540
C2C2PD monomer B	H	23.534	23.541
C3A monomer A	C	71.512	71.536
C3A monomer A	C	96.044	95.989
C3A monomer A	C	81.091	81.134
C3A monomer A	C	81.878	81.945
C3A monomer A	C	100.456	100.439
C3A monomer A	N	65.336	65.189
C3A monomer A	N	73.180	72.927
C3A monomer A	N	102.467	102.497
C3A monomer A	N	151.525	151.467
C3A monomer A	N	221.125	221.135
C3A monomer A	H	24.115	24.118

Continued on next page...

Molecule	Atom	Loose	Tight
C3A monomer A	H	25.339	25.333
C3A monomer A	H	24.208	24.203
C3A monomer A	H	27.883	27.882
C3A monomer A	H	27.281	27.274
C3GC monomer A	C	71.011	71.006
C3GC monomer A	C	100.151	100.182
C3GC monomer A	C	125.582	125.551
C3GC monomer A	C	63.275	63.257
C3GC monomer A	C	65.937	65.916
C3GC monomer A	C	73.261	73.305
C3GC monomer A	C	79.198	79.274
C3GC monomer A	C	98.707	98.636
C3GC monomer A	C	103.860	103.869
C3GC monomer A	N	161.320	161.318
C3GC monomer A	N	103.802	103.629
C3GC monomer A	N	186.877	186.851
C3GC monomer A	N	145.490	145.411
C3GC monomer A	N	134.123	133.997
C3GC monomer A	N	103.400	103.455
C3GC monomer A	N	152.045	151.968
C3GC monomer A	N	219.064	219.101
C3GC monomer A	O	92.328	92.190
C3GC monomer A	O	91.416	91.147
C3GC monomer A	H	26.070	26.076
C3GC monomer A	H	26.097	26.112
C3GC monomer A	H	27.255	27.268
C3GC monomer A	H	27.706	27.707
C3GC monomer A	H	19.841	19.821
C3GC monomer A	H	18.268	18.238
C3GC monomer A	H	25.837	25.837
C3GC monomer A	H	24.704	24.704
C3GC monomer A	H	28.491	28.489
C3GC monomer A	H	23.256	23.242
CBH monomer A	C	189.651	189.643
CBH monomer A	C	180.241	180.233
CBH monomer A	C	171.967	171.961
CBH monomer A	C	172.760	172.755
CBH monomer A	C	172.526	172.521
CBH monomer A	C	172.565	172.560
CBH monomer A	C	172.458	172.453
CBH monomer A	C	172.624	172.618
CBH monomer A	C	172.405	172.400
CBH monomer A	C	172.660	172.654
CBH monomer A	C	172.438	172.437
CBH monomer A	C	172.581	172.575
CBH monomer A	C	172.493	172.488
CBH monomer A	C	172.640	172.635
CBH monomer A	C	172.675	172.671
CBH monomer A	C	171.991	171.989
CBH monomer A	C	180.218	180.212
CBH monomer A	C	189.588	189.580
CBH monomer A	H	30.638	30.637
CBH monomer A	H	30.854	30.854
CBH monomer A	H	30.652	30.652
CBH monomer A	H	30.873	30.873
CBH monomer A	H	30.540	30.540
CBH monomer A	H	30.531	30.531
CBH monomer A	H	30.570	30.570
CBH monomer A	H	30.572	30.572
CBH monomer A	H	30.559	30.558
CBH monomer A	H	30.558	30.558
CBH monomer A	H	30.518	30.518
CBH monomer A	H	30.552	30.552
CBH monomer A	H	30.543	30.542
CBH monomer A	H	30.529	30.529
CBH monomer A	H	30.533	30.533
CBH monomer A	H	30.558	30.558
CBH monomer A	H	30.554	30.554
CBH monomer A	H	30.528	30.528
CBH monomer A	H	30.537	30.537
CBH monomer A	H	30.560	30.560
CBH monomer A	H	30.560	30.560
CBH monomer A	H	30.538	30.538
CBH monomer A	H	30.550	30.550
CBH monomer A	H	30.532	30.532
CBH monomer A	H	30.524	30.524
CBH monomer A	H	30.542	30.541
CBH monomer A	H	30.540	30.540
CBH monomer A	H	30.538	30.538
CBH monomer A	H	30.539	30.538
CBH monomer A	H	30.538	30.538
CBH monomer A	H	30.548	30.547
CBH monomer A	H	30.555	30.555
CBH monomer A	H	30.561	30.561
CBH monomer A	H	30.579	30.578
CBH monomer A	H	30.535	30.534
CBH monomer A	H	30.541	30.540
CBH monomer A	H	30.867	30.867
CBH monomer A	H	30.849	30.848

Continued on next page...

Molecule	Atom	Loose	Tight
CBH monomer B	C	172.732	172.728
CBH monomer B	C	171.950	171.946
CBH monomer B	C	180.231	180.224
CBH monomer B	C	189.622	189.616
CBH monomer B	C	172.499	172.493
CBH monomer B	C	172.642	172.637
CBH monomer B	C	172.446	172.441
CBH monomer B	C	172.623	172.617
CBH monomer B	C	172.451	172.446
CBH monomer B	C	172.627	172.621
CBH monomer B	C	172.434	172.428
CBH monomer B	C	172.651	172.645
CBH monomer B	C	172.447	172.441
CBH monomer B	C	172.625	172.619
CBH monomer B	C	172.683	172.677
CBH monomer B	C	171.964	171.958
CBH monomer B	C	180.226	180.217
CBH monomer B	C	189.562	189.554
CBH monomer B	H	30.639	30.639
CBH monomer B	H	30.847	30.847
CBH monomer B	H	30.534	30.533
CBH monomer B	H	30.541	30.541
CBH monomer B	H	30.569	30.569
CBH monomer B	H	30.561	30.561
CBH monomer B	H	30.566	30.566
CBH monomer B	H	30.542	30.541
CBH monomer B	H	30.541	30.541
CBH monomer B	H	30.551	30.551
CBH monomer B	H	30.542	30.542
CBH monomer B	H	30.546	30.545
CBH monomer B	H	30.542	30.542
CBH monomer B	H	30.559	30.559
CBH monomer B	H	30.560	30.559
CBH monomer B	H	30.555	30.555
CBH monomer B	H	30.555	30.554
CBH monomer B	H	30.547	30.547
CBH monomer B	H	30.537	30.537
CBH monomer B	H	30.559	30.559
CBH monomer B	H	30.568	30.567
CBH monomer B	H	30.528	30.528
CBH monomer B	H	30.510	30.510
CBH monomer B	H	30.553	30.553
CBH monomer B	H	30.571	30.571
CBH monomer B	H	30.536	30.536
CBH monomer B	H	30.526	30.526
CBH monomer B	H	30.552	30.552
CBH monomer B	H	30.556	30.556
CBH monomer B	H	30.556	30.556
CBH monomer B	H	30.596	30.596
CBH monomer B	H	30.541	30.541
CBH monomer B	H	30.531	30.530
CBH monomer B	H	30.653	30.652
CBH monomer B	H	30.835	30.835
CBH monomer B	H	30.888	30.887
CBH monomer B	H	30.533	30.533
CBH monomer B	H	30.846	30.846
GCGC monomer A	C	74.407	74.404
GCGC monomer A	C	101.629	101.648
GCGC monomer A	C	127.796	127.769
GCGC monomer A	C	67.233	67.210
GCGC monomer A	N	168.454	168.451
GCGC monomer A	N	105.994	105.841
GCGC monomer A	N	200.115	200.111
GCGC monomer A	O	95.993	95.928
GCGC monomer A	H	26.724	26.730
GCGC monomer A	H	26.361	26.374
GCGC monomer A	H	27.574	27.583
GCGC monomer A	H	28.514	28.514
GCGC monomer A	H	23.684	23.667
GCGC monomer A	C	70.658	70.650
GCGC monomer A	C	77.345	77.368
GCGC monomer A	C	82.154	82.204
GCGC monomer A	C	101.690	101.624
GCGC monomer A	C	104.908	104.910
GCGC monomer A	N	154.871	154.815
GCGC monomer A	N	144.994	144.894
GCGC monomer A	N	102.771	102.802
GCGC monomer A	N	155.840	155.782
GCGC monomer A	N	226.887	226.919
GCGC monomer A	O	94.203	94.024
GCGC monomer A	H	21.670	21.646
GCGC monomer A	H	26.082	26.082
GCGC monomer A	H	25.323	25.325
GCGC monomer A	H	29.192	29.192
GCGC monomer A	H	25.811	25.800
GCGC monomer B	C	70.657	70.649
GCGC monomer B	C	77.344	77.368
GCGC monomer B	C	82.153	82.204
GCGC monomer B	C	101.688	101.623

Continued on next page...

Molecule	Atom	Loose	Tight
GCGC monomer B	C	104.904	104.909
GCGC monomer B	N	154.870	154.817
GCGC monomer B	N	144.991	144.891
GCGC monomer B	N	102.762	102.794
GCGC monomer B	N	155.841	155.785
GCGC monomer B	N	226.883	226.921
GCGC monomer B	O	94.188	94.024
GCGC monomer B	H	21.670	21.646
GCGC monomer B	H	26.082	26.082
GCGC monomer B	H	25.324	25.325
GCGC monomer B	H	29.192	29.192
GCGC monomer B	H	25.811	25.800
GCGC monomer B	C	74.410	74.402
GCGC monomer B	C	101.629	101.648
GCGC monomer B	C	127.794	127.769
GCGC monomer B	C	67.232	67.209
GCGC monomer B	N	168.454	168.453
GCGC monomer B	N	105.990	105.839
GCGC monomer B	N	200.115	200.114
GCGC monomer B	O	96.019	95.916
GCGC monomer B	H	26.725	26.730
GCGC monomer B	H	26.361	26.374
GCGC monomer B	H	27.575	27.583
GCGC monomer B	H	28.514	28.514
GCGC monomer B	H	23.684	23.667
GGG monomer A	N	152.748	152.651
GGG monomer A	C	102.999	103.021
GGG monomer A	N	98.479	98.583
GGG monomer A	C	100.410	100.250
GGG monomer A	C	77.107	77.102
GGG monomer A	O	45.008	44.665
GGG monomer A	N	151.533	151.539
GGG monomer A	C	77.687	77.759
GGG monomer A	N	230.502	230.546
GGG monomer A	N	134.167	134.070
GGG monomer A	C	79.218	79.302
GGG monomer A	H	24.664	24.628
GGG monomer A	H	25.564	25.643
GGG monomer A	H	25.823	25.789
GGG monomer A	H	28.373	28.380
GGG monomer A	H	28.366	28.406
GGG monomer B	N	157.707	157.635
GGG monomer B	C	106.278	106.236
GGG monomer B	N	95.659	95.703
GGG monomer B	C	100.332	100.121
GGG monomer B	C	77.356	77.338
GGG monomer B	O	20.036	19.686
GGG monomer B	N	154.711	154.530
GGG monomer B	C	83.488	83.473
GGG monomer B	N	231.748	231.555
GGG monomer B	N	145.620	145.406
GGG monomer B	C	83.631	83.708
GGG monomer B	H	24.968	24.931
GGG monomer B	H	26.892	26.837
GGG monomer B	H	25.956	25.897
GGG monomer B	H	29.331	29.277
GGG monomer B	H	28.907	28.891
GGG monomer B	N	154.315	154.348
GGG monomer B	C	105.500	105.585
GGG monomer B	N	94.029	94.299
GGG monomer B	C	99.849	99.691
GGG monomer B	C	78.187	78.080
GGG monomer B	O	-10.178	-10.639
GGG monomer B	N	154.428	154.322
GGG monomer B	C	81.293	81.506
GGG monomer B	N	229.881	229.936
GGG monomer B	N	142.980	142.752
GGG monomer B	C	87.367	87.522
GGG monomer B	H	24.967	25.079
GGG monomer B	H	25.227	25.301
GGG monomer B	H	26.322	26.383
GGG monomer B	H	28.286	28.328
GGG monomer B	H	28.522	28.611
PHE monomer A	C	98.522	98.467
PHE monomer A	C	91.250	91.198
PHE monomer A	C	97.814	97.763
PHE monomer A	C	100.898	100.898
PHE monomer A	C	100.190	100.158
PHE monomer A	C	100.280	100.265
PHE monomer A	C	161.384	161.391
PHE monomer A	C	150.683	150.670
PHE monomer A	C	65.299	65.293
PHE monomer A	C	74.366	74.341
PHE monomer A	C	184.592	184.606
PHE monomer A	N	201.885	201.863
PHE monomer A	N	176.713	176.662
PHE monomer A	O	32.616	32.519
PHE monomer A	O	48.173	48.099
PHE monomer A	H	30.100	30.109

Continued on next page...

Molecule	Atom	Loose	Tight
PHE monomer A	H	28.011	28.006
PHE monomer A	H	30.355	30.358
PHE monomer A	H	30.746	30.753
PHE monomer A	H	26.871	26.881
PHE monomer A	H	28.313	28.294
PHE monomer A	H	29.427	29.438
PHE monomer A	H	29.315	29.310
PHE monomer A	H	25.088	25.097
PHE monomer A	H	25.131	25.120
PHE monomer A	H	25.120	25.126
PHE monomer A	H	25.080	25.074
PHE monomer A	H	25.109	25.108
PHE monomer A	H	27.415	27.406
PHE monomer B	C	66.678	66.665
PHE monomer B	C	151.006	150.979
PHE monomer B	C	68.379	68.350
PHE monomer B	C	162.115	162.120
PHE monomer B	C	91.519	91.508
PHE monomer B	C	97.734	97.695
PHE monomer B	C	101.836	101.858
PHE monomer B	C	99.760	99.724
PHE monomer B	C	99.640	99.641
PHE monomer B	C	97.171	97.127
PHE monomer B	C	182.950	182.962
PHE monomer B	C	59.282	59.253
PHE monomer B	C	149.623	149.596
PHE monomer B	C	68.153	68.136
PHE monomer B	C	167.619	167.612
PHE monomer B	C	90.029	90.012
PHE monomer B	C	99.000	98.975
PHE monomer B	C	100.119	100.103
PHE monomer B	C	101.012	100.995
PHE monomer B	C	100.943	100.945
PHE monomer B	C	95.898	95.878
PHE monomer B	C	185.649	185.656
PHE monomer B	N	161.011	160.967
PHE monomer B	N	193.938	193.911
PHE monomer B	N	188.311	188.276
PHE monomer B	N	164.465	164.419
PHE monomer B	O	49.231	49.128
PHE monomer B	O	64.165	64.069
PHE monomer B	O	68.747	68.609
PHE monomer B	O	45.454	45.348
PHE monomer B	H	30.329	30.332
PHE monomer B	H	27.808	27.799
PHE monomer B	H	30.516	30.517
PHE monomer B	H	28.093	28.086
PHE monomer B	H	30.131	30.139
PHE monomer B	H	29.891	29.896
PHE monomer B	H	30.582	30.586
PHE monomer B	H	23.645	23.646
PHE monomer B	H	28.028	28.019
PHE monomer B	H	29.160	29.161
PHE monomer B	H	29.222	29.220
PHE monomer B	H	26.042	26.048
PHE monomer B	H	25.060	25.054
PHE monomer B	H	25.404	25.407
PHE monomer B	H	25.052	25.048
PHE monomer B	H	25.166	25.165
PHE monomer B	H	24.688	24.678
PHE monomer B	H	30.596	30.604
PHE monomer B	H	26.547	26.558
PHE monomer B	H	27.718	27.705
PHE monomer B	H	29.389	29.398
PHE monomer B	H	28.966	28.964
PHE monomer B	H	25.257	25.264
PHE monomer B	H	24.918	24.911
PHE monomer B	H	25.172	25.178
PHE monomer B	H	25.919	25.914
PHE monomer B	H	25.442	25.441
PHE monomer B	H	24.557	24.549

References

[1] P. Jurečka, J. Šponer, J. Černỳ, and P. Hobza, Phys. Chem. Chem. Phys. **8**, 1985 (2006).

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

[3] "TURBOMOLE V7.4 2019, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from `http://www.turbomole.com`." .

[4] M. Kollwitz and J. Gauss, Chem. Phys. Lett. **260**, 639 (1996).

[5] M. Kollwitz, M. Häser, and J. Gauss, J. Chem. Phys. **108**, 8295 (1998).

[6] R. Sedlak, T. Janowski, M. Pitonak, J. Rezac, P. Pulay, and P. Hobza, J. Chem. Theory Comput. **9**, 3364 (2013).

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

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

[9] J. Kussmann and C. Ochsenfeld, J. Chem. Theory Comput. **13**, 3153 (2017).

3.4 Benchmarking the Accuracy of the Direct Random Phase Approximation and σ-Functionals for NMR Shieldings

M. Glasbrenner, D. Graf, C. Ochsenfeld
J. Chem. Theory Comput. **18**, 192 (2022).

Abstract: A method for computing NMR shieldings with the direct random phase approximation (RPA) and the closely related σ-functionals [Trushin, E.; Thierbach, A.; Görling, A. Toward chemical accuracy at low computational cost: density functional theory with σ-functionals for the correlation energy. *J. Chem. Phys.* **2021**, *154*, 014104] is presented, which is based on a finite-difference approach. The accuracy is evaluated in benchmark calculations using high-quality coupled cluster values as a reference. Our results show that the accuracy of the computed NMR shieldings using direct RPA is strongly dependent on the density functional theory reference orbitals and improves with increasing amounts of exact Hartree–Fock exchange in the functional. NMR shieldings computed with direct RPA using a Hartree-Fock reference are significantly more accurate than MP2 shieldings and comparable to CCSD shieldings. Also, the basis set convergence is analyzed and it is shown that at least triple-zeta basis sets are required for reliable results.

Reprinted with permission from:

M. Glasbrenner, D. Graf, C. Ochsenfeld.
"Benchmarking the Accuracy of the Direct Random Phase Approximation and
σ-Functionals for NMR Shieldings."
J. Chem. Theory Comput. **18**, 192 (2022).

`https://pubs.acs.org/doi/abs/10.1021/acs.jctc.1c00866`

JCTC Journal of Chemical Theory and Computation

pubs.acs.org/JCTC Article

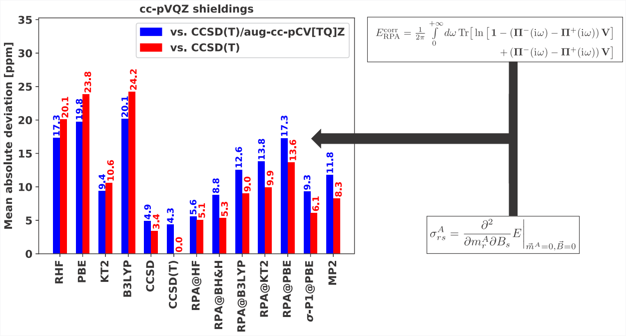
Benchmarking the Accuracy of the Direct Random Phase Approximation and σ-Functionals for NMR Shieldings

Michael Glasbrenner, Daniel Graf, and Christian Ochsenfeld*

Cite This: J. Chem. Theory Comput. 2022, 18, 192−205 Read Online

ACCESS | Metrics & More | Article Recommendations | SI Supporting Information

ABSTRACT: A method for computing NMR shieldings with the direct random phase approximation (RPA) and the closely related σ-functionals [Trushin, E.; Thierbach, A.; Görling, A. Toward chemical accuracy at low computational cost: density functional theory with σ-functionals for the correlation energy. *J. Chem. Phys.* **2021**, *154*, 014104] is presented, which is based on a finite-difference approach. The accuracy is evaluated in benchmark calculations using high-quality coupled cluster values as a reference. Our results show that the accuracy of the computed NMR shieldings using direct RPA is strongly dependent on the density functional theory reference orbitals and improves with increasing amounts of exact Hartree−Fock exchange in the functional. NMR shieldings computed with direct RPA using a Hartree−Fock reference are significantly more accurate than MP2 shieldings and comparable to CCSD shieldings. Also, the basis set convergence is analyzed and it is shown that at least triple-zeta basis sets are required for reliable results.

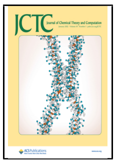


1. INTRODUCTION

NMR shieldings are one of the most often measured spectroscopic parameters in experimental chemistry. Approaches for the accurate theoretical prediction of NMR shieldings are thus of great practical relevance. Throughout the last decades, much work has been done on developing quantum-chemical methods for this spectroscopic quantity.[1,2] Methods for computing NMR shieldings with Hartree−Fock (HF)[3−8] or density functional theory (DFT)[9−11] provide reasonable accuracy at moderate computational cost. Using reduced-scaling techniques, the computation of NMR spectra for molecules with more than 1000 atoms is possible.[12−14] Computationally much more expensive, but also significantly more accurate are methods based on high-level wave functions such as coupled cluster[15−17] or multi-configurational self-consistent field theory.[18,19] Intermediate—both in terms of accuracy and computational cost—are methods for NMR shielding calculations based on second-order Møller−Plesset perturbation theory (MP2). The first analytical derivative method for computing NMR shieldings at the MP2 level was presented by Gauss,[2,20,21] using suitable *Z*-vector[22,23] equations. Later, the applicability of NMR shielding calculations with MP2 was extended using a memory-efficient integral-direct approach[24] or by using local-correlation approximations.[25−27] Another efficient method for computing MP2 shieldings using Cholesky-decomposed two-electron integrals was presented by Burger et al. very recently.[28] Also, linear and sublinear scaling methods based on atomic-orbital MP2 (AO-MP2) were developed.[29] More recently, approaches for computing NMR shieldings with methods closely related to MP2 were also developed including the spin-component-scaled MP2 (SCS-MP2)[30,31] and scaled-opposite-spin MP2 (SOS-MP2)[30,32] as well as double-hybrid DFT.[33,34]

To the best of our knowledge, no NMR shielding calculations with post-Kohn−Sham methods based on the adiabatic-connection fluctuation−dissipation theorem (ACFDT)[35−39] have been carried out to date and we want to fill this gap in our present work. The ACFDT reduces the problem of determining the exact ground-state correlation energy of a many-electron system to integrating an expression involving the frequency-dependent response function along the adiabatic connection path. The various ACFDT-based methods differ in the specific approximation used for calculating the frequency-dependent response functions for the systems with scaled electron−electron interactions. The arguably simplest approximation is the (direct) random phase approximation (RPA),[40] which is obtained by neglecting the exchange−correlation kernel in the Dyson-like equations for the interacting response functions. The earliest formulations of RPA scaled as $O(N^6)$ with system size N and were thus limited

Received: August 27, 2021
Published: December 13, 2021



https://doi.org/10.1021/acs.jctc.1c00866

to rather small molecules.[38] Renewed interest in the RPA method arose after Furche and co-workers[41] showed in 2010 that the scaling can be reduced to $\mathcal{O}(N^4)$ with the help of the resolution-of-the-identity (RI) approximation.[42] RI-RPA thus belongs to the correlation methods with the lowest formal scaling. It is an attractive alternative to MP2, which scales as $\mathcal{O}(N^5)$, or the SOS-MP2[32,43] method, which also scales as $\mathcal{O}(N^4)$ when used in conjunction with the RI approximation and a Laplace transform. In recent years, also several linear-scaling methods based on RI-RPA were developed, which allow to perform RPA calculations even on systems with more than 1000 atoms.[44−47] The accuracy of RPA can be further improved by including exchange effects.[48−54] Alternatively, one might employ the promising σ-functionals, which were recently presented by Görling and co-workers[55,56] and were shown to substantially improve upon the accuracy of RPA, while requiring only minor adaptions to existing implementations and having virtually identical computational cost.

In this work, we present a method for computing NMR shieldings with RPA and the closely related σ-functionals using a finite-field approach, in which the NMR shieldings are computed from a numerical second derivative of the energy. Note that some methods for computing NMR shieldings with RPA have been described and applied in the literature;[1,57−62] however, in these works, RPA has been used as a synonym for "time-dependent HF"; this should not be confused with the ACFDT-based post-Kohn−Sham RPA method that is used in the current work and leads to different results.

Our method is based on RI-RPA, and the working equations for computing the RI-RPA energy in terms of complex molecular orbitals are derived. The method scales as $\mathcal{O}(N^4)$, if the shielding tensor of a single nucleus is computed, and as $\mathcal{O}(N^5)$, if all shieldings are to be computed, but is in general limited to small molecules due to a large prefactor. We expect that a significantly higher efficiency could be obtained with an analytical derivative formulation in combination with a Z-vector approach;[22,23] however, this is outside of the scope of this study and left for future work. Here, we focus on the accuracy for NMR shieldings that can be achieved with RPA or the related σ-functionals. We present calculations on the benchmark set used in refs 63 and 64 and compare the results to accurate CCSD(T) shieldings. The accuracies obtained with RPA and σ-functionals are also compared to other methods, including CCSD, MP2, and several popular density functionals. In addition, the basis set convergence and the influence of the density functional used as a reference for RPA are analyzed.

2. THEORY

2.1. Numerical Differentiation Approach for NMR Shieldings.

The elements of the NMR shielding tensor σ^A of a nucleus A can be computed by differentiating the electronic energy E twice; once with respect to the magnetic field $\vec{B}$ and once with respect to the nuclear magnetic moment $\vec{m}^A$

$$\sigma_{\mathrm{rs}}^A = \frac{\partial^2}{\partial m_{\mathrm{r}}^A \partial B_{\mathrm{s}}} E \bigg|_{\vec{m}^A=0,\vec{B}=0} \quad (1)$$

Both the magnetic field and the nuclear magnetic moment are vector quantities with an x-, y-, and z-component. Therefore, the shielding tensor for any given nucleus A in the molecule is represented by a 3 × 3 matrix. The mixed second derivative from eq 1 can be computed either analytically or numerically. In general, analytical derivatives are preferable for the computation of molecular properties because the corresponding implementations are usually much more efficient. Analytical derivatives also avoid the errors due to a finite step size and the potential numerical instabilities occurring in a finite-difference based scheme.[1,65] The only major advantage of numerical derivatives is the lower complexity of the working equations and the, in consequence, significantly lower programming and debugging effort.[1,65]

A numerical computation of NMR shieldings is possible with a finite-difference scheme as shown, for example, by Fukui et al.[66] Also, nonperturbative approaches for computing the energy in the presence of strong magnetic fields can be used in finite-field calculations of NMR shieldings and other magnetic properties.[67,68] A finite-difference calculation of NMR shieldings requires to perform several energy computations with different finite values for the magnetic field strength and the size of the nuclear magnetic moment. If the energy is considered as a function of the components of the magnetic field and the nuclear magnetic moment:

$$E = E(\vec{B}, \vec{m}^A) = E((B_x, B_y, B_z)^{\mathrm{T}}, (m_x^A, m_y^A, m_z^A)^{\mathrm{T}}) \quad (2)$$

then the component σ_{yx}^A of the shielding tensor, for example, can be approximated as follows[66]

$$\sigma_{yx}^A \approx \frac{1}{2m_y^A B_x}[E((+B_x, 0, 0)^{\mathrm{T}}, (0, +m_y^A, 0)^{\mathrm{T}}) - E((+B_x, 0, 0)^{\mathrm{T}}, (0, -m_y^A, 0)^{\mathrm{T}})] \quad (3)$$

The exact value of σ_{yx}^A would be obtained in the limit $\lim_{m_y^A \to 0}$ and $\lim_{B_x \to 0}$. In practice, if floating-point arithmetic is used, finite values for m_y^A and B_x need to be used. In our implementation, we set both m_y^A and B_x to a constant value δ

$$\sigma_{yx}^A \approx \frac{E((\delta, 0, 0)^{\mathrm{T}}, (0, \delta, 0)^{\mathrm{T}}) - E((\delta, 0, 0)^{\mathrm{T}}, (0, -\delta, 0)^{\mathrm{T}})}{2\delta^2} \quad (4)$$

The other components of the shielding tensor can be computed in an analogous manner. All values presented in this study are isotropic shieldings, which can be computed as follows

$$\sigma_{\mathrm{iso}}^A = \frac{1}{3}(\sigma_{xx}^A + \sigma_{yy}^A + \sigma_{zz}^A) \quad (5)$$

As both the magnetic field and the nuclear magnetic moment perturbations are imaginary, one needs to employ complex orbitals and density matrices in evaluating the energy from eq 2. In order to ensure gauge origin independence, we use gauge-including atomic orbitals (GIAOs)[3,6,69−72] defined as

$$\chi_\mu(\vec{r}, \vec{B}) = \chi_\mu(\vec{r}) \exp\left[-\frac{\mathrm{i}}{2}(\vec{B} \times (\vec{R}_\mu - \vec{R}_0))\cdot\vec{r}\right] \quad (6)$$

in all of our calculations, where $\chi_\mu(\vec{r})$ is a magnetic field-independent basis function and $\vec{R}_\mu$ and $\vec{R}_0$ are the coordinates of the center of basis function χ_μ and of the gauge origin, respectively. The use of GIAOs thus also renders the basis functions complex quantities. Our implementation therefore employs complex matrices and higher order tensors and uses routines for complex linear algebra. This is in contrast to most analytical implementations of NMR shieldings, where real and imaginary components are treated separately and the explicit use of complex linear algebra is usually avoided.[2,6]

2.2. Numerical NMR Shieldings with HF. In order to compute numerical NMR shieldings at the HF level, the HF energy needs to be evaluated in the presence of the magnetic field and nuclear magnetic moment perturbations. For this, one needs to add the matrix representations of all relevant perturbation operators to the one-electron part of the Fock matrix. Following Fukui et al.,[66] the matrix elements $h_{\mu\nu}$ of the core Hamiltonian matrix are computed like this

$$h_{\mu\nu}(\vec{B}, \vec{m}^A) = h_{\mu\nu}^{(0)} + h_{\mu\nu}^{(1,0)}(\vec{B}) + h_{\mu\nu}^{(0,1)}(\vec{m}^A) + h_{\mu\nu}^{(1,1)}(\vec{B}, \vec{m}^A) \tag{7}$$

$h_{\mu\nu}^{(0)}$ contains the kinetic energy integrals $T_{\mu\nu}$ and the nuclear attraction integrals $V_{\mu\nu}$

$$h_{\mu\nu}^{(0)} = T_{\mu\nu} + V_{\mu\nu} \tag{8}$$

$h_{\mu\nu}^{(1,0)}(\vec{B})$ contains the magnetic field derivative of the integrals from $h_{\mu\nu}^{(0)}$ and, in addition, the integrals of the angular momentum operator

$$h_{\mu\nu}^{(1,0)}(\vec{B}) = -\frac{\mathrm{i}}{2}\langle\chi_\mu|\vec{r}_\nu \times \vec{\nabla}|\chi_\nu\rangle\cdot\vec{B} + \left(\frac{\partial T_{\mu\nu}}{\partial \vec{B}} + \frac{\partial V_{\mu\nu}}{\partial \vec{B}}\right)\Bigg|_{\vec{B}=0}\cdot\vec{B} \tag{9}$$

where $\vec{r}_\nu = \vec{r} - \vec{R}_\nu$. Note that here and in the following, all magnetic field-dependent integrals containing GIAOs like $T_{\mu\nu}(\vec{B})$ and $V_{\mu\nu}(\vec{B})$ are not computed exactly, but approximated to first order in the magnetic field. This is sufficient for a property like NMR shieldings that is of first order in the magnetic field. It is also convenient because the first derivatives of GIAO integrals are usually available in programs that allow for NMR shielding calculations with GIAOs.[13]

$h_{\mu\nu}^{(0,1)}(\vec{m}^A)$ is built from the paramagnetic spin−orbit integrals

$$h_{\mu\nu}^{(0,1)}(\vec{m}^A) = -\mathrm{i}\alpha^2\left\langle\chi_\mu\left|\frac{\vec{r}_A \times \vec{\nabla}}{r_A^3}\right|\chi_\nu\right\rangle\cdot\vec{m}^A \tag{10}$$

where $\vec{r}_A = \vec{r} - \vec{R}_A$ and $\vec{R}_A$ is the position of nucleus A and α denotes the fine-structure constant.

Finally, $h_{\mu\nu}^{(1,1)}$ is computed using the diamagnetic spin−orbit integrals and the magnetic field derivative of the paramagnetic spin−orbit integrals from eq 10

$$h_{\mu\nu}^{(1,1)}(\vec{B}, \vec{m}^A) = \frac{\alpha^2}{2}\sum_{rs}\left[\left\langle\chi_\mu\left|\frac{\vec{r}_\nu\cdot\vec{r}_A\delta_{rs} - r_{\nu,r}r_{A,s}}{r_A^3}\right|\chi_\nu\right\rangle + \left\langle\chi_\mu\left|\frac{(\vec{R}_{\mu\nu}\times\vec{r})_r(\vec{r}_A\times\vec{\nabla})_s}{r_A^3}\right|\chi_\nu\right\rangle\right]B_r m_s^A \tag{11}$$

where $\vec{R}_{\mu\nu} = \vec{R}_\mu - \vec{R}_\nu$.

In addition, the magnetic field dependence of the overlap matrix and the electron repulsion integrals needs to be accounted for:

$$S_{\mu\nu}(\vec{B}) \approx S_{\mu\nu}(\vec{B} = 0) + \frac{\partial S_{\mu\nu}}{\partial \vec{B}}\Bigg|_{\vec{B}=0}\cdot\vec{B} \tag{12}$$

$$(\mu\nu|\lambda\sigma)(\vec{B}) \approx (\mu\nu|\lambda\sigma)(\vec{B} = 0) + \frac{\partial(\mu\nu|\lambda\sigma)}{\partial \vec{B}}\Bigg|_{\vec{B}=0}\cdot\vec{B} \tag{13}$$

where $(\mu\nu|\lambda\sigma)$ are the two-electron integrals in the atomic orbital basis.

With the one-electron part from eq 7, the two-electron integrals from eq 13, and the density matrix **P**, whose entries are given by

$$P_{\mu\nu}(\vec{B}, \vec{m}^A) = \sum_i C_{\mu i}^*(\vec{B}, \vec{m}^A)C_{\nu i}(\vec{B}, \vec{m}^A) \tag{14}$$

the complex Fock matrix can be constructed

$$F_{\mu\nu}(\vec{B}, \vec{m}^A) = h_{\mu\nu}(\vec{B}, \vec{m}^A) + J_{\mu\nu}(\vec{B}, \vec{m}^A) - K_{\mu\nu}(\vec{B}, \vec{m}^A) \tag{15}$$

$$J_{\mu\nu}(\vec{B}, \vec{m}^A) = \sum_{\lambda\sigma} P_{\lambda\sigma}(\vec{B}, \vec{m}^A)(\mu\nu|\lambda\sigma)(\vec{B}) \tag{16}$$

$$K_{\mu\nu}(\vec{B}, \vec{m}^A) = \frac{1}{2}\sum_{\lambda\sigma} P_{\lambda\sigma}(\vec{B}, \vec{m}^A)(\mu\sigma|\lambda\nu)(\vec{B}) \tag{17}$$

Updated molecular orbital coefficients in each SCF iteration are then obtained by solving the complex Roothaan−Hall equation

$$\sum_\nu F_{\mu\nu}(\vec{B}, \vec{m}^A)C_{\nu i}(\vec{B}, \vec{m}^A) = \sum_\nu S_{\mu\nu}(\vec{B})C_{\nu i}(\vec{B}, \vec{m}^A)\epsilon_i(\vec{B}, \vec{m}^A) \tag{18}$$

From the converged density matrix, the HF energy can then be computed

$$E_{\mathrm{HF}}(\vec{B}, \vec{m}^A) = \frac{1}{2}\sum_{\mu\nu}(h_{\mu\nu}(\vec{B}, \vec{m}^A) + F_{\mu\nu}(\vec{B}, \vec{m}^A))P_{\mu\nu}(\vec{B}, \vec{m}^A) \tag{19}$$

2.3. Numerical NMR Shieldings with DFT. While molecular orbitals obtained with HF can be used in an RPA calculation, more commonly, density functionals are used as a reference for RPA. For this reason, we also want to discuss here how NMR shieldings can be calculated numerically at the DFT level. In this work, we focus on density functionals that only depend on the electron density (and molecular orbitals in the case of hybrid functionals) and not on other variables such as the (paramagnetic) current density[73] or the magnetic field.[74] Computing numerical DFT shieldings requires several small modifications compared to the described approach for HF. The Kohn−Sham matrix $\mathbf{F}^{\mathrm{KS}}$ for DFT is constructed as follows

$$F_{\mu\nu}^{\mathrm{KS}}(\vec{B}, \vec{m}^A) = h_{\mu\nu}(\vec{B}, \vec{m}^A) + J_{\mu\nu}(\vec{B}, \vec{m}^A) - \beta K_{\mu\nu}(\vec{B}, \vec{m}^A) + V_{\mu\nu}^{\mathrm{xc}}(\vec{B}, \vec{m}^A) \tag{20}$$

where β is the fraction of exact exchange in the functional (0.0 in the case of pure density functionals) and $V_{\mu\nu}^{\mathrm{xc}}$ is the matrix representation of the remaining exchange−correlation potential V_{xc}

$$V^{\mathrm{xc}}_{\mu\nu}(\vec{B}, \vec{m}^A) = \int \chi^*_\mu(\vec{r}, \vec{B}) V_{\mathrm{xc}}(\vec{r}; \vec{B}, \vec{m}^A) \chi_\nu(\vec{r}, \vec{B}) \mathrm{d}\vec{r} \tag{21}$$

In our implementation, the exchange−correlation potential matrix is approximated to first order in the magnetic field

$$V^{\mathrm{xc}}_{\mu\nu}(\vec{B}) \approx V^{\mathrm{xc}}_{\mu\nu}(\vec{B} = 0) + \left.\frac{\partial V^{\mathrm{xc}}_{\mu\nu}}{\partial \vec{B}}\right|_{\vec{B}=0} \cdot \vec{B} \tag{22}$$

where the first magnetic field derivative of $V^{\mathrm{xc}}_{\mu\nu}$ is also commonly available in codes that allow for computing NMR shieldings with DFT.[13] After achieving convergence in the SCF iterations, the DFT energy can then be computed as follows

$$\begin{aligned} E_{\mathrm{DFT}}(\vec{B}, \vec{m}^A) = \sum_{\mu\nu} \Big(& h_{\mu\nu}(\vec{B}, \vec{m}^A) \\ & + \frac{1}{2}(J_{\mu\nu}(\vec{B}, \vec{m}^A) \\ & - \beta K_{\mu\nu}(\vec{B}, \vec{m}^A)) \Big) P_{\mu\nu}(\vec{B}, \vec{m}^A) \\ & + E_{\mathrm{xc}}[\mathrm{Re}(\rho(\vec{r}; \vec{B}, \vec{m}^A))] \end{aligned} \tag{23}$$

where E_{xc} is the exchange−correlation energy and $\rho(\vec{r}; \vec{B}, \vec{m}^A)$ is the real-space electron density, which can be computed according to the following equation

$$\rho(\vec{r}; \vec{B}, \vec{m}^A) = \sum_{\mu\nu} \chi^*_\mu(\vec{r}, \vec{B}) P_{\mu\nu}(\vec{B}, \vec{m}^A) \chi_\nu(\vec{r}, \vec{B}) \tag{24}$$

For a complex density matrix and GIAO basis functions, the electron density also has a nonvanishing imaginary component, which, however, is unphysical and an artifact of the employed approach. For this reason, only the real part of the electron density is used in eq 23. Note that the product of the imaginary part of **P** and the imaginary part of the GIAO basis functions also results in a contribution to the real part of the electron density, which needs to be taken into account.

2.4. Numerical NMR Shieldings with Correlation Methods.

The complex orbitals obtained from the SCF procedures described in Sections 2.2 and 2.3 can be used to compute correlation energies in the presence of the perturbations of the magnetic field and nuclear magnetic moment. These perturbed energies can then be inserted into eq 3 in order to obtain the correlation contribution to the NMR shieldings.

As an example, consider the MP2 method, for which the correlation energy $E^{\mathrm{corr}}_{\mathrm{MP2}}$ can be computed from complex spin orbitals as follows[66]

$$E^{\mathrm{corr}}_{\mathrm{MP2}} = \frac{1}{4} \sum_{ijab} \frac{|(ai|bj) - (aj|bi)|^2}{\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j} \tag{25}$$

The ordering of the MO indices in the four-center integrals is important here because it determines which spin orbitals need to be complex-conjugated. In the employed Mulliken notation, the first MO of bra and ket is complex-conjugated

$$(pq|rs) = \iint \varphi^*_p(\vec{x}_1) \varphi_q(\vec{x}_1) \frac{1}{|r_{12}|} \varphi^*_r(\vec{x}_2) \varphi_s(\vec{x}_2) \mathrm{d}\vec{x}_1 \, \mathrm{d}\vec{x}_2 \tag{26}$$

where p, q, r, s are general (occupied or virtual) MO indices. $\vec{x}_1$ and $\vec{x}_2$ denote combined spatial and spin coordinates of electrons 1 and 2, respectively. The full MP2 shieldings can then be obtained by using the total MP2 energy

$$E_{\mathrm{MP2}} = E_{\mathrm{HF}} + E^{\mathrm{corr}}_{\mathrm{MP2}} \tag{27}$$

in eq 3.

The focus of the current work lies on NMR shieldings obtained with RPA, which requires computations of RPA correlation energies in terms of complex orbitals. With a scaling of $O(N^6)$, this is possible using equations that are similar to equations presented in ref 75:

$$E^{\mathrm{corr}}_{\mathrm{RPA}} = \frac{1}{2}\left\{ \sum_n \omega_n - \mathrm{Tr}[\mathbf{A}] \right\} \tag{28}$$

where the positive excitation energies ω_n can be obtained by solving the RPA eigenvalue problem

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ -\mathbf{B}^* & -\mathbf{A}^* \end{bmatrix} \begin{pmatrix} x \\ y \end{pmatrix}_n = \omega_n \begin{pmatrix} x \\ y \end{pmatrix}_n \tag{29}$$

where

$$A_{ia,jb} = (\epsilon_a - \epsilon_i)\delta_{ij}\delta_{ab} + (ia|bj) \tag{30}$$

$$B_{ia,jb} = (ai|bj) \tag{31}$$

In order to obtain the total RPA energy, one must add the correlation energy from eq 28 to the expectation value of the Hamilton operator evaluated with a single Slater determinant Ψ_{SD} (build from the MOs obtained from the HF or DFT reference calculation) as a wave function

$$E_{\mathrm{RPA}} = \frac{\langle \Psi_{\mathrm{SD}} | \hat{H} | \Psi_{\mathrm{SD}} \rangle}{\langle \Psi_{\mathrm{SD}} | \Psi_{\mathrm{SD}} \rangle} + E^{\mathrm{corr}}_{\mathrm{RPA}} \tag{32}$$

Using the energy from eq 32 in eq 3 then allows to numerically compute NMR shieldings with RPA.

For efficiency reasons, our implementation does not employ eqs 28 to 31 for computing $E^{\mathrm{corr}}_{\mathrm{RPA}}$, but is based on the RI-RPA method of Furche and co-workers[41] instead, which scales only as $O(N^4)$. We are not using our low- and linear-scaling RI-RPA methods[45−47] because we focus here only on very small molecules. In the following, a brief derivation of the working equations will be provided. We start from the following expression for the RPA correlation energy[76,77]

$$\begin{aligned} E^{\mathrm{corr}}_{\mathrm{RPA}} = & \frac{1}{2\pi}\int_0^{+\infty} \mathrm{d}\omega \iint \mathrm{d}\vec{x}_1 \, \mathrm{d}\vec{x}_2 [\ln(1 - V(\vec{x}_1, \vec{x}_2)\chi_0(\vec{x}_1, \vec{x}_2, \mathrm{i}\omega))] \\ & + \frac{1}{2\pi}\int_0^{+\infty} \mathrm{d}\omega \iint \mathrm{d}\vec{x}_1 \, \mathrm{d}\vec{x}_2 [V(\vec{x}_1, \vec{x}_2)\chi_0(\vec{x}_1, \vec{x}_2, \mathrm{i}\omega)] \end{aligned} \tag{33}$$

where $V(\vec{x}_1,\vec{x}_2) = 1/r_{12}$ is the interelectron potential and $\chi_0(\vec{x}_1,\vec{x}_2,\mathrm{i}\omega)$ is the frequency-dependent Kohn−Sham response function[76,77]

$$\begin{aligned} \chi_0(\vec{x}_1, \vec{x}_2, \mathrm{i}\omega) = \sum_{ia} \Bigg[& \frac{\varphi^*_i(\vec{x}_1)\varphi_a(\vec{x}_1)\varphi^*_a(\vec{x}_2)\varphi_i(\vec{x}_2)}{+\mathrm{i}\omega - \Omega_{ia}} \\ & - \frac{\varphi^*_a(\vec{x}_1)\varphi_i(\vec{x}_1)\varphi^*_i(\vec{x}_2)\varphi_a(\vec{x}_2)}{+\mathrm{i}\omega + \Omega_{ia}} \Bigg] \end{aligned} \tag{34}$$

with

$$\Omega_{ia} = \epsilon_a - \epsilon_i \tag{35}$$

Let us now consider the second term of eq 33 and perform the integration over the electronic coordinates

$$\iint \mathrm{d}\vec{x}_1\, \mathrm{d}\vec{x}_2 V(\vec{x}_1, \vec{x}_2)\chi_0(\vec{x}_1, \vec{x}_2, \mathrm{i}\omega) = \sum_{ia}\left[\frac{(ia|ai)}{\mathrm{i}\omega - \Omega_{ia}} - \frac{(ai|ia)}{\mathrm{i}\omega + \Omega_{ia}}\right] \tag{36}$$

The two-electron integrals appearing in this equation can be approximated using RI

$$(ia|ai) = \sum_{PQ} (ia|P)(P|Q)^{-1}(Q|ai) \tag{37}$$

Note that also in the three-center integrals, the left MO of an orbital pair in either bra or ket is complex-conjugated (e.g., φ_i in $(ia|P)$). In contrast, the auxiliary functions χ_P and χ_Q are always taken to be real. Inserting the RI approximation in eq 36 gives

$$\iint \mathrm{d}\vec{x}_1\, \mathrm{d}\vec{x}_2 V(\vec{x}_1, \vec{x}_2)\chi_0(\vec{x}_1, \vec{x}_2, \mathrm{i}\omega) = \sum_{ia}\sum_{PQ}\left[\frac{(ia|P)(P|Q)^{-1}(Q|ai)}{\mathrm{i}\omega - \Omega_{ia}} - \frac{(ai|P)(P|Q)^{-1}(Q|ia)}{\mathrm{i}\omega + \Omega_{ia}}\right] \tag{38}$$

Next, the summation over i and a is performed

$$\iint \mathrm{d}\vec{r}_1\, \mathrm{d}\vec{x}_2\ V(\vec{x}_1, \vec{x}_2)\chi_0(\vec{x}_1, \vec{x}_2, \mathrm{i}\omega) = \sum_{PQ} [\Pi^-_{QP}(\mathrm{i}\omega) - \Pi^+_{QP}(\mathrm{i}\omega)]V_{PQ} \tag{39}$$

where $V_{PQ} = (P|Q)^{-1}$ and

$$\Pi^-_{QP}(\mathrm{i}\omega) = \sum_{ia} (Q|ai)\frac{1}{\mathrm{i}\omega - \Omega_{ia}}(ia|P) \tag{40}$$

$$\Pi^+_{QP}(\mathrm{i}\omega) = \sum_{ia} (Q|ia)\frac{1}{\mathrm{i}\omega + \Omega_{ia}}(ai|P) \tag{41}$$

A similar series of manipulations, as shown in eqs 36 to 39, can be applied to the first term of eq 33; combined with the result from eq 39, this leads to the final formula for the RI-RPA correlation energy in terms of complex orbitals

$$E^{\mathrm{corr}}_{\mathrm{RPA}} = \frac{1}{2\pi}\int_0^{+\infty} \mathrm{d}\omega\ \mathrm{Tr}[\ln[1 - (\mathbf{\Pi}^-(\mathrm{i}\omega) - \mathbf{\Pi}^+(\mathrm{i}\omega))\mathbf{V}] + (\mathbf{\Pi}^-(\mathrm{i}\omega) - \mathbf{\Pi}^+(\mathrm{i}\omega))\mathbf{V}] \tag{42}$$

The frequency integration over ω can be approximated using numerical integration.[41]

3. COMPUTATIONAL DETAILS

The described method has been implemented in a development version of the FermiONs++ program.[78−80] For the field-strength δ in eq 4, a value of 0.1 a.u. is used in all calculations shown in Section 2; this value might appear to be rather large, but one needs to also consider that the perturbation operators and the GIAO integral derivatives contain factors of the fine-structure constant α or factors of α^2. As $\alpha \approx 7 \times 10^{-3}$, the magnitude of the perturbations is thus reduced substantially. With $\delta = 0.1$ a.u., changes of the energy (compared to the unperturbed case) on the order of 10 μHartree are obtained in most cases. The numerical frequency integration is carried out using the Clenshaw−Curtis scheme[41] with 120 integration points. We include all electrons in the correlation treatment and do not apply a frozen-core approximation. As a starting guess for the complex SCF, we use the converged real density matrix from the standard SCF in FermiONs++ and set the imaginary part to zero. The SCF calculations are converged tightly to an energy difference of below 10^{-16}. No integral screening or sparse linear algebra is used. As a numerical integration grid for DFT, we employ the "g5" grid from Laqua et al. described in detail in ref 81. Several different density functionals are used as a starting point for the RPA correlation energy calculations, including the PBE,[82] KT2,[83] B3LYP,[84] B97-2,[85] and the Becke half-and-half functional[86] (which we will in the following abbreviate as BH&H).

In the presented calculations, we employ correlation-consistent basis sets from the cc-pVXZ series[87,88] and the cc-pCVXZ series[89,90] with additional core functions. For each of these basis sets, we use the corresponding auxiliary basis set[91,92] in the RI-RPA calculations. In addition, we also use basis sets from the def2 series[93] with the corresponding RI-basis sets[94] and basis sets of the pcSseg series[95,96] (in combination with the cc-pVQZ/mp2fit basis set).

In order to confirm the correctness of our implementation and in order to show that the errors due to numerical differentiation and the numerical frequency integration are insignificant, we performed test calculations with a cc-pVTZ basis set, in which we varied the field strength and the number of frequency integration points used for RPA as well as the number of DFT grid points. In addition, we tested that the gauge origin independence is not violated by carrying out additional calculations after shifting the gauge origin by 5.0 Å in x-, y-, and z-direction. The data from all these calculations are shown in the Supporting Information. The results show that gauge origin independence is fulfilled to a very good approximation, as the changes of isotropic shieldings upon translation of the molecules did not exceed 0.1 ppm in all cases. Furthermore, for all molecules and basis sets, the computed HF or DFT contributions to the numerical shieldings agree well with the analytically computed values from ref 63. Therefore, we conclude that the presented numerical differentiation approach with the employed parameters is accurate enough for judging the performance of direct RPA and σ-functionals for NMR shieldings. For the pcSseg-0 basis set, we also compared calculations with different auxiliary basis sets (see Supporting Information) and concluded that the RI errors are small compared to the method and basis set errors.

All calculations in this study are performed on molecules from the benchmark set used by Teale et al.[63] We only exclude the molecules O_3 and SO_2, as it has also been done in most of the analysis from ref 63. All together, the employed benchmark set comprises 68 nuclei, for which shieldings were computed. For comparison, we performed CCSD(T) shielding calculations with the CFOUR program.[97,98] Values for extrapolated CCSD(T) shieldings were taken from ref 63. In addition, we performed MP2 shielding calculations with the Turbomole program package.[24,99,100] All computed NMR shieldings can be found in the Supporting Information.

4. RESULTS

4.1. Basis Set Convergence.

As Teale et al.[63] found that diffuse basis functions have only a very small impact on correlated NMR shieldings, we focus on basis sets without diffuse functions in the present study. Figure 1 displays the basis set errors obtained with RPA@HF for basis sets from the

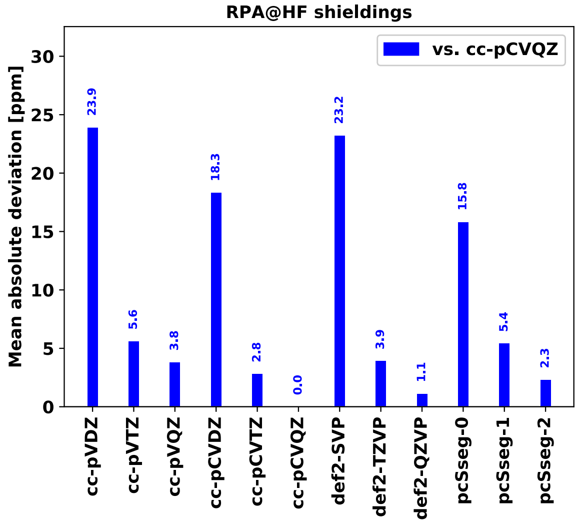


Figure 1. Basis set convergence of NMR shieldings at the RPA@HF level of theory. The blue bars display the deviations from RPA@HF shieldings with a cc-pCVQZ basis set.

cc-pVXZ and cc-pCVXZ series ($X \in \{D,T,Q\}$) as well as the def2 and pcSseg series. The mean absolute deviations (MADs) are computed relative to RPA@HF with the large cc-pCVQZ basis set, which we assume to give results close to the basis set limit. RPA@HF stands for RPA with a HF reference; a similar notation will be employed in the following for RPA calculations with several different density functionals as a reference. The basis set convergence of RPA@PBE is shown in Figure 2, again using shieldings obtained in a cc-pCVQZ basis as a reference. For both a HF and PBE reference, the basis sets from the cc-pCVXZ series with additional core functions provide slightly better accuracy than the corresponding basis sets from the cc-pVXZ series, which is consistent with the findings from Teale et al.[63] The convergence behavior for the basis sets from the def2 series is similar to that for the cc-pVXZ series. In particular, the def2-QZVP basis set gives very accurate results and a smaller mean error than the cc-pVQZ basis set, which has a similar number of basis functions per atom. Also, the basis sets from the pcSseg series appear highly suitable for RPA shielding calculations. The pcSseg-2 basis set, for example, leads to significantly smaller errors than the cc-pVTZ basis set, but contains only slightly more basis functions per atom.

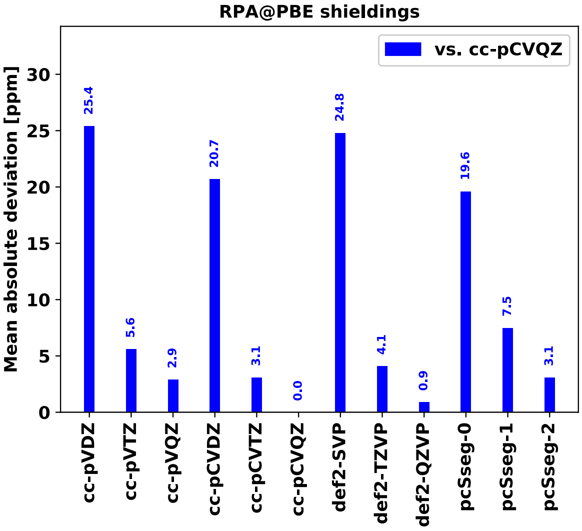


Figure 2. Basis set convergence of NMR shieldings at the RPA@PBE level of theory. The blue bars display the deviations from RPA@PBE shieldings with a cc-pCVQZ basis set.

In order to analyze potential systematic compensations between basis set and method errors, the RPA@HF shielding values are also compared to CCSD(T) shieldings, as shown in Figure 3. The deviations to CCSD(T) with the same basis set, which reflect the method error of RPA@HF, are relatively constant across different basis sets and amount to roughly 5 ppm. Also, the deviations from CCSD(T) extrapolated to the basis set limit (details for the extrapolation can be found in ref 63) are displayed, which can be attributed to a combination of basis set incompleteness error and method error. The results show that the double-zeta basis sets cc-pVDZ and cc-pCVDZ and the def2-SVP and pcSseg-0 basis sets are not sufficiently large for obtaining accurate shieldings, as the basis set error clearly exceeds the method error with these basis sets. For all other basis sets, the deviations from basis set extrapolated CCSD(T) are rather similar in magnitude and are between 4.9 and 6.2 ppm. Triple-zeta basis sets like cc-pVTZ might therefore be particularly useful in practice, as they offer a good compromise between accuracy and computational cost.

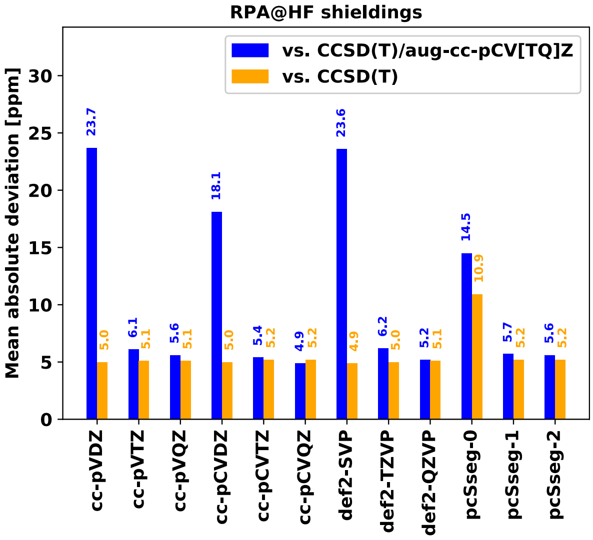


Figure 3. Mean absolute deviations of NMR shieldings computed with RPA@HF on the employed test set with several different basis sets. The blue bars display the deviations from basis set extrapolated (aug-cc-pCV[TQ]Z) CCSD(T)-shieldings; the orange bars display the deviations from CCSD(T) with the same basis set as used for the RPA calculation.

Figure 4 displays the MADs obtained with the RPA@PBE method and different basis sets. In general, the deviations from the coupled cluster references are significantly larger than for RPA@HF, but the trends are similar to RPA@HF. Similar data for the σ-functional[55] with parametrization "σ-P1" are shown in the Supporting Information. Also for this functional, large errors are obtained with double-zeta basis sets (MADs of 31.2 ppm for cc-pVDZ and 26.6 ppm for cc-pCVDZ with respect to CCSD(T)/aug-cc-pCV[TQ]Z, respectively). Significantly smaller errors are obtained with triple-zeta (MAD of

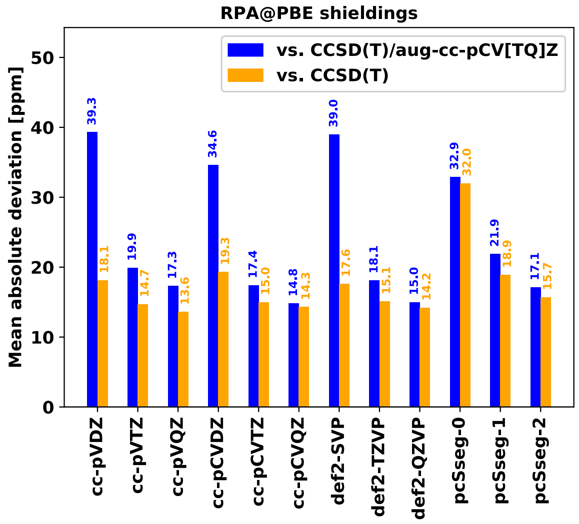


Figure 4. Mean absolute deviations of NMR shieldings computed with RPA@PBE on the employed test set with several different basis sets. The blue bars display the deviations from basis set extrapolated (aug-cc-pCV[TQ]Z) CCSD(T)-shieldings; the orange bars display the deviations from CCSD(T) with the same basis set as used for the RPA calculation.

11.5 ppm with cc-pVTZ and 9.5 ppm with cc-pCVTZ) and quadruple-zeta basis sets (MAD of 9.3 ppm with cc-pVQZ and 7.0 ppm with cc-pCVQZ).

4.2. Influence of the Chosen Density Functional.

Figures 5 and 6 illustrate (amongst others) the mean absolute errors in the NMR shieldings obtained with RPA using different reference functionals. For both basis sets, the dependence of the accuracy of RPA shieldings on the employed DFT reference is rather pronounced. There seems to be a strong correlation between the accuracy and the fraction of exact exchange in the functional used as a reference. The worst accuracy is obtained with the PBE functional as a reference, which results in a mean absolute error of 19.9 ppm relative to extrapolated CCSD(T) shieldings using the cc-pVTZ basis set. RPA@KT2 performs slightly better with an MAD of 16.1 ppm. Like PBE, the KT2 functional is a pure density functional without exact exchange. Significantly higher accuracy is observed if the DFT reference includes a fraction of exact exchange. The MAD drops to 15.2 ppm for B3LYP with 20% exact exchange, to 13.8 ppm for B97-2 with 21% exact exchange, and to 10.7 ppm for Becke-half-and-half with 50% exact exchange. The best accuracy and an MAD of only 6.1 ppm is observed using a HF reference with 100% exact exchange.

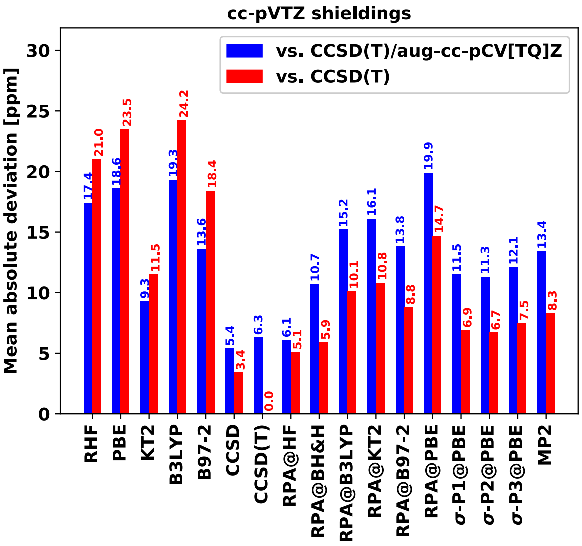


Figure 5. Mean absolute deviations of NMR shieldings on the employed test set using the cc-pVTZ basis set in all calculations. The blue bars display the deviation from basis set extrapolated (aug-cc-pCV[TQ]Z) CCSD(T) shieldings; the red bars display deviations from CCSD(T) shieldings obtained with a cc-pVTZ basis set. The cc-pVTZ/mp2fit auxiliary basis set is used in all RPA and σ-functional calculations. The data for RHF, CCSD, PBE, KT2, B3LYP, and B97-2 were taken from ref 63.

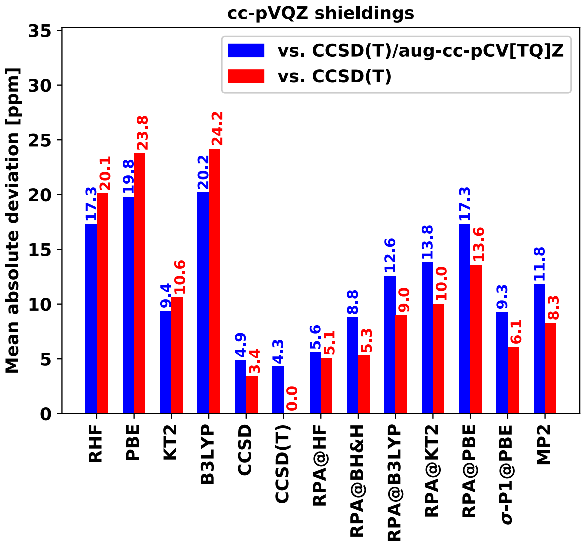


Figure 6. Mean absolute deviations of NMR shieldings on the employed test set using the cc-pVQZ basis set in all calculations. The blue bars display the deviation from basis-set extrapolated (aug-cc-pCV[TQ]Z) CCSD(T) shieldings; the red bars display deviations from CCSD(T) shieldings obtained with a cc-pVQZ basis set. The cc-pVQZ/mp2fit auxiliary basis set is used in all RPA and σ-functional calculations. The data for RHF, CCSD, PBE, KT2, B3LYP, and B97-2 were taken from ref 63.

It is known that NMR shieldings are a challenging molecular property for DFT and many popular density functionals fail to give accurate shieldings. This can also be seen in the data for DFT from Teale et al.,[63] which is also displayed in Figures 5 and 6. With the exception of the KT2 functional (and to some extent B97-2), none of the used functionals provides a substantial improvement over HF in the quality of the shieldings. This is in contrast to total ground-state energies, where DFT often surpasses HF in accuracy. The rather poor performance of many density functionals for NMR shieldings stems from errors in the paramagnetic terms, which has been explained by Reimann et al. with missing paramagnetic current contributions in conventional DFT.[101] It has also been shown that common density functionals tend to systematically overestimate the magnitude of the paramagnetic contributions, which is linked to a systematic underestimation of the electronic gap especially for pure density functionals.[102] More generally, conventional Kohn–Sham DFT suffers from fundamental limitations in the treatment of magnetic fields. Several extensions of KS-DFT have been proposed,[73,74,103] of

https://doi.org/10.1021/acs.jctc.1c00866
J. Chem. Theory Comput. 2022, 18, 192−205

which current DFT[73] is the most popular. Thus, Kohn−Sham orbitals and orbital energies obtained with functionals that only depend on the density might be suboptimal for a subsequent RPA calculation in the presence of magnetic fields. This could possibly explain the lower accuracy of the computed shieldings compared to RPA@HF and might also apply to other magnetic properties at the RPA level of theory.

For both a cc-pVTZ and a cc-pVQZ basis, RPA reduces the errors compared to HF and DFT, if CCSD(T) with the same basis set is used as a reference (see Figures 5 and 6). This reduction is small for KT2, but significant for HF and all other considered density functionals. The improvement of the RPA methods over the underlying reference methods is not as pronounced, if the deviations from extrapolated CCSD(T) shieldings are considered instead. In some cases like, for example, PBE with a cc-pVTZ basis set, the mean absolute error is even slightly larger for the corresponding RPA@PBE method. Strikingly, for all RPA methods, the errors compared to CCSD(T) with the same basis set are significantly smaller than the errors compared to basis set extrapolated CCSD(T). In contrast, for HF and the considered density functionals, the deviations from CCSD(T) with the same basis set are consistently larger. This suggests that HF and DFT seem to benefit from a compensation of basis set incompleteness and method errors. It also suggests that RPA does, in principle, improve the accuracy compared to the employed references, but this improvement seems reduced by a less favorable cancelation of errors. Among the considered density functionals, KT2 is special because it is fitted to give good NMR shieldings.[83] However, it leads to significant errors in the diamagnetic and paramagnetic shieldings and achieves good performance for NMR shieldings only by significant error cancelation.[101] KT2 orbitals and orbital energies thus might not constitute a good reference for RPA shielding calculations. This is the likely reason, why shieldings obtained with RPA@KT2 are not significantly more accurate than KT2 shieldings.

In order to examine the influence of different DFT references in more detail, we also computed mean unsigned and signed errors for the diamagnetic and paramagnetic contributions shown in Figures 7 and 8 as well as mean signed errors for the total shieldings shown in Figure 9. The diamagnetic contributions with RPA and σ-functionals were computed numerically as described in Section 2, but setting $\mathbf{h}^{(1,0)}$ and $\mathbf{h}^{(0,1)}$ in eq 7 and the GIAO integral derivatives from eqs 12, 13, and 22 to zero; the paramagnetic contributions were then computed by subtracting the diamagnetic contributions from the total NMR shieldings. With this definition of the diamagnetic and paramagnetic contributions, all contributions that would require a perturbed density from a CPSCF in an analytical derivative approach are defined to be paramagnetic. The diamagnetic contributions include all terms that would only involve the unperturbed ground-state density in an analytical derivative approach. The employed definition for diamagnetic and paramagnetic contributions is also consistent with the CCSD(T) implementation in CFOUR,[97,98] which is used for computing reference values for these contributions. Note that the diamagnetic and paramagnetic contributions are not gauge origin-independent if defined like this; gauge origin independence could be achieved by using alternative definitions of diamagnetic and paramagnetic contributions (see refs 14, 104, and 105). However, as we only compare different methods with exactly the same molecular geometry and choice of gauge origin, the missing gauge origin independence is not problematic here. In all calculations of diamagnetic and paramagnetic shielding contributions, the position of the considered nucleus was used as gauge origin.

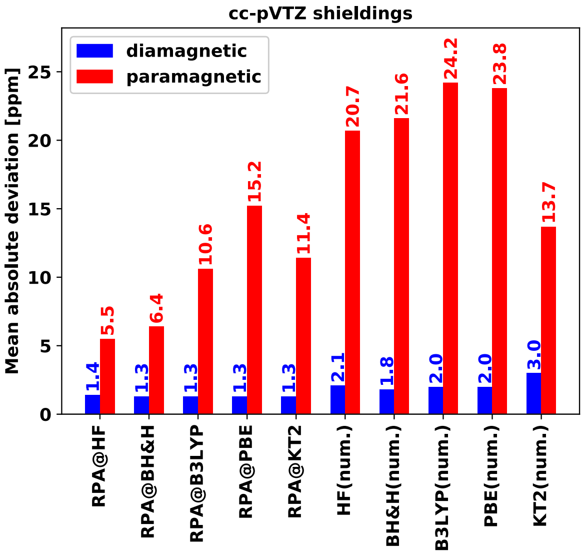


Figure 7. Mean absolute errors in the diamagnetic (blue bars) and paramagnetic (red bars) contributions to the NMR shieldings using the cc-pVTZ basis set in all calculations. The cc-pVTZ/mp2fit auxiliary basis set is used in all RPA and σ-functional calculations. The diamagnetic and paramagnetic contributions to the CCSD(T) shieldings in a cc-pVTZ basis are used as a reference, respectively.

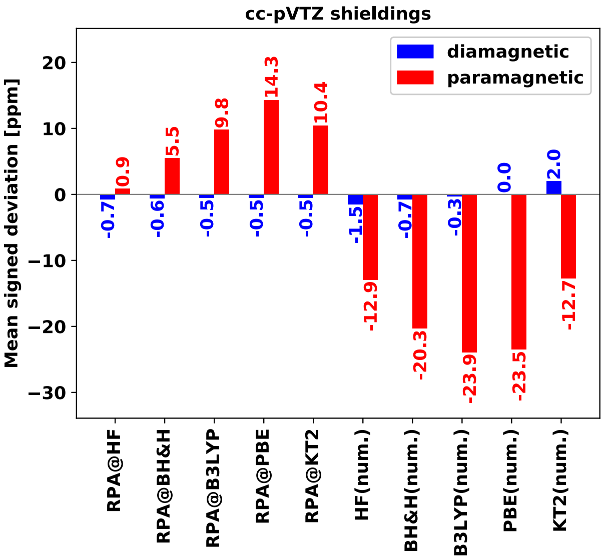


Figure 8. Mean signed errors in the diamagnetic (blue bars) and paramagnetic (red bars) contributions to the NMR shieldings using the cc-pVTZ basis set in all calculations. The cc-pVTZ/mp2fit auxiliary basis set is used in all RPA and σ-functional calculations. The diamagnetic and paramagnetic contributions to the CCSD(T) shieldings in a cc-pVTZ basis are used as a reference, respectively.

For all considered RPA methods, the errors in the diamagnetic contributions are of similar magnitude and significantly smaller than the errors in the paramagnetic contributions. Irrespective of the reference functional, RPA tends to underestimate the diamagnetic contributions to a small extent with mean absolute errors not larger than 1.4 ppm.

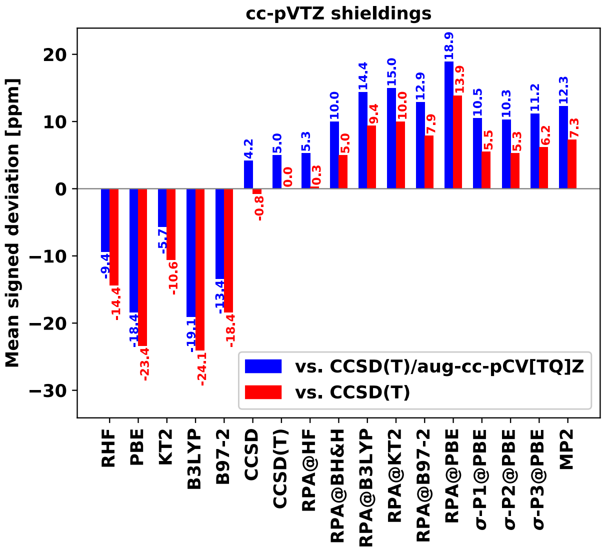


Figure 9. Mean signed errors of NMR shieldings on the employed test set using the cc-pVTZ basis set in all calculations. The blue bars display the deviation from basis-set extrapolated (aug-cc-pCV[TQ]Z) CCSD(T)-shieldings; the red bars display deviations from CCSD(T)-shieldings obtained with a cc-pVQZ basis set. The cc-pVTZ/mp2fit auxiliary basis set is used in all RPA and σ-functional calculations.

With a HF reference, slightly larger errors in the diamagnetic contributions are obtained than with the considered density functionals as references, which is in contrast to the total shieldings. The diamagnetic contributions do not depend on the magnetic field response of the density, but depend only on the unperturbed ground-state density, which is typically more accurate for a Kohn−Sham reference as for HF due to included correlation. This might explain the better performance of RPA with Kohn−Sham references for the diamagnetic contributions and is in line with the findings from Burow et al.[106] for first-order molecular properties, where RPA with a Kohn−Sham reference leads to more accurate results.

While the absolute errors in the paramagnetic contributions are comparably small for RPA@HF and amount to 5.5 ppm on average, significantly larger errors are observed for the employed density functionals as a reference. Additionally, a strong correlation with the fraction of exact exchange in the functional is observed. The RPA@PBE method with a pure density functional as a reference shows the largest errors in the paramagnetic contributions with a MAD of 15.2 ppm. As the errors in the paramagnetic contributions are significantly larger than the errors in the diamagnetic shieldings, they are also the major source of error for the total shieldings. Accordingly, because of the too high values for the paramagnetic contributions (see Figure 8), most RPA methods also give total shieldings that are significantly too high, which can be seen from the mean signed errors shown in Figure 9. Figures 7 and 8 also show the errors in the diamagnetic and paramagnetic contributions for the HF method and several density functionals. In contrast to RPA, HF and all density functional methods lead to a significant underestimation of the paramagnetic contributions. For the KT2 functional, subsequent RPA calculations only marginally improve the accuracy of the paramagnetic contributions. For all other references, RPA leads to significantly reduced errors in the paramagnetic contributions. For HF and all considered density functionals, RPA leads to a significant reduction of the errors in the diamagnetic contributions.

The way RPA improves upon different density functionals or HF references regarding the description of the electronic structure in the presence of magnetic fields remains an interesting topic for further investigation. For this purpose, it might be fruitful to exploit the formal equivalence of direct RPA and ring-coupled cluster theory.[107] A wave function at the RPA level of theory could be constructed by computing doubles amplitudes with the Riccati equation,[107] from which reduced density matrices and an electron density could be computed. Analyzing wave function and density could allow for further insights into the description of magnetic field perturbations by the RPA method. Such an analysis, however, is not possible with our current implementation and thus out of the scope of the present study.

Figures 5 and 6 further show the MADs obtained with one of the σ-functionals introduced by Görling and co-workers[55] using the parametrizations "σ-P1", "σ-P2", and "σ-P3". These σ-functionals were parametrized for a PBE reference and all give significantly more accurate shieldings than the corresponding RPA-based method RPA@PBE. The σ-functional with parametrization "σ-P3" gives slightly less accurate shieldings than "σ-P1" and "σ-P2", but the differences are small. The results suggest that σ-functionals may have great potential not only for accurately computing energies as shown in ref 55, but also for computing accurate NMR shieldings. However, the employed σ-functionals provide inferior accuracy compared to RPA@HF. It might be possible to further increase the accuracy of σ-functionals by developing new parametrizations, which are tailored for NMR shieldings, or by parametrizing σ-functionals for other DFT or HF references.

Due to its excellent accuracy, we recommend at this stage that RPA@HF should be the method of choice for NMR shielding calculations at the RPA level of theory. The use of HF as a reference for the RPA calculation is also desirable from a puristic point of view, as RPA@HF does not involve any empirical parameters and thus can be regarded as a true ab initio method.

4.3. Comparison with Other Methods.

In this section, we compare the accuracy of RPA, DFT, and several wavefunction-based methods; among the RPA-based methods, we focus on the most accurate RPA@HF. Compared to basis set extrapolated CCSD(T), RPA@HF is almost as accurate as CCSD and also CCSD(T) for both a cc-pVTZ and cc-pVQZ basis set (see Figures 5 and 6). This is presumably due to a more fortunate cancellation of basis set and method errors for RPA@HF. The errors in the shieldings obtained with RPA@HF are also considerably smaller than the errors obtained with HF or all of the considered density functionals. The comparison with MP2 is particularly interesting because MP2 belongs to the computationally cheapest methods for treating electron correlation with a formal scaling of $O(N^5)$ and at the same time providing accurate relative NMR shieldings (see, e.g., ref 108). RPA@HF in contrast scales only as $O(N^4)$ and significantly outperforms MP2 in the presented benchmark calculations; the MAD relative to basis-set extrapolated CCSD(T) amounts to 11.8 ppm for MP2 in a cc-pVQZ basis, while it amounts to only 5.6 ppm for RPA@HF. This makes RPA@HF a very attractive alternative to MP2 for the computation of NMR shieldings. Similar trends, concerning the relative accuracy of the considered methods, are observed for the pcSseg-2 basis set; the corresponding plot is displayed in the Supporting Information.

In Figure 10, the errors of different methods are analyzed separately for different types of nuclei (second-row elements carbon to fluorine). The molecules in the employed benchmark set together contain 17 carbon, 7 nitrogen, 10 oxygen, and 9 fluorine nuclei. In general, all RPA-based methods produce the largest errors for oxygen. RPA@HF provides excellent accuracy for carbon and nitrogen shieldings with MADs below 4 ppm and slightly larger errors for oxygen and fluorine shieldings. The data suggest that RPA@HF is a good alternative to MP2 especially for nitrogen shieldings, for which MP2 leads to a much larger MAD of 32.3 ppm.

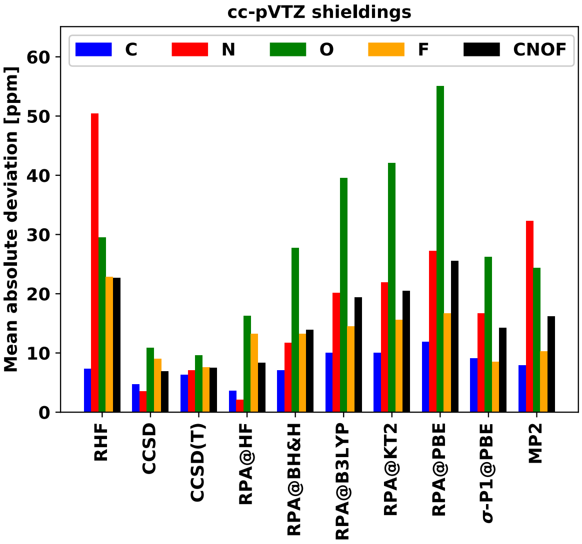


Figure 10. Mean absolute deviations of carbon, nitrogen, oxygen, and fluorine NMR shieldings using the cc-pVTZ basis set in all calculations. "CNOF" includes shieldings from all of the considered elements. The cc-pVTZ/mp2fit auxiliary basis set is used in all RPA and σ-functional calculations. Basis-set extrapolated (aug-cc-pCV[TQ]Z) CCSD(T)-shieldings are used as a reference.

Both in practical applications of NMR spectroscopy and theoretical simulations of NMR spectra, NMR shifts relative to a reference compound are most widely used as opposed to absolute shieldings considered so far. The performance of the analyzed methods for NMR shifts of the second-row elements carbon to fluorine is shown in Figure 11 using the respective hydrides as a reference. In general, the differences between the mean errors shown in Figures 10 and 11 are rather small. On the employed test set, MP2 benefits more from systematic error cancellation than the other methods and shows the greatest reduction in the mean errors between absolute and relative shieldings. Nevertheless, MP2 is slightly less accurate than RPA@HF also for relative shieldings and significantly less accurate for relative nitrogen shieldings.

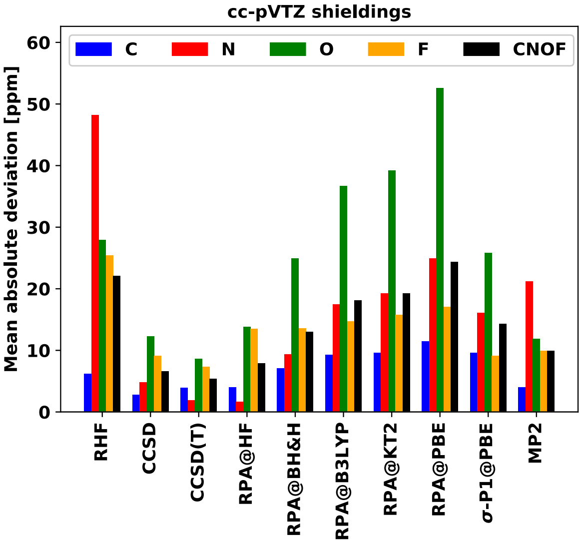


Figure 11. Mean absolute deviations of relative NMR shieldings for carbon, nitrogen, oxygen, and fluorine using the cc-pVTZ basis set in all calculations. The relative shifts were computed using the following references: CH_4 for carbon, NH_3 for nitrogen, H_2O for oxygen, and HF for fluorine. "CNOF" includes relative shieldings from all of the considered elements. The cc-pVTZ/mp2fit auxiliary basis set is used in all RPA and σ-functional calculations. Basis-set extrapolated (aug-cc-pCV[TQ]Z) CCSD(T)-shieldings are used as a reference.

5. CONCLUSIONS

We presented a method for computing NMR shielding constants within the direct RPA and the recently introduced σ-functionals from Trushin et al.[55] using a finite-difference approach. Benchmark calculations on a set of small molecules indicated that the accuracy of RPA NMR shieldings strongly depends on the employed reference orbitals and orbital energies; a quite systematic improvement was observed with increasing amounts of exact HF exchange in the functional; the best results were obtained with the HF state as a reference. In this work, density functionals were employed that depend only on the electron density and the molecular orbitals; an interesting question for further studies is whether orbitals obtained with current density functionals[73] and similar approaches[74,103] would enable more accurate RPA shielding calculations.

An analysis of the basis set convergence showed that at least triple-zeta basis sets are required for obtaining reliable NMR shieldings at the RPA level. With HF as a reference and using a quadruple-zeta basis, the accuracy of RPA shieldings are significantly higher than the accuracy of MP2 shieldings and comparable to CCSD shieldings. NMR shieldings computed with the σ-functionals are more accurate than the corresponding RPA calculations with a PBE reference, but do not reach the accuracy of RPA@HF with the employed parametrizations. Because of the high accuracy and the fact that these methods can be implemented with formally quartic scaling using RI, direct RPA and σ-functionals appear promising for the computation of NMR shielding constants. Several lines of further improvement could be pursued in the future, such as including exchange effects in the calculation of the interacting response function or developing new σ-functionals with different parametrizations. Also, self-consistent RPA schemes[109,110] might enable more accurate shielding calculations. We hope that this benchmark study encourages the development of efficient analytical derivative methods for RPA, σ-functionals, and related post-Kohn−Sham methods in order to allow for highly accurate NMR shielding calculations on extended molecular systems.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.1c00866.

> Computed NMR shieldings and statistics for all considered methods and basis sets, diamagnetic and paramagnetic shieldings in a cc-pVTZ basis, test calculations for analyzing gauge origin dependence and the influence of the employed parameters, comparison of analytical and numerical shieldings in a cc-pVTZ basis (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Christian Ochsenfeld − *Chair of Theoretical Chemistry, Department of Chemistry, University of Munich (LMU), D-81377 Munich, Germany; Max Planck Institute for Solid State Research, D-70569 Stuttgart, Germany;* orcid.org/0000-0002-4189-6558; Email: christian.ochsenfeld@uni-muenchen.de

Authors

Michael Glasbrenner − *Chair of Theoretical Chemistry, Department of Chemistry, University of Munich (LMU), D-81377 Munich, Germany*

Daniel Graf − *Chair of Theoretical Chemistry, Department of Chemistry, University of Munich (LMU), D-81377 Munich, Germany*

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jctc.1c00866

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors thank Gökcen Savasci (LMU Munich) for helpful discussions and acknowledge financial support by the "Deutsche Forschungsgemeinschaft" (DFG) via the SFB 1309-32587107 and the cluster of excellence (EXC2111-390814868) "Munich Center for Quantum Science and Technology" (MCQST). M.G. thanks the "Studienstiftung des Deutschen Volkes" for a graduate fellowship. C.O. acknowledges additional financial support as a Max-Planck-Fellow at MPI-FKF Stuttgart.

■ REFERENCES

(1) Helgaker, T.; Jaszuński, M.; Ruud, K. Ab Initio Methods for the Calculation of NMR Shielding and Indirect Spin−Spin Coupling Constants. *Chem. Rev.* **1999**, *99*, 293−352.

(2) Gauss, J.; Stanton, J. F. Electron-correlated approaches for the calculation of NMR chemical shifts. *Adv. Chem. Phys.* **2002**, *123*, 355−422.

(3) Ditchfield, R. Self-consistent perturbation theory of diamagnetism: I. A gauge-invariant LCAO method for NMR chemical shifts. *Mol. Phys.* **1974**, *27*, 789−807.

(4) Kutzelnigg, W. Theory of magnetic susceptibilities and NMR chemical shifts in terms of localized quantities. *Isr. J. Chem.* **1980**, *19*, 193−200.

(5) Schindler, M.; Kutzelnigg, W. Theory of magnetic susceptibilities and NMR chemical shifts in terms of localized quantities. II. Application to some simple molecules. *J. Chem. Phys.* **1982**, *76*, 1919−1933.

(6) Wolinski, K.; Hinton, J. F.; Pulay, P. Efficient implementation of the gauge-independent atomic orbital method for NMR chemical shift calculations. *J. Am. Chem. Soc.* **1990**, *112*, 8251−8260.

(7) Häser, M.; Ahlrichs, R.; Baron, H.; Weis, P.; Horn, H. 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*, 455−470.

(8) Stevens, R. M.; Pitzer, R. M.; Lipscomb, W. N. Perturbed Hartree-Fock Calculations. I. Magnetic Susceptibility and Shielding in the LiH Molecule. *J. Chem. Phys.* **1963**, *38*, 550−560.

(9) Schreckenbach, G.; Ziegler, T. Calculation of NMR shielding tensors using gauge-including atomic orbitals and modern density functional theory. *J. Phys. Chem.* **1995**, *99*, 606−611.

(10) Cheeseman, J. R.; Trucks, G. W.; Keith, T. A.; Frisch, M. J. A comparison of models for calculating nuclear magnetic resonance shielding tensors. *J. Chem. Phys.* **1996**, *104*, 5497−5509.

(11) Rauhut, G.; Puyear, S.; Wolinski, K.; Pulay, P. Comparison of NMR Shieldings Calculated from Hartree−Fock and Density Functional Wave Functions Using Gauge-Including Atomic Orbitals. *J. Phys. Chem.* **1996**, *100*, 6310−6316.

(12) Ochsenfeld, C.; Kussmann, J.; Koziol, F. Ab Initio NMR Spectra for Molecular Systems with a Thousand and More Atoms: A Linear-Scaling Method. *Angew. Chem.* **2004**, *116*, 4585−4589.

(13) 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.

(14) Beer, M.; Kussmann, J.; Ochsenfeld, C. Nuclei-selected NMR shielding calculations: A sublinear-scaling quantum-chemical method. *J. Chem. Phys.* **2011**, *134*, 074102.

(15) Gauss, J.; Stanton, J. F. Gauge-invariant calculation of nuclear magnetic shielding constants at the coupled-cluster singles and doubles level. *J. Chem. Phys.* **1995**, *102*, 251−253.

(16) Gauss, J.; Stanton, J. F. Perturbative treatment of triple excitations in coupled-cluster calculations of nuclear magnetic shielding constants. *J. Chem. Phys.* **1996**, *104*, 2574−2583.

(17) Kállay, M.; Gauss, J. Analytic second derivatives for general coupled-cluster and configuration-interaction models. *J. Chem. Phys.* **2004**, *120*, 6841−6848.

(18) Hoffmann, M. R.; Fox, D. J.; Gaw, J. F.; Osamura, Y.; Yamaguchi, Y.; Grev, R. S.; Fitzgerald, G.; Schaefer, H. F., III; Knowles, P. J.; Handy, N. C. Analytic energy second derivatives for general MCSCF wave functions. *J. Chem. Phys.* **1984**, *80*, 2660−2668.

(19) Ruud, K.; Helgaker, T.; Kobayashi, R.; Jørgensen, P.; Bak, K. L.; Jensen, H. J. A. Multiconfigurational self-consistent field calculations of nuclear shieldings using London atomic orbitals. *J. Chem. Phys.* **1994**, *100*, 8178−8185.

(20) Gauss, J. Calculation of NMR chemical shifts at second-order many-body perturbation theory using gauge-including atomic orbitals. *Chem. Phys. Lett.* **1992**, *191*, 614−620.

(21) Gauss, J. Effects of electron correlation in the calculation of nuclear magnetic resonance chemical shifts. *J. Chem. Phys.* **1993**, *99*, 3629−3643.

(22) 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.

(23) Yamaguchi, Y.; Osamura, Y.; Goddard, J. D.; Schaefer, H. F., III *A New Dimension to Quantum Chemistry: Analytic Derivative Methods in Ab Initio Molecular Electronic Structure Theory*; Oxford University Press: USA, 1994.

(24) Kollwitz, M.; Gauss, J. A direct implementation of the GIAO-MBPT(2) method for calculating NMR chemical shifts. Application to the naphthalenium and anthracenium ions. *Chem. Phys. Lett.* **1996**, *260*, 639−646.

(25) Gauss, J.; Werner, H.-J. NMR chemical shift calculations within local correlation methods: the GIAO-LMP2 approach. *Phys. Chem. Chem. Phys.* **2000**, *2*, 2083−2090.

(26) Loibl, S.; Schütz, M. 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) Stoychev, G. L.; Auer, A. A.; Gauss, J.; Neese, F. DLPNO-MP2 second derivatives for the computation of polarizabilities and NMR shieldings. *J. Chem. Phys.* **2021**, *154*, 164110.

(28) Burger, S.; Lipparini, F.; Gauss, J.; Stopkowicz, S. NMR chemical shift computations at second-order Møller-Plesset perturbation theory using gauge-including atomic orbitals and Cholesky-decomposed two-electron integrals. *J. Chem. Phys.* **2021**, *155*, 074105.

(29) 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.

(30) 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.

(31) 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*, 9095−9102.

(32) 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*, 9793−9802.

(33) Stoychev, G. L.; Auer, A. A.; Neese, F. Efficient and accurate prediction of nuclear magnetic resonance shielding tensors with double-hybrid density functional theory. *J. Chem. Theory Comput.* **2018**, *14*, 4756−4771.

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

(35) Langreth, D. C.; Perdew, J. P. The exchange-correlation energy of a metallic surface. *Solid State Commun.* **1975**, *17*, 1425−1429.

(36) Langreth, D. C.; Perdew, J. P. Exchange-correlation energy of a metallic surface: Wave-vector analysis. *Phys. Rev. B: Solid State* **1977**, *15*, 2884.

(37) Ángyán, J. G.; Liu, R.-F.; Toulouse, J.; Jansen, G. Correlation Energy Expressions from the Adiabatic-Connection Fluctuation-Dissipation Theorem Approach. *J. Chem. Theory Comput.* **2011**, *7*, 3116−3130.

(38) Furche, F. Molecular tests of the random phase approximation to the exchange-correlation energy functional. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2001**, *64*, 195120.

(39) Furche, F. Developing the random phase approximation into a practical post-Kohn-Sham correlation model. *J. Chem. Phys.* **2008**, *129*, 114105.

(40) Bohm, D.; Pines, D. A collective description of electron interactions: III. Coulomb interactions in a degenerate electron gas. *Phys. Rev.* **1953**, *92*, 609.

(41) Eshuis, H.; Yarkony, J.; Furche, F. Fast computation of molecular random phase approximation correlation energies using resolution of the identity and imaginary frequency integration. *J. Chem. Phys.* **2010**, *132*, 234114.

(42) 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.

(43) Almlöf, J. Elimination of energy denominators in Møller-Plesset perturbation theory by a Laplace transform approach. *Chem. Phys. Lett.* **1991**, *181*, 319−320.

(44) Kállay, M. Linear-scaling implementation of the direct random-phase approximation. *J. Chem. Phys.* **2015**, *142*, 204105.

(45) Schurkus, H. F.; Ochsenfeld, C. Communication: An effective linear-scaling atomic-orbital reformulation of the random-phase approximation using a contracted double-Laplace transformation. *J. Chem. Phys.* **2016**, *144*, 031101.

(46) 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.

(47) Graf, D.; Beuerle, M.; Schurkus, H. F.; Luenser, A.; Savasci, G.; Ochsenfeld, C. Accurate and Efficient Parallel Implementation of an Effective Linear-Scaling Direct Random Phase Approximation Method. *J. Chem. Theory Comput.* **2018**, *14*, 2505−2515.

(48) Freeman, D. L. Coupled-cluster expansion applied to the electron gas: Inclusion of ring and exchange effects. *Phys. Rev. B: Solid State* **1977**, *15*, 5512.

(49) Mussard, B.; Rocca, D.; Jansen, G.; Ángyán, J. G. Dielectric matrix formulation of correlation energies in the random phase approximation: inclusion of exchange effects. *J. Chem. Theory Comput.* **2016**, *12*, 2191−2202.

(50) Dixit, A.; Ángyán, J. G.; Rocca, D. Improving the accuracy of ground-state correlation energies within a plane-wave basis set: The electron-hole exchange kernel. *J. Chem. Phys.* **2016**, *145*, 104105.

(51) Grüneis, A.; Marsman, M.; Harl, J.; Schimka, L.; Kresse, G. Making the random phase approximation to electronic correlation accurate. *J. Chem. Phys.* **2009**, *131*, 154115.

(52) Heßelmann, A.; Görling, A. Random phase approximation correlation energies with exact Kohn-Sham exchange. *Mol. Phys.* **2010**, *108*, 359−372.

(53) Bates, J. E.; Furche, F. Communication: Random phase approximation renormalized many-body perturbation theory. *J. Chem. Phys.* **2013**, *139*, 171103.

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

(55) Trushin, E.; Thierbach, A.; Görling, A. Toward chemical accuracy at low computational cost: Density-functional theory with σ-functionals for the correlation energy. *J. Chem. Phys.* **2021**, *154*, 014104.

(56) Fauser, S.; Trushin, E.; Neiss, C.; Görling, A. Chemical accuracy with σ-functionals for the Kohn-Sham correlation energy optimized for different input orbitals and eigenvalues. *J. Chem. Phys.* **2021**, *155*, 134111.

(57) Hansen, A. E.; Bouman, T. D. Localized orbital/local origin method for calculation and analysis of NMR shieldings. Applications to13C shielding tensors. *J. Chem. Phys.* **1985**, *82*, 5035−5047.

(58) Visscher, L.; Enevoldsen, T.; Saue, T.; Jensen, H. J. A.; Oddershede, J. Full four-component relativistic calculations of NMR shielding and indirect spin-spin coupling tensors in hydrogen halides. *J. Comput. Chem.* **1999**, *20*, 1262−1273.

(59) Sauer, S. P. A.; Špirko, V.; Paidarová, I.; Kraemer, W. P. The vibrational dependence of the hydrogen and oxygen nuclear magnetic shielding constants in OH− and OH− · H2O. *Chem. Phys.* **1997**, *214*, 91−101.

(60) Gimenez, C. A.; Kozioł, K.; Aucar, G. A. Quantum electrodynamics effects on NMR magnetic shielding constants of He-like and Be-like atomic systems. *Phys. Rev. A* **2016**, *93*, 032504.

(61) Tossell, J. A.; Lazzeretti, P. Ab initio calculation of 77Se NMR shielding constants in H2Se and SeF6 and other second-order properties of H2Se. *J. Magn. Reson.* **1988**, *80*, 39−44.

(62) Tossell, J. A. Calculation of the 29 Si NMR shielding tensor in forsterite. *Phys. Chem. Miner.* **1992**, *19*, 338−342.

(63) Teale, A. M.; Lutnæs, O. B.; Helgaker, T.; Tozer, D. J.; Gauss, J. Benchmarking density-functional theory calculations of NMR shielding constants and spin-rotation constants using accurate coupled-cluster calculations. *J. Chem. Phys.* **2013**, *138*, 024111.

(64) Lutnæs, O. B.; Teale, A. M.; Helgaker, T.; Tozer, D. J.; Ruud, K.; Gauss, J. Benchmarking density-functional-theory calculations of rotational g tensors and magnetizabilities using accurate coupled-cluster calculations. *J. Chem. Phys.* **2009**, *131*, 144104.

(65) Gauss, J. *Molecular Properties. Modern Methods and Algorithms of Quantum Chemistry*; Jülich, 2000; pp 541−592.

(66) Fukui, H.; Baba, T.; Matsuda, H.; Miura, K. Calculation of nuclear magnetic shieldings. IX. Electron correlation effects. *J. Chem. Phys.* **1994**, *100*, 6608−6613.

(67) Tellgren, E. I.; Soncini, A.; Helgaker, T. Nonperturbative ab initio calculations in strong magnetic fields using London orbitals. *J. Chem. Phys.* **2008**, *129*, 154114.

https://doi.org/10.1021/acs.jctc.1c00866

(68) Tellgren, E. I.; Teale, A. M.; Furness, J. W.; Lange, K. K.; Ekström, U.; Helgaker, T. Non-perturbative calculation of molecular magnetic properties within current-density functional theory. *J. Chem. Phys.* **2014**, *140*, 034101.

(69) London, F. Théorie quantique des courants interatomiques dans les combinaisons aromatiques. *J. Phys. Radium* **1937**, *8*, 397−409.

(70) Hameka, H. F. On the nuclear magnetic shielding in the hydrogen molecule. *Mol. Phys.* **1958**, *1*, 203−215.

(71) Hameka, H. F. Theory of magnetic properties of molecules with particular emphasis on the hydrogen molecule. *Rev. Mod. Phys.* **1962**, *34*, 87.

(72) Helgaker, T.; Jørgensen, P. An electronic Hamiltonian for origin independent calculations of magnetic properties. *J. Chem. Phys.* **1991**, *95*, 2595−2601.

(73) Vignale, G.; Rasolt, M. Density-functional theory in strong magnetic fields. *Phys. Rev. Lett.* **1987**, *59*, 2360.

(74) Grayce, C. J.; Harris, R. A. Magnetic-field density-functional theory. *Phys. Rev. A: At., Mol., Opt. Phys.* **1994**, *50*, 3089.

(75) Zhu, W.; Zhang, L.; Trickey, S. B. Random phase approximation with second-order screened exchange for current-carrying atomic states. *J. Chem. Phys.* **2016**, *145*, 224106.

(76) Niquet, Y.; Fuchs, M.; Gonze, X. Exchange-correlation potentials in the adiabatic connection fluctuation-dissipation framework. *Phys. Rev. A: At., Mol., Opt. Phys.* **2003**, *68*, 032507.

(77) Grundei, M. M. J.; Burow, A. M. Random phase approximation for periodic systems employing direct coulomb lattice summation. *J. Chem. Theory Comput.* **2017**, *13*, 1159−1175.

(78) Kussmann, J.; Ochsenfeld, C. Pre-selective screening for matrix elements in linear-scaling exact exchange calculations. *J. Chem. Phys.* **2013**, *138*, 134114.

(79) 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.

(80) Kussmann, J.; Ochsenfeld, C. Hybrid CPU/GPU integral engine for strong-scaling ab initio methods. *J. Chem. Theory Comput.* **2017**, *13*, 3153−3159.

(81) Laqua, H.; Kussmann, J.; Ochsenfeld, C. An improved molecular partitioning scheme for numerical quadratures in density functional theory. *J. Chem. Phys.* **2018**, *149*, 204−111.

(82) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **1996**, *77*, 3865.

(83) Keal, T. W.; Tozer, D. J. The exchange-correlation potential in Kohn-Sham nuclear magnetic resonance shielding calculations. *J. Chem. Phys.* **2003**, *119*, 3015−3024.

(84) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields. *J. Phys. Chem.* **1994**, *98*, 11623−11627.

(85) Wilson, P. J.; Bradley, T. J.; Tozer, D. J. Hybrid exchange-correlation functional determined from thermochemical data and ab initio potentials. *J. Chem. Phys.* **2001**, *115*, 9233−9242.

(86) Becke, A. D. A new mixing of Hartree-Fock and local density-functional theories. *J. Chem. Phys.* **1993**, *98*, 1372−1377.

(87) Dunning, T. H., Jr. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007−1023.

(88) Woon, D. E.; Dunning, T. H., Jr. Gaussian basis sets for use in correlated molecular calculations. III. The atoms aluminum through argon. *J. Chem. Phys.* **1993**, *98*, 1358−1371.

(89) Woon, D. E.; Dunning, T. H., Jr. Gaussian basis sets for use in correlated molecular calculations. V. Core-valence basis sets for boron through neon. *J. Chem. Phys.* **1995**, *103*, 4572−4585.

(90) Peterson, K. A.; Dunning, T. H., Jr. Accurate correlation consistent basis sets for molecular core-valence correlation effects: The second row atoms Al-Ar, and the first row atoms B-Ne revisited. *J. Chem. Phys.* **2002**, *117*, 10548−10560.

(91) Hättig, C. Optimization of auxiliary basis sets for RI-MP2 and RI-CC2 calculations: Core-valence and quintuple-ζ basis sets for H to Ar and QZVPP basis sets for Li to Kr. *Phys. Chem. Chem. Phys.* **2005**, *7*, 59−66.

(92) Weigend, F.; Köhn, A.; Hättig, C. Efficient use of the correlation consistent basis sets in resolution of the identity MP2 calculations. *J. Chem. Phys.* **2002**, *116*, 3175−3183.

(93) Weigend, F.; Ahlrichs, R. Balanced basis sets 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.

(94) 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.

(95) Jensen, F. Unifying general and segmented contracted basis sets. Segmented polarization consistent basis sets. *J. Chem. Theory Comput.* **2014**, *10*, 1074−1085.

(96) Pritchard, B. P.; Altarawy, D.; Didier, B.; Gibson, T. D.; Windus, T. L. New basis set exchange: An open, up-to-date resource for the molecular sciences community. *J. Chem. Inf. Model.* **2019**, *59*, 4814−4820.

(97) Matthews, D. A.; Cheng, L.; Harding, M. E.; Lipparini, F.; Stopkowicz, S.; Jagau, T.-C.; Szalay, P. G.; Gauss, J.; Stanton, J. F. Coupled-cluster techniques for computational chemistry: The CFOUR program package. *J. Chem. Phys.* **2020**, *152*, 214108.

(98) Stanton, J. F.; Gauss, J.; Cheng, L.; Harding, M. E.; Matthews, D. A.; Szalay, P. G. *CFOUR, Coupled-Cluster Techniques for Computational Chemistry, a Quantum-Chemical Program Package.* With contributions from A. A. Auer, R. J. Bartlett, U. Benedikt, C. Berger, D. E. Bernholdt, S. Blaschke, Y. J. Bomble, S. Burger, O. Christiansen, D. Datta, F. Engel, R. Faber, J. Greiner, M. Heckert, O. Heun, M. Hilgenberg, C. Huber, T.-C. Jagau, D. Jonsson, J. Jusélius, T. Kirsch, K. Klein, G.M. KopperW. J. Lauderdale, F. Lipparini, T. Metzroth, L. A. Mück, D. P. O'Neill, T. Nottoli, 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.

(99) *TURBOMOLE V7.4 2019*, a Development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989−2007, TURBOMOLE GmbH, since 2007; available from http://www.turbomole.com.

(100) Kollwitz, M.; Häser, M.; Gauss, J. Non-Abelian point group symmetry in direct second-order many-body perturbation theory calculations of NMR chemical shifts. *J. Chem. Phys.* **1998**, *108*, 8295−8301.

(101) Reimann, S.; Ekström, U.; Stopkowicz, S.; Teale, A. M.; Borgoo, A.; Helgaker, T. The importance of current contributions to shielding constants in density-functional theory. *Phys. Chem. Chem. Phys.* **2015**, *17*, 18834−18842.

(102) Teale, A. M.; Tozer, D. J. Exchange representations in Kohn-Sham NMR shielding calculations. *Chem. Phys. Lett.* **2004**, *383*, 109−114.

(103) Tellgren, E. I. Density-functional theory for internal magnetic fields. *Phys. Rev. A* **2018**, *97*, 012504.

(104) Prado, F. R.; Giessner-Prettre, C.; Daudey, J.-P.; Pullman, A.; Hinton, J. F.; Young, G.; Harpool, D. Nuclear magnetic resonance spectroscopy applied to Li+ complexation by small ligands: An ab initio and experimental study of 7Li chemical shifts. *J. Magn. Reson.* **1980**, *37*, 431−440.

(105) Fukui, H.; Miura, K.; Shinbori, H. Calculation of NMR chemical shifts. VI. Gauge invariant and Hermitian condition. *J. Chem. Phys.* **1985**, *83*, 907−908.

(106) Burow, A. M.; Bates, J. E.; Furche, F.; Eshuis, H. Analytical first-order molecular properties and forces within the adiabatic connection random phase approximation. *J. Chem. Theory Comput.* **2014**, *10*, 180−194.

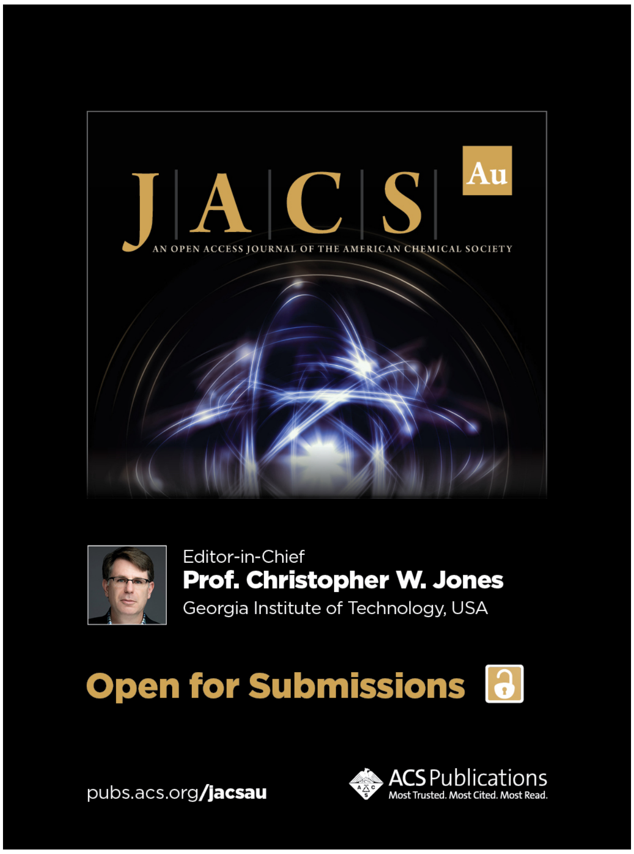
https://doi.org/10.1021/acs.jctc.1c00866

(107) Scuseria, G. E.; Henderson, T. M.; Sorensen, D. C. The ground state correlation energy of the random phase approximation from a ring coupled cluster doubles approach. *J. Chem. Phys.* **2008**, *129*, 231101.

(108) Flaig, D.; Maurer, M.; Hanni, M.; Braunger, K.; Kick, L.; Thubauville, M.; Ochsenfeld, C. Benchmarking hydrogen and carbon NMR chemical shifts at HF, DFT, and MP2 levels. *J. Chem. Theory Comput.* **2014**, *10*, 572−578.

(109) Graf, D.; Beuerle, M.; Ochsenfeld, C. Low-scaling self-consistent minimization of a density matrix based random phase approximation method in the atomic orbital space. *J. Chem. Theory Comput.* **2019**, *15*, 4468−4477.

(110) Yu, J. M.; Nguyen, B. D.; Tsai, J.; Hernandez, D. J.; Furche, F. Selfconsistent random phase approximation methods. *J. Chem. Phys.* **2021**, *155*, 040902.



Supplementary Material for: Benchmarking the accuracy of the direct random phase approximation and $\boldsymbol{\sigma}$-functionals for NMR shieldings

Michael Glasbrenner[1], Daniel Graf[1], and Christian Ochsenfeld[1,2]

[1]Chair of Theoretical Chemistry, Department of Chemistry, University of Munich (LMU), Butenandtstr. 7, 81377 Munich, Germany
[2]Max Planck Institute for Solid State Research, D-70569 Stuttgart, Germany

November 7, 2021

List of Figures and Tables

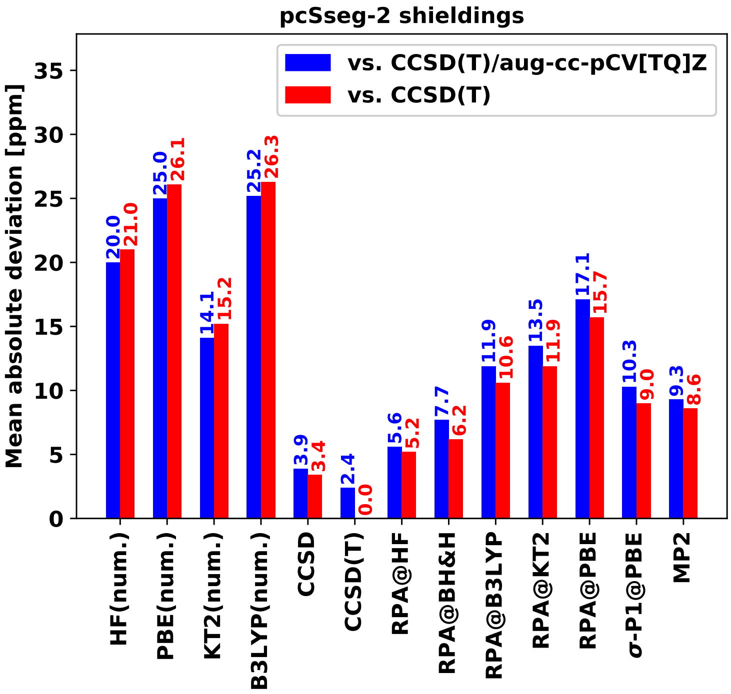


Figure 1: Mean absolute deviations of NMR-shieldings on the employed test set using the pcSseg-2 basis set in all calculations. The blue bars display the deviation from basis-set extrapolated (aug-cc-pCV[TQ]Z) CCSD(T)-shieldings; the red bars display deviations from CCSD(T)-shieldings obtained with a pcSseg-2 basis set. "(num.)" denotes shieldings computed with a numerical derivative.

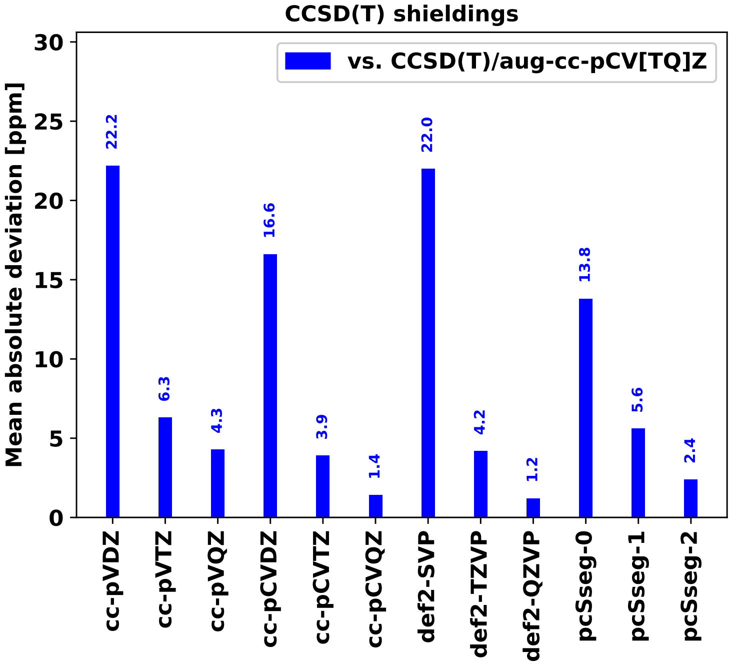


Figure 2: Mean absolute deviations of CCSD(T) shieldings using different basis sets. Basis-set extrapolated (aug-cc-pCV[TQ]Z) CCSD(T)-shieldings are used as a reference. All CCSD(T) shielding calculations were done with the CFOUR program[1, 2].

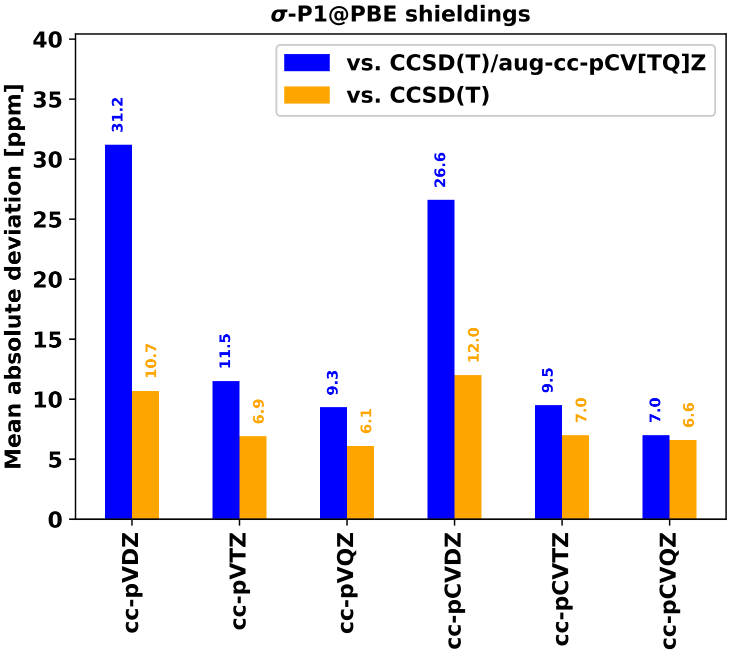


Figure 3: Mean absolute deviations of σ-P1@PBE shieldings using different basis sets. Basis-set extrapolated (aug-cc-pCV[TQ]Z) CCSD(T)-shieldings and CCSD(T) with the same respective basis set are used as references. All CCSD(T) shielding calculations were done with the CFOUR program[1, 2].

Table 1: Absolute isotropic NMR shieldings in a cc-pVDZ basis. All values are given in ppm. “Extrap.” stands for basis-set extrapolated CCSD(T)-shieldings in an aug-cc-pCV[TQ]Z basis set. The data for “Extrap.” was taken from Ref. 3. CCSD(T) calculations were done with the CFOUR program[1, 2].

Mol.	Nuc.	CCSD(T)	Extrap.	RPA@HF	RPA@PBE	σ-P1@PBE
HF	H	30.5	28.8	30.4	31.1	31.4
	F	427.0	420.3	427.2	425.1	422.0
CO	C	34.3	2.2	32.2	55.8	46.9
	O	-16.5	-55.0	-10.2	18.6	0.1
N_2	N	-25.6	-60.4	-25.4	3.7	-11.8
H_2O	O	354.9	338.0	353.8	351.8	347.9
	H	31.9	30.6	31.9	32.3	32.5
HCN	H	29.4	29.0	29.6	29.6	29.7
	C	111.5	84.6	108.7	120.1	117.4
	N	21.5	-14.1	20.3	45.0	37.9
HOF	O	-5.6	-68.9	9.1	76.6	43.1
	H	21.8	19.6	22.6	24.5	23.8
	F	183.1	192.2	206.6	222.0	213.2
NH_3	N	283.9	270.7	281.8	281.3	279.0
	H	32.3	31.4	32.4	32.5	32.6
CH_2O	O	-339.3	-378.6	-309.5	-204.3	-259.4
	C	36.9	1.5	37.1	59.3	54.2
	H	22.7	22.0	23.1	23.7	23.4
CH_4	C	209.6	198.9	207.6	208.9	207.2
	H	31.6	31.3	31.7	31.7	31.7
C2H4	C	95.6	69.7	92.3	103.4	102.7
	H	26.6	26.0	26.7	27.0	27.0
AlF	Al	599.3	572.9	604.1	610.6	607.0
	F	190.9	211.8	212.3	207.8	204.7
CH_3F	C	142.9	122.1	142.9	147.2	146.0
	F	475.9	482.9	476.7	470.5	471.0
	H	27.9	27.3	28.1	28.2	28.3
C_3H_4	C	205.0	192.1	203.9	205.6	205.4
	C	107.5	83.7	105.6	115.5	113.9
	H	25.1	24.4	25.2	25.6	25.5
	H	31.2	30.6	31.3	31.3	31.4
FCCH	C	192.5	179.9	190.7	194.5	195.6
	C	125.5	100.1	125.0	129.5	129.2
	H	31.0	30.5	31.1	31.2	31.3
	F	429.6	423.5	433.1	431.2	431.9
FCN	F	385.1	374.1	392.3	391.3	387.2
	C	110.7	82.2	108.5	116.0	114.6
	N	140.5	117.9	137.8	153.0	151.2
H_2S	S	783.0	739.0	776.4	785.2	768.2
	H	31.4	30.5	31.5	31.7	31.9
HCP	H	29.4	29.6	29.7	29.6	29.7
	C	67.3	37.6	61.2	78.4	76.6
	P	471.7	388.0	456.1	510.3	499.6
HFCO	O	-58.6	-94.3	-48.5	-15.6	-33.6
	C	68.0	39.6	66.9	76.7	74.3
	F	187.0	165.3	204.5	207.1	196.6
	H	24.7	23.9	24.9	25.0	24.9
H_2C_2O	C	205.3	193.3	204.4	205.8	205.2
	C	26.1	-6.3	26.5	37.5	31.0
	O	22.2	-5.9	36.3	51.0	35.1
	H	30.0	29.2	30.2	30.3	30.4
LiF	Li	94.4	89.3	94.5	94.6	94.6
	F	350.0	382.5	359.8	352.2	351.4
LiH	H	26.3	26.6	26.5	26.2	26.3
	Li	92.4	89.3	92.4	93.7	93.6
N_2O	N	129.6	106.4	127.9	144.2	139.8
	N	45.4	12.6	41.0	60.3	55.0
	O	209.8	199.0	215.9	223.7	217.2
OCS	O	123.0	96.8	128.1	140.9	132.8
	C	58.3	30.2	54.6	69.5	66.4
	S	840.0	796.7	845.6	857.4	848.2
OF_2	O	-339.7	-447.1	-306.1	-130.0	-229.2
	F	18.7	-24.0	49.5	79.9	58.4
H_4C_2O	O	364.8	363.2	369.5	364.7	364.2
	C	170.7	153.2	170.6	173.5	173.0
	H	29.7	29.1	30.0	30.0	30.0
PN	N	-282.0	-344.0	-273.9	-192.3	-245.0
	P	189.0	50.6	176.9	302.7	224.5

Table 2: Statistics for absolute isotropic NMR shieldings in a cc-pVDZ basis. CCSD(T)/aug-cc-pCV[TQ]Z shieldings are used as a reference. MSD=mean signed deviation; MAD=mean absolute deviation; MRD=mean relative deviation [%]; MARD=mean absolute relative deviation [%]; RMSD=root mean squared deviation.

	MSD	MAD	MRD	MARD	RMSD
RHF	1.6	15.8	8.9	39.5	28.7
CCSD	20.0	21.3	92.4	92.8	32.4
CCSD(T)	20.2	22.2	98.7	99.5	33.4
RPA@HF	22.9	23.7	103.0	103.2	36.4
RPA@PBE	37.9	39.3	165.9	166.3	69.2
σ-P1@PBE	29.7	31.2	140.8	141.2	51.0
MP2	28.6	30.5	126.1	126.9	48.3

Table 3: Statistics for absolute isotropic NMR shieldings in a cc-pVDZ basis. CCSD(T) shieldings are used as a reference. MSD=mean signed deviation; MAD=mean absolute deviation; MRD=mean relative deviation [%]; MARD=mean absolute relative deviation [%]; RMSD=root mean squared deviation.

	MSD	MAD	MRD	MARD	RMSD
RHF	-18.5	24.8	-47.3	54.5	46.6
CCSD	-0.1	2.8	-0.1	3.9	5.4
CCSD(T)	0.0	0.0	0.0	0.0	0.0
RPA@HF	2.7	5.0	8.3	10.3	9.2
RPA@PBE	17.7	18.1	44.7	44.8	39.5
σ-P1@PBE	9.6	10.7	26.2	26.5	21.2
MP2	8.4	9.0	16.8	17.3	19.1

Table 4: Absolute isotropic NMR shieldings in a cc-pVTZ basis. All values are given in ppm. “Extrap.” stands for basis-set extrapolated CCSD(T)-shieldings in an aug-cc-pCV[TQ]Z basis set. The data for “Extrap.” was taken from Ref. 3. CCSD(T) calculations were done with the CFOUR program[1, 2].

Mol.	Nuc.	CCSD(T)	Extrap.	RPA@HF	RPA@PBE	RPA@KT2	RPA@B3LYP	RPA@BH&H	RPA@B97-2
HF	H	29.5	28.8	29.4	30.1	30.1	29.9	29.7	29.9
	F	419.8	420.3	419.0	417.9	418.2	418.3	418.6	418.0
CO	C	11.3	2.2	5.6	28.4	23.6	21.5	14.3	19.7
	O	-49.8	-55.0	-48.4	-21.4	-26.2	-29.7	-37.8	-31.7
N_2	N	-52.5	-60.4	-58.2	-28.8	-34.7	-37.8	-47.3	-39.8
H_2O	O	344.3	338.0	341.7	341.1	341.6	341.5	341.5	341.0
	H	31.3	30.6	31.3	31.7	31.7	31.6	31.5	31.6
HCN	H	29.2	29.0	29.5	29.5	29.5	29.5	29.5	29.5
	C	91.5	84.6	87.5	98.7	96.4	95.7	92.2	94.8
	N	-7.4	-14.1	-11.8	12.5	6.6	6.5	-1.3	3.8
HOF	O	-44.4	-68.9	-40.6	24.1	0.4	-5.2	-26.7	-10.5
	H	20.7	19.6	21.3	22.9	22.3	22.0	21.5	22.0
	F	188.3	192.2	212.3	213.7	210.5	210.2	209.9	210.0
NH_3	N	276.2	270.7	272.8	273.0	273.3	273.3	273.0	272.6
	H	32.0	31.4	32.1	32.2	32.2	32.2	32.2	32.2
CH_2O	O	-376.5	-378.6	-351.6	-266.4	-292.3	-293.5	-316.4	-302.2
	C	13.2	1.5	12.3	31.6	25.4	24.8	18.6	22.9
	H	22.3	22.0	22.9	23.3	23.1	23.0	22.9	23.0
CH_4	C	201.3	198.9	198.4	199.3	199.3	199.6	198.9	198.7
	H	31.5	31.3	31.6	31.6	31.6	31.6	31.6	31.6
C2H4	C	77.3	69.7	73.3	83.0	80.7	80.8	77.6	79.3
	H	26.3	26.0	26.5	26.7	26.7	26.7	26.6	26.7
AlF	Al	576.5	572.9	580.9	586.1	583.4	583.3	580.7	581.7
	F	200.2	211.8	217.9	217.6	215.6	216.3	216.0	214.9
CH_3F	C	127.4	122.1	127.3	130.0	129.0	129.5	128.1	128.3
	F	479.3	482.9	477.4	472.0	472.5	473.3	474.5	473.4
	H	27.7	27.3	28.0	28.0	28.0	28.0	27.9	27.9
C_3H_4	C	194.9	192.1	193.2	194.3	193.6	194.4	193.5	193.3
	C	90.3	83.7	87.0	96.4	94.1	94.2	90.8	92.5
	H	24.6	24.4	24.8	25.2	25.1	25.1	25.0	25.1
	H	30.9	30.6	31.1	31.1	31.1	31.1	31.1	31.1
FCCH	C	183.1	179.9	180.7	183.8	182.6	183.7	182.7	183.0
	C	105.3	100.1	104.8	108.4	107.3	107.9	106.4	107.0
	H	30.7	30.5	30.9	31.0	31.0	31.0	30.9	31.0
	F	421.2	423.5	423.2	421.4	421.3	421.8	422.5	422.3
FCN	F	373.2	374.1	377.9	376.7	376.8	376.4	376.7	376.7
	C	88.2	82.2	85.4	92.4	90.8	91.3	88.8	90.3
	N	122.4	117.9	116.7	131.5	128.1	129.6	125.1	127.6
H_2S	S	758.5	739.0	744.8	753.3	752.2	752.1	747.7	748.9
	H	30.9	30.5	31.1	31.3	31.3	31.2	31.2	31.2
HCP	H	29.7	29.6	30.0	29.9	30.0	29.9	30.0	30.0
	C	45.4	37.6	38.2	54.1	50.9	50.5	45.7	48.6
	P	412.9	388.0	394.9	446.8	435.9	434.9	416.8	428.4
HFCO	O	-88.0	-94.3	-82.9	-54.6	-59.0	-58.7	-65.9	-61.9
	C	47.0	39.6	45.0	53.3	51.3	51.7	48.8	50.6
	F	167.2	165.3	184.0	184.5	180.3	183.2	181.6	181.9
	H	24.3	23.9	24.6	24.6	24.6	24.6	24.6	24.6
H_2C_2O	C	196.3	193.3	194.1	195.4	195.3	195.6	194.9	194.9
	C	3.4	-6.3	1.8	11.8	10.3	9.9	6.6	8.7
	O	-3.9	-5.9	4.1	18.2	17.4	16.7	13.5	14.9
	H	29.5	29.2	29.7	29.9	29.9	29.8	29.8	29.8
LiF	Li	90.6	89.3	90.9	90.9	90.5	90.8	90.8	90.7
	F	362.8	382.5	371.1	369.7	362.9	371.0	370.6	367.9
LiH	H	26.5	26.6	26.8	26.5	26.5	26.6	26.8	26.6
	Li	90.0	89.3	89.9	90.9	90.6	90.5	90.2	90.3
N_2O	N	114.3	106.4	107.8	125.4	122.9	121.8	116.2	120.0
	N	22.1	12.6	13.9	34.5	32.3	29.8	23.3	28.1
	O	197.0	199.0	197.7	207.2	206.4	205.3	202.6	204.5
OCS	O	98.7	96.8	100.6	112.6	111.3	110.1	106.6	108.9
	C	37.9	30.2	32.2	47.0	45.2	43.9	39.3	42.6
	S	810.9	796.7	811.1	823.4	820.4	822.1	816.6	818.7
OF_2	O	-402.8	-447.1	-377.4	-230.0	-300.5	-309.3	-356.8	-320.7
	F	0.0	-24.0	27.2	49.3	42.4	38.6	31.1	36.0
H_4C_2O	O	362.1	363.2	364.7	359.0	357.7	361.5	362.9	360.7
	C	158.1	153.2	157.5	159.2	158.3	159.2	158.0	158.0
	H	29.4	29.1	29.7	29.7	29.6	29.6	29.7	29.6
PN	N	-336.0	-344.0	-348.5	-268.5	-286.4	-293.1	-317.7	-301.4
	P	89.4	50.6	58.4	177.5	152.2	137.8	100.8	129.9

Table 5: Absolute isotropic NMR shieldings in a cc-pVTZ basis. All values are given in ppm. “Extrap.” stands for basis-set extrapolated CCSD(T)-shieldings in an aug-cc-pCV[TQ]Z basis set. The data for “Extrap.” was taken from Ref. 3. CCSD(T) calculations were done with the CFOUR program[1, 2]. MP2 calculations were done with the Turbomole program[4–6].

Mol.	Nuc.	CCSD(T)	Extrap.	σ-P1@PBE	σ-P2@PBE	σ-P3@PBE	MP2
HF	H	29.5	28.8	30.0	30.0	29.9	29.2
	F	419.8	420.3	415.7	415.7	416.1	425.2
CO	C	11.3	2.2	18.4	18.2	18.8	16.8
	O	-49.8	-55.0	-36.5	-36.6	-34.4	-43.3
N_2	N	-52.5	-60.4	-43.0	-43.5	-42.2	-35.8
H_2O	O	344.3	338.0	338.5	338.6	338.9	352.5
	H	31.3	30.6	31.6	31.6	31.5	31.0
HCN	H	29.2	29.0	29.5	29.6	29.6	29.2
	C	91.5	84.6	94.8	94.5	95.1	93.0
	N	-7.4	-14.1	3.5	3.1	3.4	5.2
HOF	O	-44.4	-68.9	-19.6	-19.4	-14.0	-28.6
	H	20.7	19.6	21.8	21.8	21.9	20.8
	F	188.3	192.2	190.4	191.7	193.3	191.4
NH_3	N	276.2	270.7	271.3	271.4	271.6	281.8
	H	32.0	31.4	32.1	32.1	32.1	31.8
CH_2O	O	-376.5	-378.6	-333.6	-333.2	-325.9	-335.0
	C	13.2	1.5	24.1	24.0	24.4	13.5
	H	22.3	22.0	23.0	23.0	23.0	22.3
CH_4	C	201.3	198.9	198.4	198.4	198.6	203.0
	H	31.5	31.3	31.6	31.6	31.6	31.4
C2H4	C	77.3	69.7	80.8	80.9	80.9	75.9
	H	26.3	26.0	26.7	26.7	26.7	26.2
AlF	Al	576.5	572.9	583.0	582.9	583.4	575.8
	F	200.2	211.8	214.0	213.7	214.7	197.8
CH_3F	C	127.4	122.1	128.8	128.8	129.0	126.0
	F	479.3	482.9	475.4	475.4	475.4	485.8
	H	27.7	27.3	28.1	28.1	28.1	27.6
C_3H_4	C	194.9	192.1	194.3	194.3	194.4	195.8
	C	90.3	83.7	94.1	94.0	93.9	91.6
	H	24.6	24.4	25.1	25.1	25.1	24.6
	H	30.9	30.6	31.2	31.2	31.2	30.9
FCCH	C	183.1	179.9	182.5	182.6	182.5	185.0
	C	105.3	100.1	107.6	107.6	107.6	102.8
	H	30.7	30.5	31.0	31.0	31.0	30.7
	F	421.2	423.5	421.7	421.7	421.8	427.5
FCN	F	373.2	374.1	373.6	373.7	374.0	382.0
	C	88.2	82.2	90.5	90.3	90.3	87.2
	N	122.4	117.9	125.6	125.4	125.6	135.3
H_2S	S	758.5	739.0	739.5	739.6	739.8	774.0
	H	30.9	30.5	31.4	31.4	31.3	30.7
HCP	H	29.7	29.6	30.1	30.1	30.1	29.6
	C	45.4	37.6	52.5	51.6	52.2	50.2
	P	412.9	388.0	436.5	434.5	435.3	419.3
HFCO	O	-88.0	-94.3	-76.7	-76.3	-73.4	-57.8
	C	47.0	39.6	49.9	49.9	50.2	47.4
	F	167.2	165.3	174.4	174.2	175.5	167.4
	H	24.3	23.9	24.6	24.6	24.6	24.2
H_2C_2O	C	196.3	193.3	194.0	194.1	194.0	201.1
	C	3.4	-6.3	5.3	5.3	5.5	7.5
	O	-3.9	-5.9	3.1	2.9	5.0	29.7
	H	29.5	29.2	29.8	29.8	29.8	29.6
LiF	Li	90.6	89.3	90.8	90.8	90.8	90.5
	F	362.8	382.5	369.7	369.4	369.6	360.2
LiH	H	26.5	26.6	26.7	26.7	26.7	26.6
	Li	90.0	89.3	90.6	90.6	90.6	90.0
N_2O	N	114.3	106.4	118.0	117.4	118.2	137.1
	N	22.1	12.6	29.7	28.7	29.5	40.0
	O	197.0	199.0	200.0	200.1	201.1	214.0
OCS	O	98.7	96.8	108.6	108.8	109.3	111.7
	C	37.9	30.2	44.2	43.9	44.1	45.0
	S	810.9	796.7	811.8	812.1	812.3	830.2
OF_2	O	-402.8	-447.1	-341.9	-341.7	-328.0	-422.1
	F	0.0	-24.0	12.4	14.1	16.7	9.6
H_4C_2O	O	362.1	363.2	359.3	359.3	359.5	370.5
	C	158.1	153.2	158.3	158.3	158.4	157.3
	H	29.4	29.1	29.7	29.7	29.7	29.3
PN	N	-336.0	-344.0	-299.1	-303.1	-301.3	-248.2
	P	89.4	50.6	132.7	126.5	128.9	152.2

Table 6: Statistics for absolute isotropic NMR shieldings in a cc-pVTZ basis. CCSD(T)/aug-cc-pCV[TQ]Z shieldings are used as a reference. MSD=mean signed deviation; MAD=mean absolute deviation; MRD=mean relative deviation [%]; MARD=mean absolute relative deviation [%]; RMSD=root mean squared deviation.

	MSD	MAD	MRD	MARD	RMSD
RHF	-9.4	17.4	-40.5	51.9	34.9
PBE	-18.4	18.6	-41.3	42.0	31.7
KT2	-5.7	9.3	16.5	24.4	17.7
B3LYP	-19.1	19.3	-52.4	53.0	30.6
B97-2	-13.4	13.6	-29.6	30.5	22.7
CCSD	4.2	5.4	19.8	21.4	11.3
CCSD(T)	5.0	6.3	28.2	28.7	10.6
RPA@HF	5.3	6.1	24.2	24.4	12.8
RPA@BH&H	10.0	10.7	46.4	46.5	19.4
RPA@B3LYP	14.4	15.2	64.5	64.7	28.4
RPA@KT2	15.0	16.1	68.1	68.4	30.5
RPA@B97-2	12.9	13.8	59.0	59.2	25.9
RPA@PBE	18.9	19.9	82.0	82.3	40.4
σ-P1@PBE	10.5	11.5	53.5	53.7	21.9
σ-P2@PBE	10.3	11.3	52.9	53.1	21.4
σ-P3@PBE	11.2	12.1	55.1	55.3	23.3
MP2	12.3	13.4	50.8	51.1	23.3

Table 7: Statistics for absolute isotropic NMR shieldings in a cc-pVTZ basis. CCSD(T) shieldings are used as a reference. MSD=mean signed deviation; MAD=mean absolute deviation; MRD=mean relative deviation [%]; MARD=mean absolute relative deviation [%]; RMSD=root mean squared deviation.

	MSD	MAD	MRD	MARD	RMSD
RHF	-14.4	21.0	-45.0	47.8	40.0
PBE	-23.4	23.5	-36.3	36.5	38.6
KT2	-10.6	11.5	-3.4	11.9	21.1
B3LYP	-24.1	24.2	-47.1	47.3	38.2
B97-2	-18.4	18.4	-31.7	31.8	30.2
CCSD	-0.8	3.4	-5.7	6.7	6.4
CCSD(T)	0.0	0.0	0.0	0.0	0.0
RPA@HF	0.3	5.1	-0.5	9.1	9.0
RPA@BH&H	5.0	5.9	13.5	13.7	12.0
RPA@B3LYP	9.4	10.1	23.1	23.3	20.1
RPA@KT2	10.0	10.8	24.7	25.0	22.0
RPA@B97-2	7.9	8.8	19.8	20.0	17.6
RPA@PBE	13.9	14.7	31.1	31.3	31.9
σ-P1@PBE	5.5	6.9	12.8	13.2	13.3
σ-P2@PBE	5.3	6.7	12.4	12.8	12.8
σ-P3@PBE	6.2	7.5	14.1	14.4	14.7
MP2	7.3	8.3	23.5	24.1	16.8

Table 8: Diamagnetic (-dia) contributions to the absolute isotropic NMR shieldings in a cc-pVTZ basis set. All values are given in ppm. The diamagnetic and paramagnetic contributions to the CCSD(T)-shieldings were computed with the CFOUR[1, 2] program. The gauge-origin is located at the position of the nucleus, for which the shielding is computed.

Mol.	Nuc.	CCSD(T)-dia	RPA@HF-dia	RPA@PBE-dia	RPA@BH&H-dia	RPA@B3LYP-dia	RPA@KT2-dia
HF	H	29.0	29.0	29.2	29.2	29.2	29.2
	F	466.9	466.7	466.9	466.8	466.9	466.8
CO	C	259.7	259.9	260.2	259.9	260.0	260.2
	O	395.7	395.3	395.0	395.3	395.2	395.1
N_2	N	324.9	324.7	324.6	324.7	324.7	324.7
H_2O	O	383.2	382.5	383.3	382.7	383.0	383.1
	H	30.2	30.6	30.5	30.6	30.5	30.5
HCN	H	28.9	29.0	29.0	29.0	29.0	28.9
	C	259.1	259.3	259.6	259.4	259.5	259.6
	N	334.2	332.8	332.5	332.9	332.7	332.6
HOF	O	379.9	379.7	381.1	380.2	380.6	381.0
	H	28.6	28.0	28.0	28.0	28.0	28.0
	F	466.6	466.3	467.1	466.6	466.8	467.1
NH_3	N	310.5	309.3	310.6	309.8	310.2	310.3
	H	30.3	30.9	30.7	30.8	30.7	30.6
CH_2O	O	398.6	399.6	399.4	399.7	399.6	399.5
	C	246.4	245.8	247.2	246.4	246.9	247.2
	H	28.4	29.6	29.2	29.4	29.3	29.2
CH_4	C	247.6	246.6	248.2	247.2	247.8	247.9
	H	29.4	29.9	29.7	29.8	29.8	29.7
C2H4	C	253.3	253.1	253.9	253.4	253.7	253.8
	H	30.1	30.8	30.5	30.7	30.6	30.5
AlF	Al	791.1	790.8	791.0	790.9	790.9	791.0
	F	467.0	466.8	466.4	466.7	466.5	466.3
CH_3F	C	238.4	240.3	242.2	241.1	241.8	241.9
	F	469.9	471.4	471.2	471.4	471.3	471.3
	H	28.3	29.1	28.9	29.0	28.9	28.8
C_3H_4	C	246.5	236.1	238.0	236.9	237.5	237.7
	C	254.9	245.8	247.0	246.2	246.7	246.9
	H	29.2	25.6	25.4	25.6	25.5	25.5
	H	29.5	22.0	21.7	21.9	21.8	21.7
FCCH	C	261.1	256.4	256.6	256.5	256.6	256.4
	C	258.8	257.8	257.9	257.9	258.0	257.8
	H	29.4	30.3	30.3	30.3	30.3	30.3
	F	467.1	462.8	463.5	463.1	463.3	463.5
FCN	F	465.4	464.5	465.2	464.8	465.0	465.1
	C	263.3	263.1	263.5	263.3	263.5	263.5
	N	330.4	329.6	329.1	329.6	329.4	329.1
H_2S	S	1054.7	1054.5	1054.1	1054.4	1054.3	1054.2
	H	29.5	30.2	30.0	30.1	30.1	30.0
HCP	H	28.8	29.2	29.0	29.1	29.0	29.0
	C	261.6	260.2	260.6	260.4	260.4	260.4
	P	966.9	970.5	969.2	970.1	969.7	969.3
HFCO	O	398.4	399.6	399.6	399.8	399.7	399.6
	C	247.5	248.3	249.7	249.0	249.4	249.7
	F	471.1	472.0	472.2	472.1	472.2	472.2
	H	28.1	29.9	29.6	29.8	29.7	29.6
H_2C_2O	C	257.5	256.5	257.2	256.9	257.1	257.1
	C	256.4	255.8	256.3	256.0	256.2	256.3
	O	397.9	396.2	396.2	396.3	396.3	396.2
	H	29.9	30.2	29.9	30.1	30.0	29.9
LiF	Li	100.8	100.7	101.1	100.9	101.0	101.1
	F	471.5	468.5	469.0	468.6	468.8	468.7
LiH	H	26.0	26.5	26.0	26.3	26.2	26.0
	Li	101.7	101.2	101.3	101.2	101.2	101.3
N_2O	N	326.1	324.2	324.4	324.5	324.5	324.4
	N	320.7	319.9	320.4	320.1	320.3	320.5
	O	398.4	397.5	397.5	397.6	397.6	397.6
OCS	O	394.8	393.2	393.3	393.4	393.4	393.3
	C	260.7	258.5	259.2	258.8	259.0	259.1
	S	1055.5	1057.5	1056.5	1057.2	1056.8	1056.6
OF_2	O	378.5	378.6	380.6	379.4	380.0	380.6
	F	466.5	465.8	466.6	466.2	466.4	466.6
H_4C_2O	O	391.4	388.7	389.1	389.0	389.0	389.4
	C	246.9	245.1	246.6	245.8	246.3	246.5
	H	29.4	28.7	28.4	28.6	28.5	28.4
PN	N	330.4	329.5	329.0	329.5	329.3	329.0
	P	967.2	967.3	966.8	967.2	967.0	966.9

Table 9: Paramagnetic (-para) contributions to the absolute isotropic NMR shieldings in a cc-pVTZ basis set. All values are given in ppm. The diamagnetic and paramagnetic contributions to the CCSD(T)-shieldings were computed with the CFOUR[1, 2] program. The gauge-origin is located at the position of the nucleus, for which the shielding is computed.

Mol.	Nuc.	CCSD(T)-para	RPA@HF-para	RPA@PBE-para	RPA@BH&H-para	RPA@B3LYP-para	RPA@KT2-para
HF	H	0.5	0.4	0.9	0.5	0.7	0.9
	F	-47.2	-47.7	-49.0	-48.2	-48.6	-48.6
CO	C	-248.4	-254.3	-231.8	-245.6	-238.5	-236.6
	O	-445.5	-443.7	-416.4	-433.1	-424.9	-421.3
N_2	N	-377.5	-382.9	-353.4	-372.0	-362.5	-359.4
H_2O	O	-38.9	-40.8	-42.2	-41.2	-41.5	-41.5
	H	1.0	0.7	1.2	0.9	1.1	1.2
HCN	H	0.4	0.5	0.5	0.5	0.5	0.6
	C	-167.6	-171.8	-160.9	-167.2	-163.8	-163.2
	N	-341.6	-344.6	-320.0	-334.2	-326.2	-326.0
HOF	O	-424.4	-420.3	-357.0	-406.9	-385.8	-380.6
	H	-7.9	-6.7	-5.1	-6.5	-6.0	-5.7
	F	-278.3	-254.0	-253.4	-256.7	-256.6	-256.6
NH_3	N	-34.4	-36.5	-37.6	-36.8	-36.9	-37.0
	H	1.6	1.2	1.5	1.4	1.5	1.6
CH_2O	O	-775.0	-751.2	-665.8	-716.1	-693.1	-691.8
	C	-233.2	-233.5	-215.6	-227.8	-222.1	-221.8
	H	-6.1	-6.7	-5.9	-6.5	-6.3	-6.1
CH_4	C	-46.2	-48.2	-48.9	-48.3	-48.2	-48.6
	H	2.0	1.7	1.9	1.8	1.8	1.9
C2H4	C	-176.1	-179.8	-170.9	-175.8	-172.9	-173.1
	H	-3.8	-4.3	-3.8	-4.1	-3.9	-3.8
AlF	Al	-214.6	-209.9	-204.9	-210.2	-207.6	-207.6
	F	-266.8	-248.9	-248.8	-250.7	-250.2	-250.7
CH_3F	C	-111.0	-113.0	-112.2	-113.0	-112.3	-112.9
	F	9.4	6.0	0.8	3.1	2.0	1.2
	H	-0.6	-1.1	-0.9	-1.1	-0.9	-0.8
C_3H_4	C	-51.6	-42.9	-43.7	-43.4	-43.1	-44.1
	C	-164.6	-158.8	-150.6	-155.4	-152.5	-152.8
	H	-4.6	-0.8	-0.2	-0.6	-0.4	-0.4
	H	1.4	9.1	9.4	9.2	9.3	9.4
FCCH	C	-77.9	-75.7	-72.8	-73.8	-72.9	-73.8
	C	-153.4	-153.0	-149.5	-151.5	-150.1	-150.5
	H	1.3	0.6	0.7	0.6	0.7	0.7
	F	-45.8	-39.6	-42.1	-40.6	-41.5	-42.2
FCN	F	-92.3	-86.6	-88.5	-88.1	-88.6	-88.3
	C	-175.1	-177.7	-171.1	-174.5	-172.2	-172.7
	N	-207.9	-212.9	-197.6	-204.5	-199.8	-201.0
H_2S	S	-296.2	-309.7	-300.8	-306.7	-302.2	-302.0
	H	1.4	0.9	1.3	1.1	1.1	1.3
HCP	H	0.9	0.8	0.9	0.9	0.9	1.0
	C	-216.3	-222.0	-206.5	-214.7	-209.9	-209.5
	P	-554.0	-575.6	-522.4	-553.3	-534.8	-533.4
HFCO	O	-486.4	-482.5	-454.2	-465.7	-458.4	-458.6
	C	-200.5	-203.3	-196.4	-200.2	-197.7	-198.4
	F	-303.9	-288.0	-287.7	-290.5	-289.0	-291.9
	H	-3.8	-5.3	-5.0	-5.2	-5.1	-5.0
H_2C_2O	C	-61.3	-62.4	-61.8	-62.0	-61.5	-61.8
	C	-252.9	-254.0	-244.5	-249.4	-246.3	-246.0
	O	-401.8	-392.1	-378.0	-382.8	-379.6	-378.8
	H	-0.3	-0.5	0.0	-0.3	-0.2	0.0
LiF	Li	-10.1	-9.8	-10.2	-10.1	-10.2	-10.6
	F	-101.2	-97.4	-99.3	-98.0	-97.8	-105.8
LiH	H	0.5	0.3	0.5	0.5	0.4	0.5
	Li	-11.7	-11.3	-10.4	-11.0	-10.7	-10.7
N_2O	N	-211.8	-216.4	-199.0	-208.3	-202.7	-201.5
	N	-298.6	-306.0	-285.9	-296.8	-290.5	-288.2
	O	-201.4	-199.8	-190.3	-195.0	-192.3	-191.2
OCS	O	-296.1	-292.6	-280.7	-286.8	-283.3	-282.0
	C	-222.9	-226.3	-212.2	-219.5	-215.1	-213.9
	S	-244.6	-246.4	-233.1	-240.6	-234.7	-236.2
OF_2	O	-781.3	-756.0	-610.6	-736.2	-689.3	-681.1
	F	-466.5	-438.6	-417.3	-435.1	-427.8	-424.2
H_4C_2O	O	-29.3	-24.0	-30.1	-26.1	-27.5	-31.7
	C	-88.8	-87.6	-87.4	-87.8	-87.1	-88.2
	H	0.0	1.0	1.3	1.1	1.1	1.2
PN	N	-666.4	-678.0	-597.5	-647.2	-622.4	-615.4
	P	-877.8	-908.9	-789.3	-866.4	-829.2	-814.7

Table 10: Diamagnetic (-dia) contributions to the absolute isotropic NMR shieldings in a cc-pVTZ basis set. All values are given in ppm. The diamagnetic and paramagnetic contributions to the CCSD(T)-shieldings were computed with the CFOUR[1, 2] program. The gauge-origin is located at the the position of the nucleus, for which the shielding is computed.

Mol.	Nuc.	CCSD(T)-dia	HF(num.)-dia	PBE(num.)-dia	BH&H(num.)-dia	B3LYP(num.)-dia	KT2(num.)-dia
HF	H	29.0	28.5	29.8	29.0	29.5	29.5
	F	466.9	466.0	467.3	467.0	467.2	470.9
CO	C	259.7	259.4	259.6	259.4	259.3	261.2
	O	395.7	394.1	397.7	396.1	397.2	401.3
N_2	N	324.9	324.3	325.6	325.0	325.2	327.4
H_2O	O	383.2	381.1	382.3	382.1	382.2	386.1
	H	30.2	30.5	31.5	30.9	31.3	31.2
HCN	H	28.9	29.0	29.7	29.3	29.7	29.6
	C	259.1	259.4	258.0	258.7	257.9	259.7
	N	334.2	330.9	336.6	333.8	335.6	338.9
HOF	O	379.9	377.7	379.9	379.0	379.5	383.2
	H	28.6	27.5	28.9	28.0	28.5	28.5
	F	466.6	466.1	466.2	466.4	466.4	469.2
NH_3	N	310.5	307.2	308.9	308.3	308.4	312.5
	H	30.3	30.9	31.5	31.1	31.3	31.1
CH_2O	O	398.6	398.0	402.2	400.4	401.8	405.3
	C	246.4	243.8	245.2	244.1	244.1	247.6
	H	28.4	29.8	29.6	29.7	29.7	29.2
CH_4	C	247.6	243.7	248.0	245.4	246.0	251.1
	H	29.4	30.0	30.3	30.0	30.1	29.8
C2H4	C	253.3	251.7	253.3	252.2	252.2	255.9
	H	30.1	31.1	31.1	31.0	31.1	30.5
AlF	Al	791.1	790.8	792.4	791.7	792.1	794.7
	F	467.0	466.0	467.9	467.6	468.4	471.0
CH_3F	C	238.4	237.4	240.5	238.4	238.6	243.6
	F	469.9	470.5	473.7	472.6	473.8	476.4
	H	28.3	29.1	29.3	29.1	29.1	28.8
C_3H_4	C	246.5	233.2	236.5	234.4	235.0	238.7
	C	254.9	244.6	243.7	244.2	243.5	245.8
	H	29.2	25.7	26.4	26.0	26.3	25.8
	H	29.5	22.3	21.8	21.9	21.8	20.6
FCCH	C	261.1	255.8	256.3	256.0	255.7	258.4
	C	258.8	257.2	257.3	257.0	256.7	259.9
	H	29.4	30.2	31.1	30.6	31.0	31.0
	F	467.1	460.9	463.7	463.0	463.9	468.0
FCN	F	465.4	462.8	466.2	465.0	466.1	470.3
	C	263.3	262.6	262.3	262.3	262.0	264.1
	N	330.4	328.0	333.6	330.7	332.3	335.8
H_2S	S	1054.7	1055.1	1056.7	1056.1	1056.5	1059.5
	H	29.5	30.2	30.1	30.1	30.2	29.9
HCP	H	28.8	29.1	30.0	29.5	29.9	30.0
	C	261.6	259.4	259.0	259.7	259.6	262.9
	P	966.9	971.7	976.2	974.1	975.5	979.9
HFCO	O	398.4	398.0	402.7	400.5	401.9	405.9
	C	247.5	246.4	247.9	246.7	246.8	249.9
	F	471.1	470.7	474.4	473.2	474.4	477.5
	H	28.1	29.9	30.6	30.1	30.4	30.0
H_2C_2O	C	257.5	254.9	258.5	256.2	256.7	260.8
	C	256.4	255.5	254.7	255.2	254.8	256.9
	O	397.9	394.5	398.3	396.6	397.6	401.6
	H	29.9	30.4	30.7	30.4	30.6	30.0
LiF	Li	100.8	100.4	101.9	101.1	101.6	104.7
	F	471.5	467.8	465.9	467.3	466.7	469.6
LiH	H	26.0	26.2	25.8	26.1	26.1	25.8
	Li	101.7	101.4	101.9	101.8	101.8	104.9
N_2O	N	326.1	322.9	326.7	324.5	325.4	328.7
	N	320.7	319.6	318.1	319.1	318.6	321.1
	O	398.4	396.2	399.2	397.8	398.8	401.6
OCS	O	394.8	391.5	395.3	393.6	394.6	398.8
	C	260.7	257.5	257.2	257.6	257.5	260.1
	S	1055.5	1058.5	1060.8	1059.7	1060.4	1064.1
OF_2	O	378.5	376.2	378.9	377.5	378.2	381.8
	F	466.5	464.8	466.5	465.8	466.3	469.5
H_4C_2O	O	391.4	387.8	389.7	389.4	390.3	391.2
	C	246.9	242.8	245.4	243.6	243.8	247.8
	H	29.4	28.9	29.1	28.8	28.9	28.1
PN	N	330.4	327.3	331.1	329.8	331.3	334.3
	P	967.2	967.9	969.5	968.7	969.1	971.8

Table 11: Paramagnetic (-para) contributions to the absolute isotropic NMR shieldings in a cc-pVTZ basis set. All values are given in ppm. The diamagnetic and paramagnetic contributions to the CCSD(T)-shieldings were computed with the CFOUR[1, 2] program. The gauge-origin is located at the the position of the nucleus, for which the shielding is computed.

Mol.	Nuc.	CCSD(T)-para	HF(num.)-para	PBE(num.)-para	BH&H(num.)-para	B3LYP(num.)-para	KT2(num.)-para
HF	H	0.5	0.0	0.4	0.1	0.3	0.8
	F	-47.2	-50.4	-54.8	-52.2	-54.3	-59.0
CO	C	-248.4	-281.2	-266.0	-277.5	-272.2	-246.7
	O	-445.5	-481.3	-473.8	-480.4	-477.2	-456.3
N_2	N	-377.5	-433.1	-402.7	-421.5	-411.9	-380.3
H_2O	O	-38.9	-45.7	-44.9	-45.6	-46.4	-50.4
	H	1.0	0.4	0.2	0.3	0.2	0.6
HCN	H	0.4	0.3	-0.4	0.1	-0.3	-0.6
	C	-167.6	-185.0	-179.8	-185.1	-183.0	-167.1
	N	-341.6	-378.2	-374.2	-380.9	-379.6	-351.5
HOF	O	-424.4	-497.0	-484.5	-494.9	-494.0	-454.1
	H	-7.9	-7.9	-8.8	-8.1	-8.5	-7.8
	F	-278.3	-179.6	-316.7	-249.4	-292.8	-306.3
NH_3	N	-34.4	-38.7	-39.5	-40.0	-41.2	-41.4
	H	1.6	1.0	0.7	0.9	0.9	1.1
CH_2O	O	-775.0	-841.6	-847.6	-858.2	-855.1	-782.5
	C	-233.2	-245.9	-260.8	-255.8	-259.6	-241.9
	H	-6.1	-7.1	-8.6	-7.7	-8.1	-7.7
CH_4	C	-46.2	-47.0	-54.4	-51.5	-54.5	-52.6
	H	2.0	1.6	1.2	1.6	1.5	1.8
C2H4	C	-176.1	-187.9	-196.9	-195.0	-197.7	-184.4
	H	-3.8	-4.8	-5.4	-4.9	-5.1	-4.5
AlF	Al	-214.6	-207.7	-247.8	-227.6	-241.3	-235.3
	F	-266.8	-249.7	-336.5	-286.5	-313.9	-336.2
CH_3F	C	-111.0	-109.6	-129.2	-120.4	-127.1	-123.9
	F	9.4	14.4	-15.9	0.8	-11.0	-18.1
	H	-0.6	-1.1	-2.1	-1.4	-1.6	-1.5
C_3H_4	C	-51.6	-38.0	-56.4	-47.7	-54.1	-52.7
	C	-164.6	-169.7	-173.6	-174.5	-176.1	-162.1
	H	-4.6	-1.5	-2.1	-1.8	-1.9	-1.4
	H	1.4	8.7	8.9	9.0	9.0	10.1
FCCH	C	-77.9	-77.9	-82.6	-82.3	-83.3	-74.2
	C	-153.4	-152.9	-168.1	-162.5	-167.6	-160.0
	H	1.3	0.4	0.0	0.2	0.0	-0.1
	F	-45.8	-35.8	-77.3	-54.1	-66.5	-76.4
FCN	F	-92.3	-87.6	-128.5	-104.8	-116.7	-133.5
	C	-175.1	-183.8	-186.5	-188.3	-189.5	-175.8
	N	-207.9	-234.7	-229.6	-238.1	-236.1	-212.2
H_2S	S	-296.2	-319.9	-328.9	-333.0	-342.5	-321.5
	H	1.4	0.6	0.9	0.8	0.9	1.1
HCP	H	0.9	1.1	-0.4	0.5	0.0	-0.7
	C	-216.3	-242.3	-236.8	-244.6	-243.0	-225.1
	P	-554.0	-616.6	-638.4	-642.0	-646.7	-595.0
HFCO	O	-486.4	-524.6	-522.1	-534.7	-531.7	-496.2
	C	-200.5	-209.1	-218.3	-216.7	-219.6	-206.3
	F	-303.9	-282.4	-381.3	-326.5	-357.5	-368.5
	H	-3.8	-5.3	-7.0	-5.9	-6.5	-6.4
H_2C_2O	C	-61.3	-63.9	-69.7	-68.0	-70.0	-65.5
	C	-252.9	-264.5	-262.1	-269.0	-268.7	-249.4
	O	-401.8	-421.2	-415.1	-426.1	-423.4	-401.2
	H	-0.3	-0.9	-1.2	-0.9	-1.0	-0.4
LiF	Li	-10.1	-8.6	-13.6	-10.9	-12.5	-12.5
	F	-101.2	-94.9	-150.1	-122.7	-140.8	-129.6
LiH	H	0.5	0.3	0.5	0.5	0.4	0.6
	Li	-11.7	-11.3	-12.7	-12.5	-12.6	-11.7
N_2O	N	-211.8	-255.8	-227.4	-245.6	-237.0	-219.5
	N	-298.6	-346.7	-305.5	-331.8	-320.1	-298.4
	O	-201.4	-225.7	-225.6	-229.1	-229.0	-226.8
OCS	O	-296.1	-315.6	-321.7	-320.4	-322.2	-314.8
	C	-222.9	-245.6	-225.3	-240.8	-235.1	-215.5
	S	-244.6	-264.8	-289.5	-283.7	-291.7	-279.1
OF_2	O	-781.3	-796.1	-965.7	-869.5	-927.8	-883.3
	F	-466.5	-424.1	-530.1	-479.1	-513.5	-511.9
H_4C_2O	O	-29.3	-9.9	-59.6	-36.1	-53.1	-53.8
	C	-88.8	-84.1	-102.4	-94.2	-100.6	-96.9
	H	0.0	0.9	0.1	0.7	0.5	1.2
PN	N	-666.4	-836.4	-734.9	-796.4	-766.4	-695.7
	P	-877.8	-1059.1	-995.6	-1043.9	-1020.4	-921.9

Table 12: Testing influence of δ and DFT grid and gauge-origin independence in a cc-pVTZ basis set. All values are given in ppm. "RPA@HF0.05" denotes calculations with $\delta = 0.05$ instead of $\delta = 0.1$ as in the other calculations. RPA@HF_trans denotes calculations after translating the molecule by 5.0 Angstroem in x-, y- and z-direction, respectively. The gauge-origin is located at (0.0, 0.0, 0.0). The "RPA@PBE" calculations are performed with a g5 grid; in contrast, a g7 grid is used in the "RPA@PBE-g7" calculations.

Mol.	Nuc.	RPA@HF	RPA@HF0.05	RPA@HF_trans	RPA@PBE	RPA@PBE-g7
HF	H	29.4	29.4	29.4	30.1	30.1
	F	419.0	419.0	419.0	417.9	417.9
CO	C	5.6	5.6	5.6	28.4	28.4
	O	-48.4	-48.4	-48.4	-21.4	-21.4
N_2	N	-58.2	-58.2	-58.2	-28.8	-28.8
H_2O	O	341.7	341.7	341.7	341.1	341.1
	H	31.3	31.3	31.3	31.7	31.7
HCN	H	29.5	29.5	29.5	29.5	29.5
	C	87.5	87.5	87.5	98.7	98.7
	N	-11.8	-11.8	-11.8	12.5	12.5
HOF	O	-40.6	-40.7	-40.6	24.1	24.1
	H	21.3	21.3	21.3	22.9	22.9
	F	212.3	212.2	212.3	213.7	213.7
NH_3	N	272.8	272.8	272.8	273.0	273.0
	H	32.1	32.1	32.1	32.2	32.2
CH_2O	O	-351.6	-351.8	-351.6	-266.4	-266.4
	C	12.3	12.3	12.3	31.6	31.6
	H	22.9	22.9	22.9	23.3	23.3
CH_4	C	198.4	198.4	198.4	199.3	199.3
	H	31.6	31.6	31.6	31.6	31.6
C2H4	C	73.3	73.3	73.3	83.0	83.0
	H	26.5	26.5	26.5	26.7	26.7
AlF	Al	580.9	580.9	580.9	586.1	586.1
	F	217.9	218.0	217.9	217.6	217.6
CH_3F	C	127.3	127.3	127.3	130.0	130.0
	F	477.4	477.4	477.4	472.0	472.0
	H	28.0	28.0	28.0	28.0	28.0
C_3H_4	C	193.2	193.2	193.2	194.3	194.3
	C	87.0	87.0	87.0	96.4	96.4
	H	24.8	24.8	24.8	25.2	25.2
	H	31.1	31.1	31.1	31.1	31.1
FCCH	C	180.7	180.7	180.7	183.8	183.8
	C	104.8	104.8	104.8	108.4	108.4
	H	30.9	30.9	30.9	31.0	31.0
	F	423.2	423.2	423.2	421.4	421.4
FCN	F	377.9	377.9	377.9	376.7	376.7
	C	85.4	85.4	85.4	92.4	92.4
	N	116.7	116.7	116.7	131.5	131.5
H_2S	S	744.8	744.9	744.9	753.3	753.3
	H	31.1	31.1	31.1	31.3	31.3
HCP	H	30.0	30.0	30.0	29.9	29.9
	C	38.2	38.2	38.2	54.1	54.1
	P	394.9	394.9	394.8	446.8	446.8
HFCO	O	-82.9	-82.9	-82.9	-54.6	-54.6
	C	45.0	45.0	45.0	53.3	53.3
	F	184.0	184.1	184.1	184.5	184.5
	H	24.6	24.6	24.6	24.6	24.6
H_2C_2O	C	194.1	194.1	194.1	195.4	195.4
	C	1.8	1.8	1.8	11.8	11.8
	O	4.1	4.1	4.1	18.2	18.2
	H	29.7	29.7	29.7	29.9	29.9
LiF	Li	90.9	90.9	90.9	90.9	90.9
	F	371.1	371.1	371.1	369.7	369.7
LiH	H	26.8	26.8	26.8	26.5	26.5
	Li	89.9	89.9	89.9	90.9	90.9
N_2O	N	107.8	107.8	107.7	125.4	125.4
	N	13.9	13.9	13.9	34.5	34.5
	O	197.7	197.7	197.7	207.2	207.2
OCS	O	100.6	100.6	100.6	112.6	112.6
	C	32.2	32.2	32.2	47.0	47.0
	S	811.1	811.1	811.0	823.4	823.4
OF_2	O	-377.4	-377.5	-377.4	-230.0	-230.0
	F	27.2	27.1	27.2	49.3	49.3
H_4C_2O	O	364.7	364.7	364.7	359.0	359.0
	C	157.5	157.5	157.5	159.2	159.2
	H	29.7	29.7	29.7	29.7	29.7
PN	N	-348.5	-348.5	-348.5	-268.5	-268.5
	P	58.4	58.4	58.4	177.5	177.5

Table 13: Comparison of analytical and numerical NMR shieldings in a cc-pVTZ basis set. All values are given in ppm. The HF and PBE values were taken from Ref. 3. "HF(num.)" and "PBE(num.)" denote the numerically computed shieldings.

Mol.	Nuc.	RHF	HF(num.)	PBE	PBE(num.)
HF	H	28.5	28.5	30.2	30.2
	F	415.6	415.6	412.5	412.5
CO	C	-21.8	-21.8	-6.4	-6.4
	O	-87.2	-87.2	-76.1	-76.1
N_2	N	-108.8	-108.8	-77.1	-77.1
H_2O	O	335.4	335.4	337.4	337.4
	H	30.9	30.9	31.7	31.7
HCN	H	29.3	29.3	29.3	29.3
	C	74.4	74.4	78.2	78.2
	N	-47.3	-47.3	-37.6	-37.6
HOF	O	-119.5	-119.3	-104.9	-104.6
	H	19.6	19.6	20.1	20.1
	F	286.6	286.5	149.2	149.5
NH_3	N	268.5	268.5	269.4	269.4
	H	31.9	31.9	32.2	32.2
CH_2O	O	-443.9	-443.6	-445.7	-445.4
	C	-2.1	-2.1	-15.6	-15.6
	H	22.7	22.7	21.0	21.0
CH_4	C	196.7	196.7	193.6	193.6
	H	31.6	31.6	31.5	31.5
C2H4	C	63.8	63.8	56.4	56.4
	H	26.3	26.3	25.7	25.7
AlF	Al	583.1	583.1	544.6	544.6
	F	216.3	216.3	131.4	131.4
CH_3F	C	127.8	127.8	111.3	111.3
	F	484.9	484.9	457.9	457.8
	H	28.0	28.0	27.2	27.2
C_3H_4	C	195.2	195.2	180.1	180.1
	C	74.9	74.9	70.1	70.1
	H	24.2	24.2	24.3	24.3
	H	31.0	31.0	30.7	30.7
FCCH	C	177.9	177.9	173.7	173.7
	C	104.3	104.3	89.2	89.2
	H	30.6	30.6	31.1	31.1
	F	425.1	425.1	386.4	386.4
FCN	F	375.2	375.2	337.7	337.7
	C	78.8	78.8	75.8	75.8
	N	93.3	93.3	104.0	104.0
H_2S	S	735.2	735.2	727.8	727.8
	H	30.8	30.8	31.0	31.0
HCP	H	30.2	30.2	29.6	29.6
	C	17.1	17.1	22.2	22.2
	P	355.1	355.1	337.8	337.8
HFCO	O	-126.7	-126.6	-119.4	-119.4
	C	37.3	37.3	29.6	29.6
	F	188.3	188.3	93.1	93.1
	H	24.6	24.6	23.6	23.6
H_2C_2O	C	191.0	191.0	188.8	188.8
	C	-9.0	-9.0	-7.4	-7.4
	O	-26.7	-26.7	-16.9	-16.8
	H	29.5	29.5	29.5	29.5
LiF	Li	91.8	91.8	88.3	88.3
	F	373.0	372.9	315.9	315.8
LiH	H	26.5	26.5	26.3	26.3
	Li	90.1	90.1	89.2	89.2
N_2O	N	67.1	67.1	99.3	99.3
	N	-27.1	-27.1	12.6	12.6
	O	170.4	170.5	173.6	173.6
OCS	O	75.9	75.9	73.6	73.6
	C	11.9	11.9	31.9	31.9
	S	793.7	793.7	771.3	771.3
OF_2	O	-420.2	-419.9	-587.3	-586.8
	F	40.7	40.7	-63.8	-63.6
H_4C_2O	O	377.9	377.9	330.1	330.1
	C	158.7	158.7	143.0	143.0
	H	29.8	29.8	29.2	29.2
PN	N	-509.2	-509.1	-403.9	-403.8
	P	-91.4	-91.2	-26.2	-26.1

Table 14: Testing influence of number of frequency integration points in a cc-pVTZ basis set. The default of 120 grid points is used for the RPA frequency integration in “RPA@BH&H” and “σ-P1@PBE”. In the other calculations with suffix “-npX”, X frequency points are employed. All values are given in ppm.

Mol.	Nuc.	RPA@BH&H	RPA@BH&H-np90	σ-P1@PBE	σ-P1@PBE-np160	σ-P1@PBE-np200	σ-P1@PBE-np250
HF	H	29.7	29.7	30.0	30.0	30.0	30.0
	F	418.6	418.6	415.7	415.7	415.7	415.7
CO	C	14.3	14.3	18.4	18.4	18.7	18.4
	O	-37.8	-37.8	-36.5	-36.5	-36.5	-36.4
N_2	N	-47.3	-47.3	-43.0	-43.0	-43.0	-43.0
H_2O	O	341.5	341.5	338.5	338.5	338.5	338.5
	H	31.5	31.5	31.6	31.6	31.6	31.6
HCN	H	29.5	29.5	29.5	29.5	29.5	29.5
	C	92.2	92.2	94.8	95.2	95.2	95.2
	N	-1.3	-1.3	3.5	3.7	3.7	3.7
HOF	O	-26.7	-26.7	-19.6	-19.5	-19.6	-19.6
	H	21.5	21.5	21.8	21.8	21.8	21.8
	F	209.9	209.9	190.4	190.4	190.4	190.4
NH_3	N	273.0	273.0	271.3	271.3	271.3	271.3
	H	32.2	32.2	32.1	32.1	32.1	32.1
CH_2O	O	-316.4	-316.4	-333.6	-333.5	-333.5	-333.5
	C	18.6	18.6	24.1	24.1	23.9	24.1
	H	22.9	22.9	23.0	23.0	23.0	23.0
CH_4	C	198.9	198.9	198.4	198.4	198.4	198.4
	H	31.6	31.6	31.6	31.6	31.6	31.6
C2H4	C	77.6	77.6	80.8	80.8	80.8	80.8
	H	26.6	26.6	26.7	26.7	26.7	26.7
AlF	Al	580.7	580.7	583.0	583.0	583.0	583.0
	F	216.0	216.0	214.0	214.0	214.1	214.0
CH_3F	C	128.1	128.1	128.8	128.8	128.8	128.8
	F	474.5	474.5	475.4	475.4	475.3	475.3
	H	27.9	27.9	28.1	28.1	28.1	28.1
C_3H_4	C	193.5	193.5	194.3	194.2	194.2	194.2
	C	90.8	90.8	94.1	94.1	94.1	94.1
	H	25.0	25.0	25.1	25.1	25.1	25.1
	H	31.1	31.1	31.2	31.2	31.2	31.2
FCCH	C	182.7	182.7	182.5	182.5	182.6	182.5
	C	106.4	106.4	107.6	107.6	107.6	107.6
	H	30.9	30.9	31.0	31.0	31.0	31.0
	F	422.5	422.5	421.7	421.7	421.7	421.7
FCN	F	376.7	376.7	373.6	373.6	373.2	373.6
	C	88.8	88.8	90.5	90.5	90.5	90.5
	N	125.1	125.1	125.6	125.6	125.6	125.6
H_2S	S	747.7	747.7	739.5	739.6	739.5	739.6
	H	31.2	31.2	31.4	31.4	31.4	31.4
HCP	H	30.0	30.0	30.1	30.1	30.1	30.1
	C	45.7	45.7	52.5	52.5	52.5	52.5
	P	416.8	416.8	436.5	436.5	436.5	436.5
HFCO	O	-65.9	-65.9	-76.7	-76.7	-76.7	-76.7
	C	48.8	48.8	49.9	49.9	49.9	49.9
	F	181.6	181.6	174.4	174.4	174.3	174.4
	H	24.6	24.6	24.6	24.6	24.6	24.6
H_2C_2O	C	194.9	194.9	194.0	193.9	194.0	194.0
	C	6.6	6.6	5.3	5.3	5.3	5.4
	O	13.5	13.5	3.1	3.2	3.2	3.2
	H	29.8	29.8	29.8	29.8	29.8	29.8
LiF	Li	90.8	90.8	90.8	90.8	90.8	90.8
	F	370.6	370.6	369.7	369.7	369.7	369.7
LiH	H	26.8	26.8	26.7	26.7	26.7	26.7
	Li	90.2	90.2	90.6	90.6	90.6	90.6
N_2O	N	116.2	116.2	118.0	118.4	118.0	118.0
	N	23.3	23.3	29.7	29.7	29.7	29.7
	O	202.6	202.6	200.0	200.0	200.0	202.5
OCS	O	106.6	106.6	108.6	108.4	108.4	108.1
	C	39.3	39.3	44.2	44.2	44.2	44.0
	S	816.6	816.6	811.8	811.8	811.8	811.8
OF_2	O	-356.8	-356.8	-341.9	-341.6	-341.3	-341.9
	F	31.1	31.1	12.4	12.4	12.8	12.3
H_4C_2O	O	362.9	362.9	359.3	359.3	359.3	359.3
	C	158.0	158.0	158.3	158.3	158.3	158.3
	H	29.7	29.7	29.7	29.7	29.7	29.7
PN	N	-317.7	-317.7	-299.1	-299.1	-299.1	-299.0
	P	100.8	100.8	132.7	132.3	132.7	132.7

Table 15: Absolute isotropic NMR shieldings in a cc-pVQZ basis. All values are given in ppm. “Extrap.” stands for basis-set extrapolated CCSD(T)-shieldings in an aug-cc-pCV[TQ]Z basis set. The data for “Extrap.” was taken from Ref. 3. CCSD(T) calculations were done with the CFOUR program[1, 2]. MP2 calculations were done with the Turbomole program[4–6].

Mol.	Nuc.	CCSD(T)	Extrap.	RPA@HF	RPA@PBE	σ-P1@PBE	MP2
HF	H	29.1	28.8	29.0	29.6	29.1	28.8
	F	419.9	420.3	419.2	418.5	418.3	425.3
CO	C	5.9	2.2	-0.5	22.4	11.5	11.2
	O	-51.9	-55.0	-50.7	-25.1	-38.9	-45.2
N_2	N	-57.6	-60.4	-64.3	-34.7	-49.5	-40.6
H_2O	O	341.0	338.0	338.5	338.5	338.0	349.1
	H	30.9	30.6	30.9	31.2	30.9	30.6
HCN	H	29.1	29.0	29.3	29.3	29.3	29.0
	C	87.3	84.6	83.3	94.4	90.1	88.7
	N	-11.3	-14.1	-15.9	8.2	-1.3	1.6
HOF	O	-58.7	-68.9	-57.7	8.6	-37.1	-43.9
	H	20.0	19.6	20.6	22.0	20.7	20.1
	F	191.6	192.2	215.3	214.2	194.7	194.9
NH_3	N	273.1	270.7	269.8	270.6	269.5	278.7
	H	31.7	31.4	31.8	31.9	31.7	31.5
CH_2O	O	-377.2	-378.6	-350.5	-271.3	-333.0	-332.6
	C	6.5	1.5	5.7	24.7	15.3	6.6
	H	22.1	22.0	22.6	23.0	22.6	22.1
CH_4	C	199.9	198.9	196.9	197.9	198.0	201.5
	H	31.4	31.3	31.5	31.5	31.4	31.3
C2H4	C	72.9	69.7	69.0	78.7	75.7	71.4
	H	26.1	26.0	26.4	26.5	26.4	26.0
AlF	Al	587.7	572.9	589.8	589.7	587.5	587.0
	F	212.2	211.8	228.7	229.5	231.0	209.5
CH_3F	C	124.5	122.1	124.6	127.3	126.1	123.0
	F	482.8	482.9	480.2	475.6	481.1	488.9
	H	27.5	27.3	27.8	27.8	27.7	27.4
C_3H_4	C	193.2	192.1	191.4	192.4	192.4	194.1
	C	86.5	83.7	83.1	92.5	89.7	87.7
	H	24.5	24.4	24.6	24.9	24.8	24.4
	H	30.7	30.6	30.9	30.9	30.9	30.7
FCCH	C	181.4	179.9	179.0	181.9	179.6	183.2
	C	102.6	100.1	102.3	106.0	104.8	99.9
	H	30.6	30.5	30.7	30.8	30.6	30.5
	F	424.0	423.5	425.4	423.0	424.9	430.0
FCN	F	375.1	374.1	379.2	377.3	376.4	383.7
	C	84.8	82.2	82.0	89.1	86.6	83.7
	N	120.1	117.9	114.1	128.8	122.1	133.1
H_2S	S	759.6	739.0	744.2	752.4	743.9	775.1
	H	30.7	30.5	30.9	31.0	30.9	30.5
HCP	H	29.7	29.6	30.0	29.9	29.9	29.5
	C	41.1	37.6	34.0	49.5	46.4	45.8
	P	426.9	388.0	408.5	443.1	432.9	432.6
HFCO	O	-91.0	-94.3	-85.7	-59.0	-79.2	-60.2
	C	43.1	39.6	41.1	49.2	44.9	43.4
	F	166.9	165.3	183.7	183.9	177.2	166.6
	H	24.0	23.9	24.3	24.3	24.2	23.9
H_2C_2O	C	194.4	193.3	192.0	193.6	192.4	199.1
	C	-2.0	-6.3	-3.9	6.1	-0.7	1.7
	O	-4.3	-5.9	3.1	16.6	4.7	29.1
	H	29.3	29.2	29.5	29.6	29.4	29.4
LiF	Li	89.3	89.3	89.7	89.7	89.5	89.3
	F	375.8	382.5	383.6	383.8	387.9	373.9
LiH	H	26.6	26.6	26.9	26.6	26.8	26.7
	Li	89.3	89.3	89.2	90.2	89.4	89.3
N_2O	N	109.4	106.4	102.0	120.1	112.3	132.0
	N	16.0	12.6	7.5	28.0	23.1	33.6
	O	199.9	199.0	199.3	208.3	201.7	217.4
OCS	O	98.7	96.8	100.3	111.7	108.8	111.9
	C	33.8	30.2	27.9	42.5	39.5	40.7
	S	815.1	796.7	814.1	819.1	810.1	833.9
OF_2	O	-430.3	-447.1	-406.1	-260.8	-374.3	-451.6
	F	-12.7	-24.0	14.0	34.7	-0.5	-3.4
H_4C_2O	O	364.7	363.2	366.5	360.2	360.7	373.5
	C	155.1	153.2	154.5	156.4	155.1	154.2
	H	29.2	29.1	29.5	29.4	29.4	29.2
PN	N	-340.7	-344.0	-353.8	-276.3	-306.7	-252.2
	P	118.2	50.6	84.8	191.2	149.7	176.7

Table 16: Statistics for absolute isotropic NMR shieldings in a cc-pVQZ basis. CCSD(T)/aug-cc-pCV[TQ]Z shieldings are used as a reference. MSD=mean signed deviation; MAD=mean absolute deviation; MRD=mean relative deviation [%]; MARD=mean absolute relative deviation [%]; RMSD=root mean squared deviation.

	MSD	MAD	MRD	MARD	RMSD
RHF	-9.2	17.3	-49.9	59.8	32.9
PBE	-19.6	19.8	-58.9	59.3	32.6
KT2	-5.3	9.4	5.8	18.4	17.3
B3LYP	-20.0	20.2	-69.9	70.4	30.7
CCSD	3.4	4.9	5.7	11.1	10.2
CCSD(T)	4.1	4.3	13.9	14.0	10.8
RPA@HF	4.1	5.6	9.0	15.4	10.8
RPA@BH&H	8.5	8.8	30.7	30.8	17.3
RPA@B3LYP	12.3	12.6	48.3	48.4	25.0
RPA@KT2	13.1	13.8	51.7	51.9	27.7
RPA@PBE	16.9	17.3	65.7	65.8	36.7
σ-P1@PBE	9.0	9.3	35.5	35.6	19.2
MP2	11.3	11.8	35.8	36.0	23.8

Table 17: Statistics for absolute isotropic NMR shieldings in a cc-pVQZ basis. CCSD(T) shieldings are used as a reference. MSD=mean signed deviation; MAD=mean absolute deviation; MRD=mean relative deviation [%]; MARD=mean absolute relative deviation [%]; RMSD=root mean squared deviation.

	MSD	MAD	MRD	MARD	RMSD
RHF	-13.4	20.1	-43.5	56.2	37.9
PBE	-23.8	23.8	-52.5	52.8	38.5
KT2	-9.4	10.6	-4.5	17.9	19.4
B3LYP	-24.1	24.2	-63.9	64.1	37.3
CCSD	-0.7	3.4	-5.0	9.2	6.4
CCSD(T)	0.0	0.0	0.0	0.0	0.0
RPA@HF	-0.0	5.1	0.3	13.2	9.1
RPA@BH&H	4.4	5.3	16.4	16.7	11.6
RPA@B3LYP	8.2	9.0	28.3	28.5	18.8
RPA@KT2	9.0	10.0	30.5	30.8	20.5
RPA@PBE	12.8	13.6	39.0	39.2	30.3
σ-P1@PBE	4.9	6.1	14.3	14.6	12.0
MP2	7.2	8.3	24.0	24.6	16.8

Table 18: Absolute isotropic NMR shieldings in a cc-pCVDZ basis. All values are given in ppm. “Extrap.” stands for basis-set extrapolated CCSD(T)-shieldings in an aug-cc-pCV[TQ]Z basis set. The data for “Extrap.” was taken from Ref. 3. CCSD(T) calculations were done with the CFOUR program[1, 2].

Mol.	Nuc.	CCSD(T)	Extrap.	RPA@HF	RPA@PBE	σ-P1@PBE
HF	H	30.5	28.8	30.4	31.1	31.4
	F	426.0	420.3	426.2	424.0	420.7
CO	C	28.9	2.2	27.0	52.2	42.8
	O	-24.3	-55.0	-18.2	12.1	-7.3
N_2	N	-32.8	-60.4	-32.7	-1.8	-18.4
H_2O	O	353.7	338.0	352.5	350.5	346.2
	H	31.9	30.6	31.9	32.3	32.5
HCN	H	29.4	29.0	29.5	29.6	29.6
	C	106.3	84.6	103.6	116.0	113.0
	N	13.9	-14.1	12.8	39.4	31.5
HOF	O	-12.4	-68.9	2.6	72.9	38.1
	H	21.8	19.6	22.6	24.5	23.8
	F	179.7	192.2	204.4	219.7	210.3
NH_3	N	282.2	270.7	280.1	279.5	276.8
	H	32.3	31.4	32.4	32.5	32.6
CH_2O	O	-352.1	-378.6	-321.4	-212.0	-268.7
	C	32.3	1.5	32.7	56.4	51.0
	H	22.6	22.0	23.1	23.7	23.4
CH_4	C	206.8	198.9	204.9	206.4	204.5
	H	31.6	31.3	31.7	31.7	31.7
C2H4	C	90.3	69.7	87.2	99.2	98.3
	H	26.6	26.0	26.7	27.0	27.0
AlF	Al	578.6	572.9	584.9	592.8	592.1
	F	186.3	211.8	207.7	203.7	199.8
CH_3F	C	139.1	122.1	139.3	144.1	142.7
	F	475.6	482.9	476.4	469.8	470.1
	H	27.9	27.3	28.1	28.2	28.3
C_3H_4	C	202.2	192.1	201.2	203.2	202.9
	C	102.7	83.7	100.9	111.8	109.9
	H	25.0	24.4	25.2	25.6	25.5
	H	31.1	30.6	31.3	31.3	31.3
FCCH	C	189.6	179.9	188.0	192.3	193.2
	C	120.7	100.1	120.5	125.6	125.0
	H	31.0	30.5	31.0	31.2	31.2
	F	428.0	423.5	431.4	429.6	430.0
FCN	F	382.5	374.1	389.3	388.5	384.2
	C	105.7	82.2	103.7	111.9	110.4
	N	135.8	117.9	133.3	149.8	147.3
H_2S	S	756.4	739.0	750.5	759.2	745.5
	H	31.3	30.5	31.4	31.6	31.8
HCP	H	29.4	29.6	29.6	29.6	29.7
	C	61.6	37.6	55.5	74.2	72.3
	P	422.0	388.0	407.3	465.8	462.4
HFCO	O	-67.1	-94.3	-56.5	-22.1	-40.7
	C	64.0	39.6	63.1	73.9	71.1
	F	183.1	165.3	200.7	204.0	193.1
	H	24.7	23.9	24.9	25.0	24.9
H_2C_2O	C	202.2	193.3	201.4	203.1	202.3
	C	21.2	-6.3	21.8	33.6	26.6
	O	14.8	-5.9	29.1	44.6	27.7
	H	30.0	29.2	30.2	30.3	30.4
LiF	Li	93.0	89.3	93.2	93.4	93.5
	F	348.3	382.5	358.8	352.2	354.5
LiH	H	26.3	26.6	26.5	26.2	26.3
	Li	91.7	89.3	91.8	93.2	93.4
N_2O	N	125.2	106.4	123.3	140.8	135.9
	N	41.3	12.6	36.6	57.0	51.0
	O	206.2	199.0	211.6	220.3	213.7
OCS	O	117.0	96.8	121.8	135.6	127.3
	C	54.4	30.2	50.6	66.5	63.2
	S	817.4	796.7	823.6	836.3	830.8
OF_2	O	-349.1	-447.1	-314.6	-132.6	-234.1
	F	13.9	-24.0	45.6	77.1	55.3
H_4C_2O	O	364.0	363.2	368.9	364.1	363.5
	C	167.5	153.2	167.5	170.8	170.1
	H	29.7	29.1	30.0	29.9	30.0
PN	N	-295.2	-344.0	-289.6	-200.5	-252.9
	P	111.2	50.6	100.4	235.3	168.0

Table 19: Statistics for absolute isotropic NMR shieldings in a cc-pCVDZ basis. CCSD(T)/aug-cc-pCV[TQ]Z shieldings are used as a reference. MSD=mean signed deviation; MAD=mean absolute deviation; MRD=mean relative deviation [%]; MARD=mean absolute relative deviation [%]; RMSD=root mean squared deviation.

	MSD	MAD	MRD	MARD	RMSD
RHF	-4.0	17.1	-6.4	43.2	34.8
CCSD	14.0	15.4	74.2	74.7	23.6
CCSD(T)	14.2	16.6	81.1	82.0	23.8
RPA@HF	17.1	18.1	85.9	86.2	28.2
RPA@PBE	33.1	34.6	152.9	153.4	62.4
σ-P1@PBE	25.1	26.6	126.8	127.3	44.5

Table 20: Statistics for absolute isotropic NMR shieldings in a cc-pCVDZ basis. CCSD(T) shieldings are used as a reference. MSD=mean signed deviation; MAD=mean absolute deviation; MRD=mean relative deviation [%]; MARD=mean absolute relative deviation [%]; RMSD=root mean squared deviation.

	MSD	MAD	MRD	MARD	RMSD
RHF	-18.2	24.9	-35.4	45.2	47.5
CCSD	-0.3	3.0	-0.5	4.6	5.7
CCSD(T)	0.0	0.0	0.0	0.0	0.0
RPA@HF	2.9	5.0	7.4	9.6	9.4
RPA@PBE	18.8	19.3	37.9	38.0	41.3
σ-P1@PBE	10.9	12.0	22.5	22.8	23.1

Table 21: Absolute isotropic NMR shieldings in a cc-pCVTZ basis. All values are given in ppm. "Extrap." stands for basis-set extrapolated CCSD(T)-shieldings in an aug-cc-pCV[TQ]Z basis set. The data for "Extrap." was taken from Ref. 3. CCSD(T) calculations were done with the CFOUR program[1, 2].

Mol.	Nuc.	CCSD(T)	Extrap.	RPA@HF	RPA@PBE	σ-P1@PBE
HF	H	29.4	28.8	29.4	30.0	29.9
	F	420.4	420.3	419.5	418.2	415.9
CO	C	8.0	2.2	2.2	26.0	16.1
	O	-52.0	-55.0	-51.3	-24.0	-39.8
N_2	N	-56.3	-60.4	-62.9	-32.3	-47.5
H_2O	O	344.5	338.0	341.8	341.1	338.3
	H	31.2	30.6	31.3	31.6	31.5
HCN	H	29.2	29.0	29.4	29.4	29.5
	C	89.5	84.6	85.5	97.1	93.7
	N	-11.0	-14.1	-15.8	9.4	0.5
HOF	O	-49.0	-68.9	-46.1	20.2	-24.8
	H	20.6	19.6	21.2	22.8	21.7
	F	184.9	192.2	210.2	210.7	185.5
NH_3	N	275.5	270.7	272.1	272.4	270.5
	H	32.0	31.4	32.1	32.2	32.1
CH_2O	O	-383.6	-378.6	-359.7	-272.4	-340.4
	C	10.1	1.5	9.2	29.7	22.2
	H	22.3	22.0	22.8	23.2	22.9
CH_4	C	200.3	198.9	197.5	198.6	197.6
	H	31.5	31.3	31.6	31.6	31.6
C2H4	C	74.4	69.7	70.5	80.9	78.7
	H	26.3	26.0	26.5	26.7	26.7
AlF	Al	573.6	572.9	578.4	582.4	578.8
	F	197.8	211.8	215.8	214.6	210.7
CH_3F	C	125.8	122.1	125.8	129.1	127.8
	F	480.3	482.9	478.5	473.0	475.7
	H	27.6	27.3	28.0	28.0	28.0
C_3H_4	C	193.8	192.1	192.1	193.4	193.3
	C	87.7	83.7	84.5	94.3	92.0
	H	24.6	24.4	24.8	25.1	25.0
	H	30.9	30.6	31.1	31.0	31.1
FCCH	C	182.1	179.9	179.9	183.4	182.5
	C	103.6	100.1	103.3	107.3	106.7
	H	30.7	30.5	30.8	30.9	30.9
	F	422.0	423.5	423.7	421.8	421.4
FCN	F	373.2	374.1	377.3	375.9	371.4
	C	86.6	82.2	83.8	91.4	89.8
	N	120.3	117.9	114.5	130.1	124.5
H_2S	S	739.0	739.0	725.1	734.4	722.5
	H	30.8	30.5	30.9	31.1	31.2
HCP	H	29.6	29.6	30.0	29.8	30.0
	C	42.3	37.6	35.0	51.8	50.3
	P	393.6	388.0	376.2	426.9	417.7
HFCO	O	-92.3	-94.3	-87.8	-58.6	-81.3
	C	45.0	39.6	43.1	52.1	48.7
	F	166.3	165.3	182.9	183.2	172.3
	H	24.3	23.9	24.6	24.6	24.5
H_2C_2O	C	195.0	193.3	192.8	194.3	192.9
	C	0.4	-6.3	-1.2	9.4	2.7
	O	-6.9	-5.9	0.6	15.0	-1.0
	H	29.5	29.2	29.7	29.8	29.7
LiF	Li	90.3	89.3	90.8	91.0	91.5
	F	364.2	382.5	373.2	373.7	377.6
LiH	H	26.5	26.6	26.8	26.5	26.7
	Li	89.9	89.3	90.1	91.3	91.6
N_2O	N	111.1	106.4	103.7	122.6	114.7
	N	19.0	12.6	9.9	31.5	26.2
	O	196.6	199.0	196.0	206.1	197.7
OCS	O	96.6	96.8	97.8	110.3	105.0
	C	35.4	30.2	29.4	44.9	42.2
	S	800.2	796.7	800.5	810.1	799.2
OF_2	O	-412.4	-447.1	-387.7	-239.0	-352.7
	F	-6.2	-24.0	21.2	43.1	4.0
H_4C_2O	O	362.4	363.2	365.1	359.2	359.4
	C	156.3	153.2	155.9	158.0	157.0
	H	29.4	29.1	29.7	29.6	29.7
PN	N	-342.9	-344.0	-359.2	-273.0	-304.1
	P	59.8	50.6	27.1	151.2	105.8

Table 22: Statistics for absolute isotropic NMR shieldings in a cc-pCVTZ basis. CCSD(T)/aug-cc-pCV[TQ]Z shieldings are used as a reference. MSD=mean signed deviation; MAD=mean absolute deviation; MRD=mean relative deviation [%]; MARD=mean absolute relative deviation [%]; RMSD=root mean squared deviation.

	MSD	MAD	MRD	MARD	RMSD
RHF	-10.8	18.7	-46.0	57.1	36.9
CCSD	1.5	4.5	9.7	15.2	9.7
CCSD(T)	2.3	3.9	18.2	19.3	6.9
RPA@HF	2.4	5.4	13.7	17.1	11.3
RPA@PBE	16.5	17.4	74.2	74.5	36.9
σ-P1@PBE	7.9	9.5	45.2	45.6	17.9

Table 23: Statistics for absolute isotropic NMR shieldings in a cc-pCVTZ basis. CCSD(T) shieldings are used as a reference. MSD=mean signed deviation; MAD=mean absolute deviation; MRD=mean relative deviation [%]; MARD=mean absolute relative deviation [%]; RMSD=root mean squared deviation.

	MSD	MAD	MRD	MARD	RMSD
RHF	-13.1	20.3	-66.0	90.4	38.6
CCSD	-0.8	3.5	-7.4	14.9	6.6
CCSD(T)	0.0	0.0	0.0	0.0	0.0
RPA@HF	0.1	5.2	-1.4	20.0	9.1
RPA@PBE	14.2	15.0	70.1	70.3	32.3
σ-P1@PBE	5.6	7.0	22.4	22.8	13.4

Table 24: Absolute isotropic NMR shieldings in a cc-pCVQZ basis. All values are given in ppm. “Extrap.” stands for basis-set extrapolated CCSD(T)-shieldings in an aug-cc-pCV[TQ]Z basis set. The data for “Extrap.” was taken from Ref. 3. CCSD(T) calculations were done with the CFOUR program[1, 2].

Mol.	Nuc.	CCSD(T)	Extrap.	RPA@HF	RPA@PBE	σ-P1@PBE
HF	H	29.1	28.8	29.0	29.6	29.1
	F	419.7	420.3	419.0	418.3	418.0
CO	C	3.9	2.2	-3.0	20.6	9.7
	O	-54.9	-55.0	-54.5	-27.9	-41.7
N_2	N	-60.2	-60.4	-67.7	-37.2	-52.0
H_2O	O	340.6	338.0	338.0	338.0	337.3
	H	30.9	30.6	30.9	31.2	30.9
HCN	H	29.1	29.0	29.3	29.3	29.2
	C	86.0	84.6	81.7	93.1	88.9
	N	-13.6	-14.1	-18.7	6.1	-3.3
HOF	O	-61.8	-68.9	-61.6	5.9	-39.7
	H	20.0	19.6	20.5	22.0	20.6
	F	189.1	192.2	213.4	211.6	191.9
NH_3	N	272.7	270.7	269.4	270.2	269.5
	H	31.6	31.4	31.7	31.8	31.6
CH_2O	O	-382.9	-378.6	-357.5	-275.8	-336.7
	C	4.3	1.5	3.1	22.8	13.5
	H	22.1	22.0	22.6	23.0	22.6
CH_4	C	199.2	198.9	196.3	197.6	197.3
	H	31.4	31.3	31.5	31.4	31.4
C2H4	C	71.0	69.7	66.9	77.1	74.1
	H	26.1	26.0	26.3	26.5	26.4
AlF	Al	573.2	572.9	576.7	579.8	576.2
	F	206.8	211.8	223.6	223.5	225.1
CH_3F	C	123.2	122.1	123.3	126.5	125.1
	F	482.4	482.9	479.7	475.2	480.3
	H	27.4	27.3	27.8	27.7	27.7
C_3H_4	C	192.5	192.1	190.6	191.9	191.8
	C	84.8	83.7	81.2	90.9	88.2
	H	24.4	24.4	24.6	24.9	24.7
	H	30.7	30.6	30.9	30.9	30.8
FCCH	C	180.5	179.9	178.1	181.3	179.8
	C	101.0	100.1	100.7	104.7	103.8
	H	30.5	30.5	30.6	30.8	30.7
	F	422.6	423.5	423.7	422.3	423.9
FCN	F	373.6	374.1	377.1	376.0	374.4
	C	83.4	82.2	80.5	87.9	85.6
	N	118.4	117.9	112.1	127.3	120.8
H_2S	S	737.9	739.0	722.8	733.1	725.5
	H	30.6	30.5	30.7	30.8	30.7
HCP	H	29.6	29.6	29.9	29.8	29.8
	C	38.8	37.6	31.1	47.6	44.4
	P	390.5	388.0	373.0	421.9	409.0
HFCO	O	-94.7	-94.3	-90.2	-62.5	-82.4
	C	41.2	39.6	38.9	47.6	43.4
	F	164.9	165.3	181.4	182.2	175.3
	H	24.0	23.9	24.3	24.3	24.2
H_2C_2O	C	193.7	193.3	191.2	192.8	191.6
	C	-4.4	-6.3	-6.6	4.0	-2.9
	O	-7.2	-5.9	-0.3	13.7	1.7
	H	29.3	29.2	29.5	29.5	29.3
LiF	Li	89.6	89.3	90.1	90.3	90.9
	F	375.0	382.5	383.3	385.0	392.4
LiH	H	26.6	26.6	26.9	26.6	26.8
	Li	89.5	89.3	89.7	90.7	90.8
N_2O	N	107.9	106.4	99.8	118.7	110.8
	N	14.0	12.6	4.7	26.1	21.1
	O	197.5	199.0	195.9	206.0	199.0
OCS	O	96.2	96.8	97.1	109.2	106.1
	C	31.5	30.2	25.2	40.5	37.4
	S	798.5	796.7	797.4	806.0	795.7
OF_2	O	-435.0	-447.1	-411.6	-265.2	-378.0
	F	-15.7	-24.0	10.7	31.5	-3.7
H_4C_2O	O	363.3	363.2	364.9	359.2	359.8
	C	154.0	153.2	153.4	155.6	154.3
	H	29.2	29.1	29.5	29.4	29.3
PN	N	-345.5	-344.0	-362.3	-279.8	-309.9
	P	55.3	50.6	21.7	140.9	96.2

Table 25: Statistics for absolute isotropic NMR shieldings in a cc-pCVQZ basis. CCSD(T)/aug-cc-pCV[TQ]Z shieldings are used as a reference. MSD=mean signed deviation; MAD=mean absolute deviation; MRD=mean relative deviation [%]; MARD=mean absolute relative deviation [%]; RMSD=root mean squared deviation.

	MSD	MAD	MRD	MARD	RMSD
RHF	-12.1	19.2	-55.1	64.7	37.2
CCSD	-0.2	3.7	-3.1	10.2	7.4
CCSD(T)	0.6	1.4	5.4	6.4	2.6
RPA@HF	0.4	4.9	-0.8	13.4	9.3
RPA@PBE	14.1	14.8	58.3	58.5	32.5
σ-P1@PBE	6.2	7.0	28.0	28.2	14.0

Table 26: Statistics for absolute isotropic NMR shieldings in a cc-pCVQZ basis. CCSD(T) shieldings are used as a reference. MSD=mean signed deviation; MAD=mean absolute deviation; MRD=mean relative deviation [%]; MARD=mean absolute relative deviation [%]; RMSD=root mean squared deviation.

	MSD	MAD	MRD	MARD	RMSD
RHF	-12.7	19.7	-41.9	53.3	37.6
CCSD	-0.8	3.5	-5.4	9.1	6.6
CCSD(T)	0.0	0.0	0.0	0.0	0.0
RPA@HF	-0.2	5.2	-2.9	12.7	9.1
RPA@PBE	13.5	14.3	37.5	37.7	31.1
σ-P1@PBE	5.6	6.6	15.1	15.4	12.9

Table 27: Absolute isotropic NMR shieldings in a def2-SVP basis. All values are given in ppm. “Extrap.” stands for basis-set extrapolated CCSD(T)-shieldings in an aug-cc-pCV[TQ]Z basis set. The data for “Extrap.” was taken from Ref. 3. CCSD(T) calculations were done with the CFOUR program[1, 2].

Mol.	Nuc.	CCSD(T)	Extrap.	RPA@HF	RPA@PBE
HF	H	30.1	28.8	30.0	30.6
	F	430.6	420.3	430.7	428.6
CO	C	32.2	2.2	28.8	52.7
	O	-10.5	-55.0	-5.4	23.6
N_2	N	-20.8	-60.4	-22.0	7.0
H_2O	O	355.2	338.0	353.9	352.0
	H	31.7	30.6	31.7	32.1
HCN	H	29.8	29.0	29.9	30.0
	C	108.4	84.6	105.2	116.7
	N	17.9	-14.1	16.1	41.7
HOF	O	-4.6	-68.9	6.9	73.4
	H	21.7	19.6	22.4	24.2
	F	193.0	192.2	217.0	231.8
NH_3	N	280.7	270.7	278.4	278.0
	H	32.3	31.4	32.4	32.5
CH_2O	O	-338.7	-378.6	-307.9	-202.0
	C	33.3	1.5	32.6	54.7
	H	23.1	22.0	23.5	24.2
CH_4	C	205.5	198.9	203.3	204.2
	H	31.6	31.3	31.7	31.7
C2H4	C	92.3	69.7	88.1	99.1
	H	26.8	26.0	26.9	27.2
AlF	Al	583.4	572.9	587.7	592.2
	F	205.1	211.8	223.9	224.6
CH_3F	C	138.4	122.1	138.2	141.9
	F	481.2	482.9	481.3	475.4
	H	28.0	27.3	28.3	28.4
C_3H_4	C	202.9	192.1	201.6	203.2
	C	104.8	83.7	102.3	112.1
	H	25.2	24.4	25.4	25.8
	H	31.3	30.6	31.5	31.5
FCCH	C	188.9	179.9	186.6	190.3
	C	121.7	100.1	121.0	125.2
	H	31.2	30.5	31.4	31.5
	F	438.9	423.5	441.2	441.2
FCN	F	398.3	374.1	403.8	404.5
	C	105.3	82.2	103.0	110.2
	N	134.1	117.9	130.3	146.4
H_2S	S	786.6	739.0	780.0	787.0
	H	31.5	30.5	31.6	31.8
HCP	H	29.8	29.6	30.1	30.0
	C	63.0	37.6	55.9	72.8
	P	462.9	388.0	445.7	500.7
HFCO	O	-58.8	-94.3	-49.6	-16.2
	C	64.6	39.6	62.9	72.3
	F	196.8	165.3	212.8	217.1
	H	25.1	23.9	25.4	25.5
H_2C_2O	C	203.9	193.3	202.7	204.2
	C	21.3	-6.3	21.4	31.8
	O	21.3	-5.9	36.4	51.1
	H	30.1	29.2	30.3	30.4
LiF	Li	89.3	89.3	89.7	89.6
	F	338.0	382.5	349.7	338.9
LiH	H	26.7	26.6	26.9	26.6
	Li	90.3	89.3	90.4	92.4
N_2O	N	133.5	106.4	131.0	147.5
	N	44.5	12.6	39.8	58.8
	O	224.7	199.0	229.9	237.2
OCS	O	131.8	96.8	135.9	148.9
	C	54.7	30.2	50.4	65.0
	S	849.1	796.7	852.3	865.3
OF_2	O	-329.6	-447.1	-300.0	-132.4
	F	27.3	-24.0	58.0	88.7
H_4C_2O	O	368.4	363.2	372.4	367.4
	C	166.9	153.2	166.6	169.1
	H	29.8	29.1	30.1	30.1
PN	N	-282.6	-344.0	-278.6	-194.6
	P	183.8	50.6	167.1	292.3

Table 28: Statistics for absolute isotropic NMR shieldings in a def2-SVP basis. CCSD(T)/aug-cc-pCV[TQ]Z shieldings are used as a reference. MSD=mean signed deviation; MAD=mean absolute deviation; MRD=mean relative deviation [%]; MARD=mean absolute relative deviation [%]; RMSD=root mean squared deviation.

	MSD	MAD	MRD	MARD	RMSD
CCSD(T)	20.4	22.0	92.2	92.7	33.7
RPA@HF	22.5	23.6	94.3	94.6	36.4
RPA@PBE	37.5	39.0	157.1	157.5	68.5

Table 29: Statistics for absolute isotropic NMR shieldings in a def2-SVP basis. CCSD(T) shieldings are used as a reference. MSD=mean signed deviation; MAD=mean absolute deviation; MRD=mean relative deviation [%]; MARD=mean absolute relative deviation [%]; RMSD=root mean squared deviation.

	MSD	MAD	MRD	MARD	RMSD
CCSD(T)	0.0	0.0	0.0	0.0	0.0
RPA@HF	2.1	4.9	7.0	9.9	9.0
RPA@PBE	17.1	17.6	48.9	49.1	38.1

Table 30: Absolute isotropic NMR shieldings in a def2-TZVP basis. All values are given in ppm. "Extrap." stands for basis-set extrapolated CCSD(T)-shieldings in an aug-cc-pCV[TQ]Z basis set. The data for "Extrap." was taken from Ref. 3. CCSD(T) calculations were done with the CFOUR program[1, 2].

Mol.	Nuc.	CCSD(T)	Extrap.	RPA@HF	RPA@PBE
HF	H	29.7	28.8	29.6	30.2
	F	420.8	420.3	420.7	420.1
CO	C	10.4	2.2	4.8	27.7
	O	-50.9	-55.0	-49.2	-21.4
N_2	N	-52.2	-60.4	-58.1	-28.4
H_2O	O	343.3	338.0	341.7	341.9
	H	31.6	30.6	31.6	31.9
HCN	H	29.3	29.0	29.4	29.5
	C	90.3	84.6	86.2	97.0
	N	-8.3	-14.1	-12.8	12.8
HOF	O	-53.7	-68.9	-50.1	17.1
	H	20.9	19.6	21.4	23.0
	F	188.0	192.2	213.4	213.4
NH_3	N	275.8	270.7	273.3	274.2
	H	32.2	31.4	32.3	32.4
CH_2O	O	-374.8	-378.6	-350.4	-266.7
	C	10.7	1.5	9.3	28.2
	H	22.4	22.0	22.9	23.3
CH_4	C	201.1	198.9	198.7	200.3
	H	31.7	31.3	31.8	31.8
C2H4	C	75.8	69.7	71.6	81.5
	H	26.5	26.0	26.6	26.8
AlF	Al	574.6	572.9	579.9	584.5
	F	203.3	211.8	220.7	222.0
CH_3F	C	126.5	122.1	126.6	129.8
	F	482.1	482.9	480.6	476.0
	H	27.8	27.3	28.1	28.1
C_3H_4	C	195.3	192.1	193.8	195.6
	C	89.4	83.7	86.0	96.0
	H	24.8	24.4	24.9	25.2
	H	31.0	30.6	31.2	31.2
FCCH	C	184.5	179.9	182.2	185.4
	C	104.5	100.1	103.9	107.8
	H	30.9	30.5	31.0	31.1
	F	426.0	423.5	427.4	427.9
FCN	F	374.1	374.1	378.8	378.8
	C	87.3	82.2	84.5	91.6
	N	123.3	117.9	117.9	133.6
H_2S	S	729.2	739.0	718.4	729.4
	H	31.0	30.5	31.2	31.3
HCP	H	29.7	29.6	30.0	29.9
	C	44.1	37.6	36.7	52.4
	P	381.4	388.0	363.4	414.4
HFCO	O	-88.3	-94.3	-83.0	-53.8
	C	45.8	39.6	43.8	52.3
	F	165.7	165.3	182.3	184.1
	H	24.4	23.9	24.7	24.7
H_2C_2O	C	197.1	193.3	195.3	197.0
	C	1.8	-6.3	0.4	10.9
	O	-2.6	-5.9	6.5	22.3
	H	29.8	29.2	29.9	30.0
LiF	Li	90.6	89.3	91.0	91.3
	F	378.0	382.5	386.8	392.1
LiH	H	26.6	26.6	26.9	26.6
	Li	91.1	89.3	91.2	93.2
N_2O	N	114.1	106.4	107.7	125.8
	N	18.7	12.6	10.8	31.4
	O	200.6	199.0	201.2	211.3
OCS	O	100.4	96.8	102.6	115.4
	C	36.4	30.2	31.0	45.7
	S	797.7	796.7	799.4	808.0
OF_2	O	-419.7	-447.1	-393.0	-242.5
	F	-12.8	-24.0	15.5	37.6
H_4C_2O	O	366.1	363.2	368.8	363.5
	C	157.3	153.2	156.8	159.1
	H	29.6	29.1	29.8	29.7
PN	N	-336.2	-344.0	-347.2	-267.4
	P	40.0	50.6	11.8	134.6

Table 31: Statistics for absolute isotropic NMR shieldings in a def2-TZVP basis. CCSD(T)/aug-cc-pCV[TQ]Z shieldings are used as a reference. MSD=mean signed deviation; MAD=mean absolute deviation; MRD=mean relative deviation [%]; MARD=mean absolute relative deviation [%]; RMSD=root mean squared deviation.

	MSD	MAD	MRD	MARD	RMSD
CCSD(T)	2.9	4.2	22.0	22.9	6.1
RPA@HF	3.5	6.2	18.1	21.1	12.0
RPA@PBE	17.6	18.1	76.6	76.7	36.5

Table 32: Statistics for absolute isotropic NMR shieldings in a def2-TZVP basis. CCSD(T) shieldings are used as a reference. MSD=mean signed deviation; MAD=mean absolute deviation; MRD=mean relative deviation [%]; MARD=mean absolute relative deviation [%]; RMSD=root mean squared deviation.

	MSD	MAD	MRD	MARD	RMSD
CCSD(T)	0.0	0.0	0.0	0.0	0.0
RPA@HF	0.6	5.0	3.7	15.5	8.9
RPA@PBE	14.7	15.1	48.8	49.0	32.7

Table 33: Absolute isotropic NMR shieldings in a def2-QZVP basis. All values are given in ppm. "Extrap." stands for basis-set extrapolated CCSD(T)-shieldings in an aug-cc-pCV[TQ]Z basis set. The data for "Extrap." was taken from Ref. 3. CCSD(T) calculations were done with the CFOUR program[1, 2].

Mol.	Nuc.	CCSD(T)	Extrap.	RPA@HF	RPA@PBE
HF	H	29.0	28.8	28.9	29.6
	F	419.2	420.3	418.7	418.9
CO	C	4.4	2.2	-2.2	20.8
	O	-54.9	-55.0	-53.9	-26.9
N_2	N	-60.0	-60.4	-66.9	-36.7
H_2O	O	337.8	338.0	335.6	336.9
	H	30.8	30.6	30.9	31.2
HCN	H	29.1	29.0	29.3	29.3
	C	86.3	84.6	82.1	93.5
	N	-12.4	-14.1	-17.2	7.5
HOF	O	-63.3	-68.9	-61.8	5.0
	H	19.9	19.6	20.5	22.0
	F	190.6	192.2	213.8	211.5
NH_3	N	271.1	270.7	268.0	269.5
	H	31.6	31.4	31.7	31.8
CH_2O	O	-378.1	-378.6	-351.8	-272.6
	C	4.9	1.5	4.0	23.1
	H	22.1	22.0	22.6	23.0
CH_4	C	199.7	198.9	196.7	197.4
	H	31.4	31.3	31.5	31.5
C2H4	C	72.1	69.7	68.0	77.8
	H	26.1	26.0	26.4	26.5
AlF	Al	572.1	572.9	575.6	579.4
	F	211.2	211.8	227.8	227.6
CH_3F	C	124.0	122.1	124.0	126.0
	F	482.3	482.9	479.8	474.8
	H	27.4	27.3	27.8	27.8
C_3H_4	C	193.1	192.1	191.1	192.3
	C	85.8	83.7	82.3	92.5
	H	24.5	24.4	24.6	24.9
	H	30.7	30.6	30.9	30.9
FCCH	C	181.0	179.9	178.5	181.3
	C	101.2	100.1	100.9	104.3
	H	30.6	30.5	30.7	30.8
	F	423.7	423.5	424.9	422.8
FCN	F	374.7	374.1	378.7	377.6
	C	83.3	82.2	80.5	87.4
	N	119.4	117.9	113.2	128.0
H_2S	S	737.5	739.0	722.1	732.3
	H	30.6	30.5	30.8	30.9
HCP	H	29.6	29.6	29.9	29.8
	C	39.7	37.6	32.2	48.0
	P	385.9	388.0	368.8	417.8
HFCO	O	-93.1	-94.3	-88.1	-60.8
	C	41.7	39.6	39.6	47.3
	F	165.2	165.3	182.0	182.2
	H	24.0	23.9	24.3	24.3
H_2C_2O	C	194.3	193.3	191.8	193.7
	C	-4.1	-6.3	-6.1	4.1
	O	-6.1	-5.9	1.2	15.7
	H	29.3	29.2	29.5	29.6
LiF	Li	90.3	89.3	90.8	91.0
	F	382.2	382.5	390.2	392.6
LiH	H	26.6	26.6	26.9	26.6
	Li	90.0	89.3	90.2	91.3
N_2O	N	108.3	106.4	100.7	119.2
	N	13.5	12.6	4.9	26.0
	O	199.0	199.0	198.3	207.9
OCS	O	97.2	96.8	98.6	111.1
	C	31.7	30.2	25.8	40.5
	S	798.6	796.7	798.4	806.6
OF_2	O	-438.0	-447.1	-412.7	-267.5
	F	-19.3	-24.0	7.6	28.3
H_4C_2O	O	363.7	363.2	365.2	359.2
	C	154.7	153.2	154.0	155.5
	H	29.2	29.1	29.5	29.4
PN	N	-341.8	-344.0	-356.5	-278.4
	P	47.5	50.6	16.1	134.7

Table 34: Statistics for absolute isotropic NMR shieldings in a def2-QZVP basis. CCSD(T)/aug-cc-pCV[TQ]Z shieldings are used as a reference. MSD=mean signed deviation; MAD=mean absolute deviation; MRD=mean relative deviation [%]; MARD=mean absolute relative deviation [%]; RMSD=root mean squared deviation.

	MSD	MAD	MRD	MARD	RMSD
CCSD(T)	0.8	1.2	6.6	6.9	1.9
RPA@HF	0.8	5.2	1.4	13.9	9.7
RPA@PBE	14.3	15.0	59.2	59.4	32.3

Table 35: Statistics for absolute isotropic NMR shieldings in a def2-QZVP basis. CCSD(T) shieldings are used as a reference. MSD=mean signed deviation; MAD=mean absolute deviation; MRD=mean relative deviation [%]; MARD=mean absolute relative deviation [%]; RMSD=root mean squared deviation.

	MSD	MAD	MRD	MARD	RMSD
CCSD(T)	0.0	0.0	0.0	0.0	0.0
RPA@HF	0.0	5.1	-2.3	12.1	9.0
RPA@PBE	13.5	14.2	36.8	37.0	31.2

Table 36: Absolute isotropic NMR shieldings in a pcSseg-0 basis. All values are given in ppm. “Extrap.” stands for basis-set extrapolated CCSD(T)-shieldings in an aug-cc-pCV[TQ]Z basis set. The data for “Extrap.” was taken from Ref. 3. CCSD(T) calculations were done with the CFOUR program[1, 2].

Mol.	Nuc.	CCSD(T)	Extrap.	RPA@HF	RPA@PBE
HF	H	30.8	28.8	30.7	30.9
	F	430.9	420.3	432.5	432.1
CO	C	7.0	2.2	11.6	43.2
	O	-77.8	-55.0	-63.6	-19.8
N_2	N	-74.3	-60.4	-63.1	-25.2
H_2O	O	355.8	338.0	356.3	355.3
	H	32.8	30.6	32.8	32.8
HCN	H	30.4	29.0	30.5	30.5
	C	87.6	84.6	87.1	102.5
	N	-14.7	-14.1	-8.9	26.5
HOF	O	-38.8	-68.9	-7.3	103.6
	H	23.6	19.6	24.4	27.0
	F	130.4	192.2	168.0	212.2
NH_3	N	280.5	270.7	279.7	279.3
	H	33.5	31.4	33.7	33.5
CH_2O	O	-475.5	-378.6	-392.7	-217.6
	C	13.6	1.5	18.8	47.9
	H	23.1	22.0	23.7	24.8
CH_4	C	205.6	198.9	204.9	207.2
	H	32.7	31.3	32.8	32.6
C2H4	C	82.6	69.7	78.6	92.8
	H	27.8	26.0	27.9	28.1
AlF	Al	554.5	572.9	559.3	573.8
	F	187.8	211.8	211.7	205.1
CH_3F	C	134.9	122.1	135.4	142.1
	F	506.2	482.9	507.0	499.3
	H	28.9	27.3	29.1	29.1
C_3H_4	C	207.9	192.1	207.3	210.1
	C	96.7	83.7	96.2	108.3
	H	27.0	24.4	27.2	27.4
	H	32.1	30.6	32.3	32.1
FCCH	C	179.1	179.9	177.3	182.8
	C	108.3	100.1	108.1	114.1
	H	31.8	30.5	31.9	31.8
	F	436.2	423.5	439.0	438.2
FCN	F	391.6	374.1	400.3	402.0
	C	84.6	82.2	84.1	93.4
	N	105.5	117.9	107.7	130.6
H_2S	S	789.4	739.0	797.4	806.0
	H	32.7	30.5	32.9	32.7
HCP	H	30.9	29.6	31.1	31.1
	C	39.6	37.6	37.2	58.4
	P	405.2	388.0	404.4	469.9
HFCO	O	-139.8	-94.3	-114.1	-60.0
	C	44.0	39.6	44.8	56.5
	F	170.3	165.3	193.3	202.4
	H	25.2	23.9	25.4	25.5
H_2C_2O	C	197.7	193.3	197.2	199.8
	C	-4.4	-6.3	0.9	16.0
	O	-49.1	-5.9	-13.0	14.2
	H	31.2	29.2	31.4	31.3
LiF	Li	89.0	89.3	89.3	89.1
	F	301.8	382.5	318.5	302.8
LiH	H	26.1	26.6	26.2	25.7
	Li	89.6	89.3	89.5	91.4
N_2O	N	99.0	106.4	104.3	124.2
	N	10.0	12.6	12.2	33.9
	O	183.1	199.0	198.1	207.1
OCS	O	73.4	96.8	83.4	105.0
	C	29.1	30.2	28.7	47.1
	S	807.2	796.7	816.8	831.3
OF_2	O	-392.4	-447.1	-329.6	-27.6
	F	-38.1	-24.0	14.6	74.1
H_4C_2O	O	376.6	363.2	382.4	377.3
	C	165.1	153.2	165.6	169.9
	H	30.8	29.1	31.0	30.8
PN	N	-370.9	-344.0	-268.3	-201.9
	P	45.1	50.6	146.9	254.6

Table 37: Statistics for absolute isotropic NMR shieldings in a pcSseg-0 basis. CCSD(T)/aug-cc-pCV[TQ]Z shieldings are used as a reference. MSD=mean signed deviation; MAD=mean absolute deviation; MRD=mean relative deviation [%]; MARD=mean absolute relative deviation [%]; RMSD=root mean squared deviation.

	MSD	MAD	MRD	MARD	RMSD
CCSD(T)	-1.4	13.8	4.0	36.3	23.4
RPA@HF	9.0	14.5	33.4	40.3	26.8
RPA@PBE	30.3	32.9	118.8	119.6	70.8

Table 38: Statistics for absolute isotropic NMR shieldings in a pcSseg-0 basis. CCSD(T) shieldings are used as a reference. MSD=mean signed deviation; MAD=mean absolute deviation; MRD=mean relative deviation [%]; MARD=mean absolute relative deviation [%]; RMSD=root mean squared deviation.

	MSD	MAD	MRD	MARD	RMSD
CCSD(T)	0.0	0.0	0.0	0.0	0.0
RPA@HF	10.5	10.9	14.9	15.4	24.7
RPA@PBE	31.8	32.0	56.7	56.8	71.1

Table 39: Testing influence of the auxiliary basis set in a pcSseg-0 basis set. Calculations with the cc-pVDZ/mp2fit, cc-pVTZ/mp2fit, and cc-pVQZ/mp2fit auxiliary basis set are compared. All values are given in ppm.

Mol.	Nuc.	cc-pVDZ/mp2fit	cc-pVTZ/mp2fit	cc-pVQZ/mp2fit
HF	H	30.7	30.7	30.7
	F	432.5	432.5	432.5
CO	C	11.5	11.6	11.6
	O	-63.7	-63.7	-63.6
N_2	N	-63.2	-63.1	-63.1
H_2O	O	356.3	356.3	356.3
	H	32.8	32.8	32.8
HCN	H	30.5	30.5	30.5
	C	87.1	87.1	87.1
	N	-9.0	-8.9	-8.9
HOF	O	-7.6	-7.4	-7.3
	H	24.4	24.4	24.4
	F	167.9	168.0	168.0
NH_3	N	279.7	279.7	279.7
	H	33.7	33.7	33.7
CH_2O	O	-393.1	-392.7	-392.7
	C	18.8	18.8	18.8
	H	23.7	23.7	23.7
CH_4	C	204.8	204.8	204.9
	H	32.8	32.8	32.8
C2H4	C	78.6	78.6	78.6
	H	27.9	27.9	27.9
AlF	Al	559.2	559.3	559.3
	F	211.6	211.7	211.7
CH_3F	C	135.4	135.4	135.4
	F	507.0	507.0	507.0
	H	29.1	29.1	29.1
C_3H_4	C	207.3	207.3	207.3
	C	96.2	96.2	96.2
	H	27.2	27.2	27.2
	H	32.3	32.3	32.3
FCCH	C	177.3	177.3	177.3
	C	108.1	108.1	108.1
	H	31.9	31.9	31.9
	F	439.0	439.0	439.0
FCN	F	400.3	400.3	400.3
	C	84.1	84.1	84.1
	N	107.6	107.7	107.7
H_2S	S	797.1	797.4	797.4
	H	32.9	32.9	32.9
HCP	H	31.1	31.1	31.1
	C	37.1	37.2	37.2
	P	404.2	404.4	404.4
HFCO	O	-114.2	-114.1	-114.1
	C	44.8	44.8	44.8
	F	193.2	193.2	193.3
	H	25.4	25.4	25.4
H_2C_2O	C	197.2	197.2	197.2
	C	0.8	0.9	0.9
	O	-13.3	-13.1	-13.0
	H	31.4	31.4	31.4
LiF	Li	89.3	89.3	89.3
	F	318.5	318.5	318.5
LiH	H	26.2	26.2	26.2
	Li	89.5	89.5	89.5
N_2O	N	104.3	104.3	104.3
	N	12.2	12.2	12.2
	O	198.1	198.1	198.1
OCS	O	83.3	83.4	83.4
	C	28.6	28.7	28.7
	S	816.7	816.8	816.8
OF_2	O	-330.1	-329.7	-329.6
	F	14.5	14.6	14.6
H_4C_2O	O	382.3	382.4	382.4
	C	165.6	165.6	165.6
	H	31.0	31.0	31.0
PN	N	-268.8	-268.3	-268.3
	P	146.2	146.8	146.9

Table 40: Absolute isotropic NMR shieldings in a pcSseg-1 basis. All values are given in ppm. “Extrap.” stands for basis-set extrapolated CCSD(T)-shieldings in an aug-cc-pCV[TQ]Z basis set. The data for “Extrap.” was taken from Ref. 3. CCSD(T) calculations were done with the CFOUR program[1, 2].

Mol.	Nuc.	CCSD(T)	Extrap.	RPA@HF	RPA@PBE
HF	H	29.2	28.8	29.1	29.8
	F	419.6	420.3	419.5	418.5
CO	C	6.3	2.2	2.2	27.0
	O	-54.4	-55.0	-50.8	-18.8
N_2	N	-58.6	-60.4	-61.7	-29.9
H_2O	O	346.3	338.0	344.4	343.8
	H	31.1	30.6	31.1	31.4
HCN	H	29.2	29.0	29.4	29.5
	C	88.0	84.6	83.9	96.0
	N	-11.6	-14.1	-15.1	12.8
HOF	O	-50.7	-68.9	-38.9	36.3
	H	20.7	19.6	21.4	23.2
	F	179.2	192.2	205.2	219.3
NH_3	N	276.8	270.7	274.0	274.4
	H	31.9	31.4	32.0	32.1
CH_2O	O	-399.5	-378.6	-364.8	-257.5
	C	13.0	1.5	11.9	34.0
	H	22.6	22.0	23.0	23.7
CH_4	C	197.9	198.9	195.7	197.0
	H	31.4	31.3	31.5	31.5
C2H4	C	74.6	69.7	69.8	81.3
	H	26.5	26.0	26.6	26.9
AlF	Al	568.6	572.9	574.1	581.9
	F	199.6	211.8	216.7	224.4
CH_3F	C	126.0	122.1	126.0	129.9
	F	473.8	482.9	474.3	468.8
	H	27.6	27.3	27.9	28.0
C_3H_4	C	193.8	192.1	192.3	194.2
	C	88.1	83.7	84.6	95.3
	H	24.8	24.4	25.0	25.4
	H	31.0	30.6	31.1	31.1
FCCH	C	179.9	179.9	177.4	181.3
	C	102.2	100.1	101.4	105.5
	H	30.6	30.5	30.7	30.8
	F	422.8	423.5	425.0	425.4
FCN	F	374.4	374.1	380.0	380.6
	C	84.2	82.2	81.3	88.8
	N	119.2	117.9	114.3	131.7
H_2S	S	750.6	739.0	744.9	755.0
	H	31.0	30.5	31.1	31.3
HCP	H	29.6	29.6	29.9	29.9
	C	44.1	37.6	35.9	54.6
	P	399.4	388.0	383.6	441.7
HFCO	O	-100.1	-94.3	-91.3	-56.7
	C	45.5	39.6	43.4	52.8
	F	165.6	165.3	183.1	188.4
	H	24.4	23.9	24.6	24.8
H_2C_2O	C	192.4	193.3	190.6	192.2
	C	-2.5	-6.3	-2.9	7.5
	O	-17.3	-5.9	-2.5	13.9
	H	29.5	29.2	29.7	29.8
LiF	Li	90.0	89.3	90.3	90.5
	F	345.5	382.5	355.5	351.9
LiH	H	26.5	26.6	26.7	26.5
	Li	90.8	89.3	90.9	93.3
N_2O	N	109.6	106.4	104.7	123.5
	N	20.5	12.6	13.5	34.6
	O	194.2	199.0	197.8	207.7
OCS	O	94.7	96.8	97.3	113.2
	C	35.0	30.2	29.5	45.0
	S	805.4	796.7	808.9	823.4
OF_2	O	-411.2	-447.1	-376.5	-189.9
	F	-6.4	-24.0	27.3	60.1
H_4C_2O	O	361.3	363.2	365.7	360.7
	C	156.7	153.2	156.4	159.1
	H	29.5	29.1	29.7	29.7
PN	N	-330.3	-344.0	-329.6	-237.8
	P	83.1	50.6	69.2	200.0

Table 41: Statistics for absolute isotropic NMR shieldings in a pcSseg-1 basis. CCSD(T)/aug-cc-pCV[TQ]Z shieldings are used as a reference. MSD=mean signed deviation; MAD=mean absolute deviation; MRD=mean relative deviation [%]; MARD=mean absolute relative deviation [%]; RMSD=root mean squared deviation.

	MSD	MAD	MRD	MARD	RMSD
CCSD(T)	1.9	5.6	17.0	24.0	9.8
RPA@HF	3.9	5.7	17.5	18.6	12.8
RPA@PBE	20.4	21.9	83.1	83.5	46.7

Table 42: Statistics for absolute isotropic NMR shieldings in a pcSseg-1 basis. CCSD(T) shieldings are used as a reference. MSD=mean signed deviation; MAD=mean absolute deviation; MRD=mean relative deviation [%]; MARD=mean absolute relative deviation [%]; RMSD=root mean squared deviation.

	MSD	MAD	MRD	MARD	RMSD
CCSD(T)	0.0	0.0	0.0	0.0	0.0
RPA@HF	2.0	5.2	7.0	14.4	9.6
RPA@PBE	18.5	18.9	47.6	47.7	41.4

Table 43: Absolute isotropic NMR shieldings in a pcSseg-2 basis. All values are given in ppm. "Extrap." stands for basis-set extrapolated CCSD(T)-shieldings in an aug-cc-pCV[TQ]Z basis set. The data for "Extrap." was taken from Ref. 3. CCSD(T) calculations were done with the CFOUR program[1, 2]. MP2 calculations were done with the Turbomole program[4–6].

Mol.	Nuc.	CCSD(T)	Extrap.	RPA@HF	RPA@PBE	RPA@KT2	RPA@B3LYP	RPA@BH&H	σ-P1@PBE	MP2
HF	H	29.2	28.8	29.1	29.7	29.7	29.6	29.4	29.5	28.9
	F	419.5	420.3	418.9	418.2	418.0	418.5	418.7	417.3	425.0
CO	C	6.5	2.2	0.6	24.7	20.4	17.4	9.8	15.3	12.1
	O	-52.3	-55.0	-50.7	-22.6	-27.0	-31.3	-39.5	-36.2	-45.4
N_2	N	-56.7	-60.4	-62.8	-32.2	-37.5	-41.6	-51.2	-46.1	-39.2
H_2O	O	339.4	338.0	337.1	337.4	337.7	337.7	337.6	335.7	347.6
	H	30.9	30.6	31.0	31.3	31.3	31.2	31.1	31.2	30.7
HCN	H	29.2	29.0	29.4	29.4	29.4	29.4	29.4	29.3	29.1
	C	87.4	84.6	83.3	95.1	93.1	91.8	88.2	92.4	88.8
	N	-11.3	-14.1	-15.3	10.6	4.9	3.9	-4.1	2.2	1.9
HOF	O	-58.9	-68.9	-53.9	15.0	-9.1	-15.8	-38.4	-25.7	-44.1
	H	20.1	19.6	20.8	22.3	21.7	21.5	21.0	21.2	20.2
	F	190.0	192.2	214.6	215.7	213.2	212.6	212.3	199.6	194.3
NH_3	N	272.4	270.7	269.1	270.2	270.4	270.2	269.8	269.0	278.0
	H	31.7	31.4	31.8	31.9	31.9	31.9	31.8	31.8	31.5
CH_2O	O	-377.6	-378.6	-348.7	-261.9	-287.9	-290.1	-313.2	-318.8	-331.5
	C	7.3	1.5	6.5	27.6	21.7	19.9	13.2	20.8	7.3
	H	22.3	22.0	22.8	23.2	23.0	22.9	22.8	22.9	22.3
CH_4	C	199.5	198.9	196.5	197.9	197.6	197.7	197.1	197.4	201.2
	H	31.4	31.3	31.5	31.5	31.5	31.5	31.5	31.5	31.3
C2H4	C	72.7	69.7	68.6	79.6	77.5	76.7	73.2	77.8	71.1
	H	26.2	26.0	26.4	26.6	26.6	26.6	26.5	26.6	26.1
AlF	Al	571.0	572.9	574.7	579.8	577.1	576.7	574.2	575.6	569.6
	F	213.9	211.8	229.1	239.0	236.4	233.0	229.3	242.6	210.3
CH_3F	C	124.2	122.1	124.3	127.9	126.9	126.8	125.1	127.2	122.7
	F	479.4	482.9	477.4	472.0	472.4	473.3	474.7	474.1	485.9
	H	27.5	27.3	27.9	27.9	27.9	27.8	27.8	27.9	27.4
C_3H_4	C	193.2	192.1	191.4	192.9	192.2	192.4	191.7	192.8	194.3
	C	85.9	83.7	82.5	92.7	90.7	89.9	86.5	90.7	87.2
	H	24.6	24.4	24.7	25.1	25.1	25.0	24.9	25.0	24.5
	H	30.8	30.6	31.0	31.0	31.0	31.0	31.0	31.0	30.8
FCCH	C	181.4	179.9	179.0	182.4	181.2	182.1	181.0	182.1	183.1
	C	101.8	100.1	101.5	106.0	105.0	104.9	103.2	106.0	99.1
	H	30.6	30.5	30.8	30.8	30.9	30.9	30.8	30.8	30.6
	F	424.5	423.5	425.8	425.5	425.0	425.3	425.5	425.9	430.7
FCN	F	375.2	374.1	379.2	378.9	379.1	378.0	378.2	374.7	384.1
	C	84.3	82.2	81.6	89.3	88.0	87.8	85.1	88.4	83.2
	N	120.1	117.9	114.6	130.5	127.3	128.0	123.4	126.7	133.4
H_2S	S	732.1	739.0	718.0	725.7	724.2	723.2	720.2	712.9	748.4
	H	30.7	30.5	30.9	31.1	31.1	31.1	31.0	31.1	30.5
HCP	H	29.7	29.6	30.0	29.9	29.9	29.9	29.9	30.0	29.5
	C	41.5	37.6	34.1	52.4	49.5	47.8	42.4	51.5	46.2
	P	381.7	388.0	363.1	411.7	403.2	399.0	384.3	402.8	385.9
HFCO	O	-93.0	-94.3	-86.9	-57.6	-61.8	-62.3	-69.5	-76.0	-61.3
	C	43.2	39.6	41.1	50.6	48.9	48.4	45.1	47.9	43.4
	F	166.8	165.3	183.3	186.5	182.2	184.1	181.9	179.5	166.5
	H	24.1	23.9	24.4	24.4	24.4	24.4	24.4	24.4	24.1
H_2C_2O	C	194.4	193.3	192.0	193.8	193.5	193.4	192.9	192.5	199.3
	C	-3.3	-6.3	-4.8	5.9	4.8	3.5	0.1	-0.5	0.7
	O	-6.7	-5.9	2.2	16.8	16.1	14.9	11.8	4.0	28.4
	H	29.3	29.2	29.5	29.6	29.7	29.6	29.6	29.5	29.4
LiF	Li	89.6	89.3	90.0	90.2	89.6	90.0	89.8	90.5	89.5
	F	372.0	382.5	380.9	382.9	374.3	382.1	380.3	390.5	369.7
LiH	H	26.6	26.6	26.9	26.6	26.6	26.7	26.8	26.8	26.7
	Li	89.4	89.3	89.5	91.1	90.6	90.5	90.0	91.2	89.4
N_2O	N	109.6	106.4	103.0	121.7	119.7	117.5	111.9	115.2	132.7
	N	15.3	12.6	7.2	28.4	26.7	23.2	16.8	24.3	33.5
	O	200.8	199.0	200.9	210.8	210.6	208.8	206.2	202.8	218.9
OCS	O	99.9	96.8	101.1	115.3	114.1	111.9	107.7	111.2	113.0
	C	32.9	30.2	27.0	42.6	41.2	39.1	34.3	40.4	39.9
	S	797.4	796.7	796.6	806.1	804.3	803.2	800.3	794.5	816.8
OF_2	O	-429.7	-447.1	-401.2	-245.5	-317.2	-328.9	-378.8	-349.4	-452.3
	F	-16.9	-24.0	12.9	36.9	30.3	26.0	17.9	7.5	-7.0
H_4C_2O	O	363.6	363.2	365.5	360.5	359.5	362.7	364.1	360.3	372.5
	C	154.9	153.2	154.4	157.0	156.0	156.2	155.0	156.2	154.2
	H	29.3	29.1	29.6	29.5	29.5	29.5	29.5	29.5	29.2
PN	N	-339.8	-344.0	-350.0	-262.9	-281.1	-290.5	-317.0	-289.4	-248.0
	P	42.7	50.6	10.8	129.4	105.3	87.9	51.8	82.7	104.9

Table 44: Absolute isotropic NMR shieldings in a pcSseg-2 basis. All values are given in ppm. “Extrap.” stands for basis-set extrapolated CCSD(T)-shieldings in an aug-cc-pCV[TQ]Z basis set. The data for “Extrap.” was taken from Ref. 3. CCSD(T) calculations were done with the CFOUR program[1, 2].

Mol.	Nuc.	CCSD(T)	Extrap.	HF(num.)	PBE(num.)	KT2(num.)	B3LYP(num.)
HF	H	29.2	28.8	28.3	29.9	29.9	29.4
	F	419.5	420.3	414.5	411.0	408.6	411.3
CO	C	6.5	2.2	-27.0	-17.2	-0.9	-22.8
	O	-52.3	-55.0	-91.8	-85.7	-68.1	-88.6
N_2	N	-56.7	-60.4	-114.9	-88.8	-69.1	-97.2
H_2O	O	339.4	338.0	328.8	330.4	326.9	328.4
	H	30.9	30.6	30.6	31.3	31.4	31.1
HCN	H	29.2	29.0	29.3	29.2	28.9	29.3
	C	87.4	84.6	70.3	69.6	80.2	67.1
	N	-11.3	-14.1	-53.2	-49.2	-28.0	-54.2
HOF	O	-58.9	-68.9	-135.9	-129.4	-97.8	-138.2
	H	20.1	19.6	19.2	19.5	20.2	19.5
	F	190.0	192.2	287.0	147.2	156.8	172.3
NH_3	N	272.4	270.7	263.5	264.1	263.1	261.6
	H	31.7	31.4	31.7	31.9	31.9	31.8
CH_2O	O	-377.6	-378.6	-449.6	-461.8	-398.7	-466.7
	C	7.3	1.5	-7.8	-28.6	-11.1	-26.9
	H	22.3	22.0	22.6	20.9	21.5	21.5
CH_4	C	199.5	198.9	195.1	190.1	192.0	188.4
	H	31.4	31.3	31.6	31.5	31.5	31.5
C2H4	C	72.7	69.7	58.9	46.6	58.2	45.8
	H	26.2	26.0	26.2	25.6	25.9	25.9
AlF	Al	571.0	572.9	577.2	540.4	556.7	547.2
	F	213.9	211.8	226.2	140.6	140.7	164.3
CH_3F	C	124.2	122.1	125.1	104.5	109.7	105.8
	F	479.4	482.9	485.7	459.8	458.0	464.7
	H	27.5	27.3	28.0	27.1	27.2	27.4
C_3H_4	C	193.2	192.1	193.6	176.7	180.0	177.9
	C	85.9	83.7	70.3	61.3	71.8	59.6
	H	24.6	24.4	24.1	24.2	24.3	24.3
	H	30.8	30.6	30.9	30.6	30.6	30.7
FCCH	C	181.4	179.9	176.1	169.5	177.4	168.7
	C	101.8	100.1	100.5	81.5	89.5	82.2
	H	30.6	30.5	30.5	31.0	30.8	30.9
	F	424.5	423.5	428.3	388.4	391.1	399.9
FCN	F	375.2	374.1	377.3	338.2	334.4	350.6
	C	84.3	82.2	74.9	67.8	77.4	65.1
	N	120.1	117.9	90.1	97.4	113.8	90.4
H_2S	S	732.1	739.0	712.4	710.4	725.8	694.4
	H	30.7	30.5	30.6	30.8	30.8	30.8
HCP	H	29.7	29.6	30.1	29.6	29.2	29.8
	C	41.5	37.6	11.6	11.5	23.1	7.1
	P	381.7	388.0	330.5	320.1	375.4	310.0
HFCO	O	-93.0	-94.3	-134.0	-132.2	-107.2	-141.5
	C	43.2	39.6	33.4	20.3	30.6	19.0
	F	166.8	165.3	186.3	88.2	101.2	112.9
	H	24.1	23.9	24.5	23.4	23.5	23.8
H_2C_2O	C	194.4	193.3	189.4	185.0	188.7	183.3
	C	-3.3	-6.3	-15.1	-19.7	-8.6	-25.3
	O	-6.7	-5.9	-30.9	-25.5	-11.8	-33.5
	H	29.3	29.2	29.3	29.3	29.3	29.4
LiF	Li	89.6	89.3	90.8	86.7	90.6	87.7
	F	372.0	382.5	383.2	323.1	345.8	335.6
LiH	H	26.6	26.6	26.5	26.2	26.4	26.5
	Li	89.4	89.3	89.5	88.1	92.0	88.2
N_2O	N	109.6	106.4	61.7	89.6	96.2	79.4
	N	15.3	12.6	-34.8	-0.3	6.5	-13.6
	O	200.8	199.0	173.2	173.7	171.5	170.8
OCS	O	99.9	96.8	73.9	68.3	75.4	68.2
	C	32.9	30.2	6.8	21.8	30.6	13.1
	S	797.4	796.7	784.1	762.1	778.4	759.2
OF_2	O	-429.7	-447.1	-444.5	-631.8	-548.1	-590.8
	F	-16.9	-24.0	22.4	-89.8	-71.5	-72.5
H_4C_2O	O	363.6	363.2	378.7	331.9	336.0	339.0
	C	154.9	153.2	155.7	137.1	142.3	138.0
	H	29.3	29.1	29.7	29.1	29.2	29.3
PN	N	-339.8	-344.0	-521.1	-426.6	-390.3	-454.9
	P	42.7	50.6	-127.4	-54.9	27.0	-81.9

Table 45: Statistics for absolute isotropic NMR shieldings in a pcSseg-2 basis. CCSD(T)/aug-cc-pCV[TQ]Z shieldings are used as a reference. MSD=mean signed deviation; MAD=mean absolute deviation; MRD=mean relative deviation [%]; MARD=mean absolute relative deviation [%]; RMSD=root mean squared deviation.

	MSD	MAD	MRD	MARD	RMSD
HF(num.)	-13.4	20.0	-58.1	66.7	39.5
PBE(num.)	-24.9	25.0	-75.9	76.3	39.9
KT2(num.)	-13.8	14.1	-29.0	29.6	23.3
B3LYP(num.)	-25.1	25.2	-83.9	84.3	38.9
CCSD	0.4	3.9	3.1	11.0	8.4
CCSD(T)	1.2	2.4	11.3	12.4	3.8
RPA@HF	1.6	5.6	7.8	15.3	11.4
RPA@BH&H	6.5	7.7	31.0	31.3	15.2
RPA@B3LYP	11.1	11.9	50.1	50.3	23.5
RPA@KT2	12.3	13.5	56.5	56.8	26.1
RPA@PBE	16.2	17.1	69.5	69.7	36.0
σ-P1@PBE	8.9	10.3	43.8	44.1	19.3
MP2	8.6	9.3	34.2	34.4	18.0

Table 46: Statistics for absolute isotropic NMR shieldings in a pcSseg-2 basis. CCSD(T) shieldings are used as a reference. MSD=mean signed deviation; MAD=mean absolute deviation; MRD=mean relative deviation [%]; MARD=mean absolute relative deviation [%]; RMSD=root mean squared deviation.

	MSD	MAD	MRD	MARD	RMSD
HF(num.)	-14.6	21.0	-44.7	54.2	40.0
PBE(num.)	-26.1	26.1	-54.9	55.0	41.6
KT2(num.)	-15.0	15.2	-25.5	25.8	25.2
B3LYP(num.)	-26.2	26.3	-62.2	62.3	40.4
CCSD	-0.8	3.4	-4.8	8.1	6.4
CCSD(T)	0.0	0.0	0.0	0.0	0.0
RPA@HF	0.4	5.2	-0.2	11.8	9.4
RPA@BH&H	5.4	6.2	15.0	15.2	13.0
RPA@B3LYP	9.9	10.6	26.6	26.8	21.3
RPA@KT2	11.1	11.9	30.8	31.0	23.9
RPA@PBE	15.0	15.7	38.0	38.1	33.9
σ-P1@PBE	7.7	9.0	19.4	19.7	17.1
MP2	7.4	8.6	20.7	21.4	17.5

References

[1] D. A. Matthews, L. Cheng, M. E. Harding, F. Lipparini, S. Stopkowicz, T.-C. Jagau, P. G. Szalay, J. Gauss, and J. F. Stanton, J. Chem. Phys. **152**, 214108 (2020).

[2] J. F. Stanton, J. Gauss, L. Cheng, M. E. Harding, D. A. Matthews, and P. G. Szalay, "CFOUR, Coupled-Cluster techniques for Computational Chemistry, a quantum-chemical program package," With contributions from A.A. Auer, R.J. Bartlett, U. Benedikt, C. Berger, D.E. Bernholdt, S. Blaschke, Y. J. Bomble, S. Burger, O. Christiansen, D. Datta, F. Engel, R. Faber, J. Greiner, M. Heckert, O. Heun, M. Hilgenberg, C. Huber, T.-C. Jagau, D. Jonsson, J. Jusélius, T. Kirsch, K. Klein, G.M. KopperW.J. Lauderdale, F. Lipparini, T. Metzroth, L.A. Mück, D.P. O'Neill, T. Nottoli, 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.

[3] A. M. Teale, O. B. Lutnæs, T. Helgaker, D. J. Tozer, and J. Gauss, J. Chem. Phys. **138**, 024111 (2013).

[4] "TURBOMOLE V7.4 2019, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from `http://www.turbomole.com`." .

[5] M. Kollwitz and J. Gauss, Chem. Phys. Lett. **260**, 639 (1996).

[6] M. Kollwitz, M. Häser, and J. Gauss, J. Chem. Phys. **108**, 8295 (1998).

3.5 Gauge-origin dependence in electronic g-tensor calculations

M. Glasbrenner, S. Vogler, C. Ochsenfeld.
J. Chem. Phys. **148**, 214101 (2018).

Abstract: 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.

Reproduced from:

M. Glasbrenner, S. Vogler, C. Ochsenfeld.
"Gauge-origin dependence in electronic g-tensor calculations."
J. Chem. Phys. **148**, 214101 (2018).

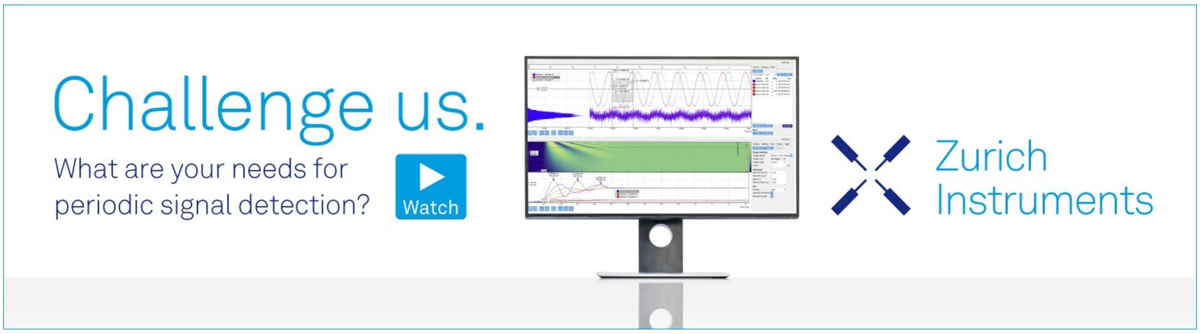
with the permission of AIP Publishing.
`https://aip.scitation.org/doi/abs/10.1063/1.5028454`

The Journal of Chemical Physics

Gauge-origin dependence in electronic g-tensor calculations

Cite as: J. Chem. Phys. **148**, 214101 (2018); https://doi.org/10.1063/1.5028454
Submitted: 10 March 2018 . Accepted: 09 May 2018 . Published Online: 04 June 2018

Michael Glasbrenner, Sigurd Vogler, and Christian Ochsenfeld



J. Chem. Phys. **148**, 214101 (2018); https://doi.org/10.1063/1.5028454

THE JOURNAL OF CHEMICAL PHYSICS **148**, 214101 (2018)

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

(Received 10 March 2018; accepted 9 May 2018; published online 4 June 2018)

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 $\vec{s}$

$$g_{pq} = \frac{2}{\alpha} \frac{\partial^2 E}{\partial B_p \partial s_q}\bigg|_{B=0,s=0}, \qquad p,q \in \{x,y,z\}, \tag{1}$$

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]

$$g_{pq} = \delta_{pq} g_{\text{el}} + g_{pq}^{\text{rmc}} + g_{pq}^{\text{dso1}} + g_{pq}^{\text{dso2}} + g_{pq}^{\text{pso1}} + g_{pq}^{\text{pso2}}. \tag{2}$$

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 $P_{\mu\nu}^{\alpha-\beta}$

$$g_{pq}^{\text{rmc}} = -\frac{\alpha^2 g_{\text{el}}}{2S} \delta_{pq} \sum_{\mu\nu} P_{\mu\nu}^{\alpha-\beta} T_{\mu\nu}. \tag{3}$$

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:

$$\begin{aligned} g_{pq}^{\text{dso1}} = \frac{\alpha^2 g'}{8S} \sum_A Z_A \sum_{\mu\nu} P_{\mu\nu}^{\alpha-\beta} & \left[\int \chi_\mu(\vec{r}) \frac{\delta_{pq}(\vec{r} - \vec{R}_A) \cdot (\vec{r} - \vec{R}_0)}{|\vec{r} - \vec{R}_A|^3} \chi_\nu(\vec{r}) dr^3 \right. \\ & \left. - \int \chi_\mu(\vec{r}) \frac{(\vec{r} - \vec{R}_A)_p (\vec{r} - \vec{R}_0)_q}{|\vec{r} - \vec{R}_A|^3} \chi_\nu(\vec{r}) dr^3 \right], \end{aligned} \tag{4}$$

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

$$\begin{aligned} g_{pq}^{\text{dso1}} + g_{pq}^{\text{dso2}} \approx \frac{\alpha^2 g'}{8S} \sum_A \tilde{Z}_A \sum_{\mu\nu} P_{\mu\nu}^{\alpha-\beta} & \left[\int \chi_\mu(\vec{r}) \frac{\delta_{pq}(\vec{r} - \vec{R}_A) \cdot (\vec{r} - \vec{R}_0)}{|\vec{r} - \vec{R}_A|^3} \chi_\nu(\vec{r}) dr^3 \right. \\ & \left. - \int \chi_\mu(\vec{r}) \frac{(\vec{r} - \vec{R}_A)_p (\vec{r} - \vec{R}_0)_q}{|\vec{r} - \vec{R}_A|^3} \chi_\nu(\vec{r}) dr^3 \right]. \end{aligned} \tag{5}$$

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

$$z_{\mu\nu}^q = h_{\mu\nu,q}^{\text{SOMF}} + J_{\mu\nu,q}^{\text{SOMF}} + K_{\mu\nu,q}^{\text{SOMF}}, \qquad q \in \{x,y,z\}, \tag{6}$$

$$h_{\mu\nu,q}^{\text{SOMF}} = \frac{\alpha^2 g'}{4} \sum_A Z_A \int \chi_\mu(\vec{r}) \frac{\left[(\vec{r} - \vec{R}_A) \times \hat{p}\right]_q}{|\vec{r} - \vec{R}_A|^3} \chi_\nu(\vec{r}) dr^3, \tag{7}$$

$$J_{\mu\nu,q}^{\text{SOMF}} = -\frac{\alpha^2 g'}{4} \sum_{\lambda\sigma} P_{\lambda\sigma} g_{\mu\nu\lambda\sigma}^{\text{soc, q}}, \tag{8}$$

$$K_{\mu\nu,q}^{\text{SOMF}} = \frac{3}{8} \alpha^2 g' \sum_{\lambda\sigma} \left(P_{\lambda\sigma} g_{\mu\lambda\sigma\nu}^{\text{soc, q}} + P_{\lambda\sigma} g_{\sigma\nu\mu\lambda}^{\text{soc, q}} \right), \tag{9}$$

$$g_{\mu\nu\lambda\sigma}^{\text{soc, q}} = \int \chi_\mu(\vec{r}) \frac{[(\vec{r} - \vec{r}\,') \times \hat{p}]_q}{|\vec{r} - \vec{r}\,'|^3} \chi_\nu(\vec{r}) \chi_\lambda(\vec{r}\,') \chi_\sigma(\vec{r}\,') dr^3 dr'^3. \tag{10}$$

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:

$$g_{pq}^{\text{pso1}} + g_{pq}^{\text{pso2}} \approx \frac{1}{\alpha S} \sum_{\mu\nu} \frac{\partial P_{\mu\nu}^{\alpha-\beta}}{\partial B_p}\bigg|_{s=0} z_{\mu\nu}^q. \tag{11}$$

The magnetic field derivative of the spin density is obtained from the difference of the perturbed α and β densities

$$\frac{\partial P_{\mu\nu}^{\alpha-\beta}}{\partial B_p}\bigg|_{s=0} = \frac{\partial P_{\mu\nu}^{\alpha}}{\partial B_p}\bigg|_{s=0} - \frac{\partial P_{\mu\nu}^{\beta}}{\partial B_p}\bigg|_{s=0}, \tag{12}$$

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]

$$\phi_\mu(\vec{r}) = \exp\left[-i\frac{\alpha}{2}(\vec{B} \times (\vec{R}_\mu - \vec{R}_0)) \cdot \vec{r}\right] \chi_\mu(\vec{r}). \tag{13}$$

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:

$$g_{pq}^{\text{dso1}} + g_{pq}^{\text{dso2}} \approx \frac{\alpha^2 g'}{8S} \sum_A \tilde{Z}_A \sum_{\mu\nu} P_{\mu\nu}^{\alpha-\beta} \left[\int \chi_\mu(\vec{r}) \frac{\delta_{pq}(\vec{r}-\vec{R}_A)\cdot(\vec{r}-\vec{R}_\nu)}{|\vec{r}-\vec{R}_A|^3} \chi_\nu(\vec{r}) dr^3 - \int \chi_\mu(\vec{r}) \frac{(\vec{r}-\vec{R}_A)_p(\vec{r}-\vec{R}_\nu)_q}{|\vec{r}-\vec{R}_A|^3} \chi_\nu(\vec{r}) dr^3 \right]. \tag{14}$$

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

$$g_{pq}^{\text{pso1}} + g_{pq}^{\text{pso2}} \approx \frac{1}{\alpha S} \sum_{\mu\nu} \left[\left.\frac{\partial P_{\mu\nu}^{\alpha-\beta}}{\partial B_p}\right|_{s=0} z_{\mu\nu}^q + P_{\mu\nu}^{\alpha-\beta} \left(\frac{\partial z_{\mu\nu}^q}{\partial B_p} \right) \right], \tag{15}$$

$$\frac{\partial z_{\mu\nu}^q}{\partial B_p} = -\frac{\alpha}{2} \left(\vec{R}_\mu \times \vec{R}_\nu \right)_p z_{\mu\nu}^q - \frac{\alpha}{2} \left[\left(\vec{R}_\mu - \vec{R}_\nu \right) \times z_{\mu\nu^+}^q \right]_p. \tag{16}$$

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

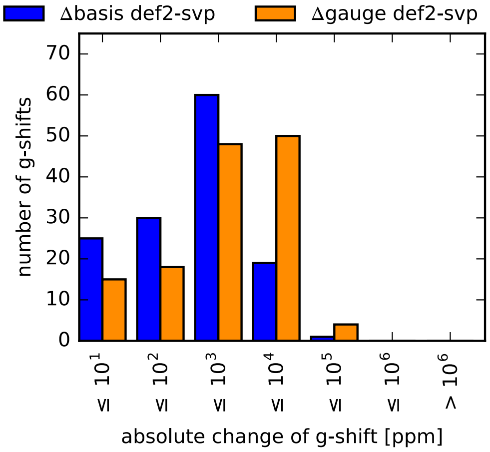


FIG. 1. Comparison of the influence of the basis set and gauge-origin on g-shifts in our test set. Δbasis is the absolute deviation of the value computed with the def2-SVP basis set from the def2-QZVPPD result obtained with the same nuclear coordinates. Δgauge is the absolute change in the individual 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).

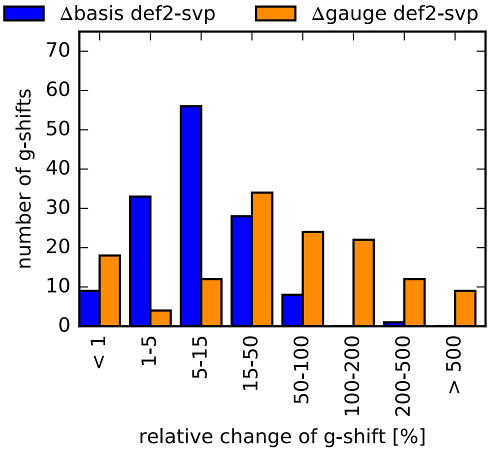


FIG. 3. Comparison of the influence of the basis set and gauge-origin on g-shifts in our test set. Δbasis is the unsigned relative deviation of the value computed with the def2-SVP basis set from the def2-QZVPPD result obtained with the same nuclear coordinates. Δgauge is the unsigned relative change 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

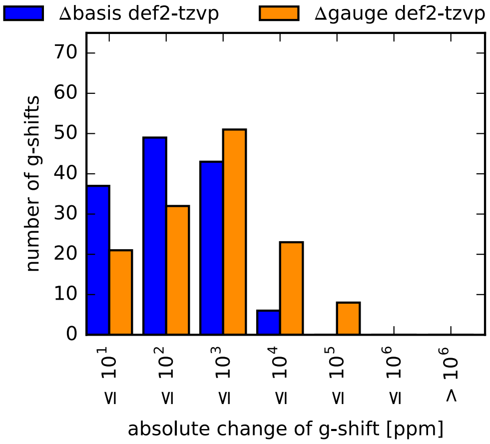


FIG. 2. Comparison of the influence of the basis set and gauge-origin on g-shifts in our test set. Δbasis is the absolute deviation of the value computed with the def2-TZVP basis set from the def2-QZVPPD result obtained with the same nuclear coordinates. Δgauge is the absolute change in the individual 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).

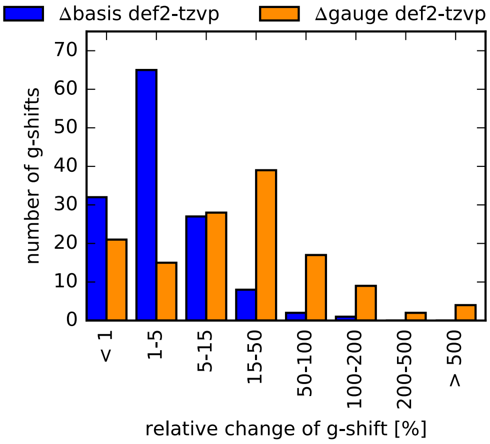


FIG. 4. Comparison of the influence of the basis set and gauge-origin on g-shifts in our test set. Δbasis is the unsigned relative deviation of the value computed with the def2-TZVP basis set from the def2-QZVPPD result obtained with the same nuclear coordinates. Δgauge is the unsigned relative change 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)"].

		def2-SVP		def2-TZVP	
Shifts		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

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 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

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]

$$r_{\mathrm{ECC},p} = \frac{1}{n_{\mathrm{el}}} \sum_{\mu\nu} P_{\mu\nu} \left\langle \mu \middle| \hat{r}_p \middle| \nu \right\rangle, \qquad p \in \{x, y, z\}, \tag{17}$$

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

$$r_{\mathrm{SDC},p} = \frac{1}{\mathrm{Tr}(|P^{\alpha-\beta}|S)} \sum_{\mu\nu} |P^{\alpha-\beta}_{\mu\nu}| \left\langle \mu \middle| \hat{r}_p \middle| \nu \right\rangle, \quad p \in \{x, y, z\}. \tag{18}$$

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

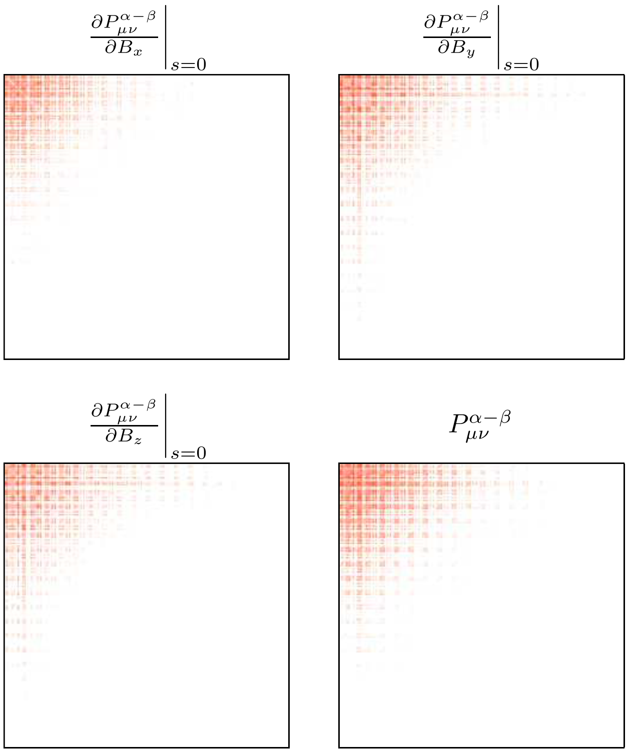


FIG. 5. Absolute values of the entries in the spin density matrix and the perturbed spin density matrices for the calculation of the $\cdot$O—[CH$_2$]$_{18}$—CH$_3$ molecule (def2-SVP basis set). Only matrix elements above a threshold of 10^{-5} are shown. The basis functions centered on the oxygen atom are located in the upper left corner.

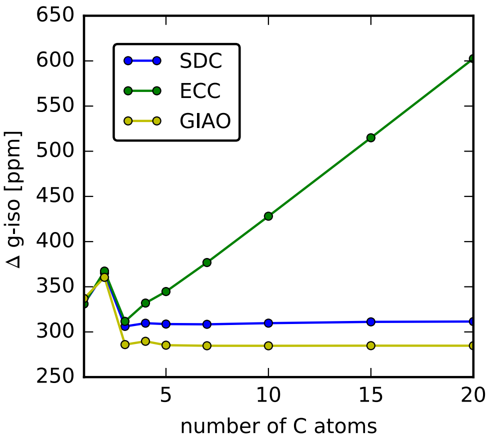


FIG. 6. Isotropic g-shifts of alkane radicals of increasing chain length with a terminal CH$_2$ group. The computations were carried out at the UDFT/B3LYP-level with a def2-TZVP basis. The GIAO ansatz is compared to the common gauge-origins SDC and ECC.

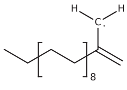
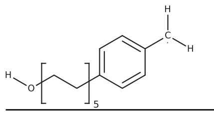
from a minor gauge error within the SDC calculation. In contrast to the SDC, the ECC is by construction always located centrally within the molecule (the ECC can even be viewed as one possible definition of the center of the molecule); for these calculations, it means that the ECC is for growing chain 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 $\cdot NH-[CH_2-O]_3-CH_3-OH$ or the $\cdot O-[CH_2]_{18}-CH_3$ 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 $\cdot NH-[CH_2-O]_7-CH_3-OH$ 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
$\cdot NH-[CH_2-O]_7-CH_2-OH$	ECC	5 680	1 505	207	2 464
	SDC	4 137	1 509	−132	1 838
	GIAO	4 119	1 540	−151	1 836
$\cdot O-[CH_2]_{18}-CH_3$	ECC	65 816	7 382	292	24 497
	SDC	65 935	6 494	−143	24 095
	GIAO	66 021	6 443	−175	24 096
$\cdot Cys-Gly_4$	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$_2$]$_x$—$\dot{\mathrm{C}}$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 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.

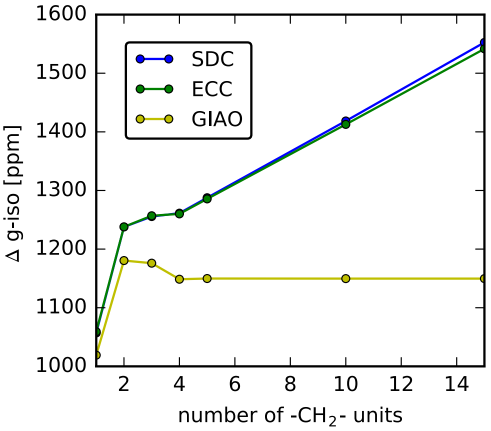


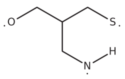
FIG. 7. Isotropic g-shifts of radicals of the type $\cdot$NH—[CH$_2$]$_x$—$\dot{\mathrm{C}}$H$_2$ with triplet spin state, where x is varied from 1 to 15. The computations were carried out at the UDFT/B3LYP-level and with a def2-TZVP basis. The GIAO ansatz is compared to the common gauge-origins SDC and ECC.

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)
$\mathrm{O{=}\dot{C}{-}[CH_2]_{17}{-}C{\equiv}C\cdot}$	ECC	1 550	−92	−3268	−603
	SDC	1 487	−96	−3248	−619
	GIAO	1 353	−118	−3253	−673
$\mathrm{\cdot NF{-}CH_2{-}[CH{=}CH{-}CH_2]_5{-}CH_2{-}CH{=}\dot{C}H}$	ECC	4 167	1 788	−35	1 973
	SDC	4 197	1 786	−28	1 985
	GIAO	3 704	1 518	−155	1 689
$\mathrm{\cdot Ala{-}Gly_2{-}Lys\cdot}$	ECC	3 252	1 031	113	1 465
	SDC	3 149	1 032	111	1 431
	GIAO	2 632	887	−32	1 162
$\mathrm{\cdot NH{-}[CH_2{-}O]_7{-}CH_2{-}S\cdot}$	ECC	108 751	9 710	−100	39 454
	SDC	108 742	9 710	−100	39 451
	GIAO	108 747	9 710	−100	39 452
$\mathrm{\cdot O{-}[CH_2]_{18}{-}\dot{N}H}$	ECC	35 012	4 516	1108	13 545
	SDC	35 002	4 513	1110	13 542
	GIAO	34 749	3 795	561	13 035
	ECC	138 717	20 958	2980	54 218
	SDC	138 718	20 968	2975	54 220
	GIAO	138 658	20 575	2805	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.

ACKNOWLEDGMENTS

The authors acknowledge financial support by the “Deutsche Forschungsgemeinschaft” (DFG) Funding Proposal No. Oc35/4-1 and the Cluster of Excellence EXC 114 “Center for Integrated Protein Science Munich” (CIPSM). S.V. thanks the Studienstiftung des Deutschen Volkes for a graduate fellowship. C.O. acknowledges additional financial support as a Max-Planck-Fellow at MPI-FKF Stuttgart.

[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. Komorovskỳ, 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

Continued on next page

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

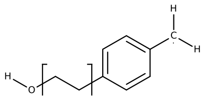
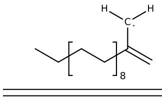
Continued on next page

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

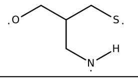
1.2 Section "Gauge-origin dependence in molecules with a single localized spin center"

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=$\dot{\mathrm{C}}$-[CH$_2$]$_{17}$-C≡C•	def2-tzvp	ECC	1550	-92	-3268
•NF-CH$_2$-[CH=CH-CH$_2$]$_5$-CH$_2$-CH=$\dot{\mathrm{C}}$H	def2-tzvp	ECC	4167	1788	-35
•Ala-Gly$_2$-Lys•	def2-tzvp	ECC	3252	1031	113
•NH-$\dot{\mathrm{C}}$H$_2$	def2-tzvp	ECC	2296	887	-6
•NH-CH$_2$-$\dot{\mathrm{C}}$H$_2$	def2-tzvp	ECC	2763	979	-28
•NH-[CH$_2$]$_2$-$\dot{\mathrm{C}}$H$_2$	def2-tzvp	ECC	2647	944	179
•NH-[CH$_2$]$_3$-$\dot{\mathrm{C}}$H$_2$	def2-tzvp	ECC	2812	901	67
•NH-[CH$_2$]$_4$-$\dot{\mathrm{C}}$H$_2$	def2-tzvp	ECC	2688	913	256
•NH-[CH$_2$]$_9$-$\dot{\mathrm{C}}$H$_2$	def2-tzvp	ECC	3155	954	129
•NH-[CH$_2$]$_{14}$-$\dot{\mathrm{C}}$H$_2$	def2-tzvp	ECC	3251	917	457
•O-[CH$_2$]$_{18}$-$\dot{\mathrm{N}}$H	def2-tzvp	ECC	35012	4516	1108
O=$\dot{\mathrm{C}}$-[CH$_2$]$_{17}$-C≡C•	def2-tzvp	SDC	1487	-96	-3248
•NF-CH$_2$-[CH=CH-CH$_2$]$_5$-CH$_2$-CH=$\dot{\mathrm{C}}$H	def2-tzvp	SDC	4197	1786	-28
•Ala-Gly$_2$-Lys•	def2-tzvp	SDC	3149	1032	111
•NH-$\dot{\mathrm{C}}$H$_2$	def2-tzvp	SDC	2288	890	-7
•NH-CH$_2$-$\dot{\mathrm{C}}$H$_2$	def2-tzvp	SDC	2764	978	-28
•NH-[CH$_2$]$_2$-$\dot{\mathrm{C}}$H$_2$	def2-tzvp	SDC	2646	943	178
•NH-[CH$_2$]$_3$-$\dot{\mathrm{C}}$H$_2$	def2-tzvp	SDC	2818	899	67
•NH-[CH$_2$]$_4$-$\dot{\mathrm{C}}$H$_2$	def2-tzvp	SDC	2692	915	255
•NH-[CH$_2$]$_9$-$\dot{\mathrm{C}}$H$_2$	def2-tzvp	SDC	3178	949	129
•NH-[CH$_2$]$_{14}$-$\dot{\mathrm{C}}$H$_2$	def2-tzvp	SDC	3289	913	456
•O-[CH$_2$]$_{18}$-$\dot{\mathrm{N}}$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=$\dot{\mathrm{C}}$-[CH$_2$]$_{17}$-C≡C•	def2-tzvp	GIAO	1353	-118	-3253
•NF-CH$_2$-[CH=CH-CH$_2$]$_5$-CH$_2$-CH=$\dot{\mathrm{C}}$H	def2-tzvp	GIAO	3704	1518	-155
•Ala-Gly$_2$-Lys•	def2-tzvp	GIAO	2632	887	-32
•NH-$\dot{\mathrm{C}}$H$_2$	def2-tzvp	GIAO	2196	901	-39
•NH-CH$_2$-$\dot{\mathrm{C}}$H$_2$	def2-tzvp	GIAO	2623	996	-78
•NH-[CH$_2$]$_2$-$\dot{\mathrm{C}}$H$_2$	def2-tzvp	GIAO	2468	948	112
•NH-[CH$_2$]$_3$-$\dot{\mathrm{C}}$H$_2$	def2-tzvp	GIAO	2564	897	-15
•NH-[CH$_2$]$_4$-$\dot{\mathrm{C}}$H$_2$	def2-tzvp	GIAO	2376	931	143
•NH-[CH$_2$]$_9$-$\dot{\mathrm{C}}$H$_2$	def2-tzvp	GIAO	2579	927	-57
•NH-[CH$_2$]$_{14}$-$\dot{\mathrm{C}}$H$_2$	def2-tzvp	GIAO	2378	928	143
•NH-[CH$_2$-O]$_7$-CH$_2$-S•	def2-tzvp	GIAO	108747	9710	-100
•O-[CH$_2$]$_{18}$-$\dot{\mathrm{N}}$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

3.6 Linear and sublinear scaling computation of the electronic g-tensor at the density functional theory level

M. Glasbrenner, S. Vogler, C. Ochsenfeld.
J. Chem. Phys. **150**, 024104 (2019).

Abstract: We present an efficient and low-scaling implementation of a density functional theory based method for the computation of electronic g-tensors. It allows for an accurate description of spin-orbit coupling effects by employing the spin-orbit mean-field operator. Gauge-origin independence is ensured by the use of gauge-including atomic orbitals. Asymptotically linear scaling with molecule size is achieved with an atomic orbital based formulation, integral screening methods, and sparse linear algebra. In addition, we introduce an ansatz that exploits the locality of the contributions to the g-tensor for molecules with local spin density. For such systems, sublinear scaling is obtained by restricting the magnetic field perturbation to the relevant subspaces of the full atomic orbital space; several criteria for selecting these subspaces are discussed and compared. It is shown that the computational cost of g-tensor calculations with the local approach can fall below the cost of the self-consistent field calculation for large molecules. The presented methods thus enable efficient, accurate, and gauge-origin independent computations of electronic g-tensors of large molecular systems.

Reproduced from:

M. Glasbrenner, S. Vogler, C. Ochsenfeld.
“Linear and sublinear scaling computation of the electronic g-tensor at the density functional theory level.”
J. Chem. Phys. **150**, 024104 (2019).

with the permission of AIP Publishing.
`https://aip.scitation.org/doi/abs/10.1063/1.5066266`

The Journal of Chemical Physics

Linear and sublinear scaling computation of the electronic g-tensor at the density functional theory level EP

Cite as: J. Chem. Phys. **150**, 024104 (2019); https://doi.org/10.1063/1.5066266
Submitted: 13 October 2018 . Accepted: 03 December 2018 . Published Online: 09 January 2019

Michael Glasbrenner, Sigurd Vogler, and Christian Ochsenfeld

COLLECTIONS

EP This paper was selected as an Editor's Pick

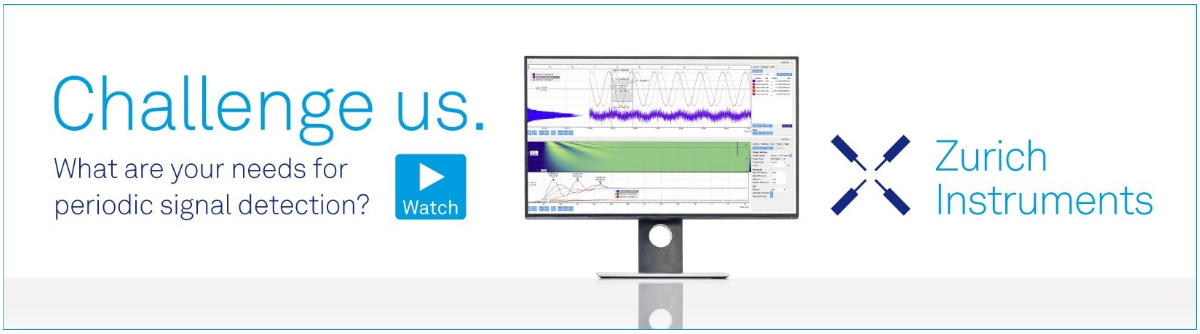
View Online Export Citation CrossMark

ARTICLES YOU MAY BE INTERESTED IN

Understanding non-covalent interactions in larger molecular complexes from first principles
The Journal of Chemical Physics **150**, 010901 (2019); https://doi.org/10.1063/1.5075487

Perspective: Computational chemistry software and its advancement as illustrated through three grand challenge cases for molecular science
The Journal of Chemical Physics **149**, 180901 (2018); https://doi.org/10.1063/1.5052551

A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu
The Journal of Chemical Physics **132**, 154104 (2010); https://doi.org/10.1063/1.3382344



AIP Publishing

The Journal of Chemical Physics ARTICLE scitation.org/journal/jcp

Linear and sublinear scaling computation of the electronic g-tensor at the density functional theory level

Cite as: J. Chem. Phys. 150, 024104 (2019); doi: 10.1063/1.5066266
Submitted: 13 October 2018 • Accepted: 3 December 2018 •
Published Online: 9 January 2019

Michael Glasbrenner, Sigurd Vogler, and Christian Ochsenfeld[a)]

AFFILIATIONS

Chair of Theoretical Chemistry and Center for Integrated Protein Science Munich (CIPSM), Department of Chemistry, University of Munich (LMU), Butenandtstr. 7, 81377 Munich, Germany

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

ABSTRACT

We present an efficient and low-scaling implementation of a density functional theory based method for the computation of electronic g-tensors. It allows for an accurate description of spin-orbit coupling effects by employing the spin-orbit mean-field operator. Gauge-origin independence is ensured by the use of gauge-including atomic orbitals. Asymptotically linear scaling with molecule size is achieved with an atomic orbital based formulation, integral screening methods, and sparse linear algebra. In addition, we introduce an ansatz that exploits the locality of the contributions to the g-tensor for molecules with local spin density. For such systems, sublinear scaling is obtained by restricting the magnetic field perturbation to the relevant subspaces of the full atomic orbital space; several criteria for selecting these subspaces are discussed and compared. It is shown that the computational cost of g-tensor calculations with the local approach can fall below the cost of the self-consistent field calculation for large molecules. The presented methods thus enable efficient, accurate, and gauge-origin independent computations of electronic g-tensors of large molecular systems.

I. INTRODUCTION

The g-factor of the free electron relates its magnetic moment to the electronic spin. It is one of the most precisely known constants in physics,[1,2] and its accurate prediction has historically been one of the largest successes of quantum electrodynamics.[3–6]

In a molecular environment, the effective magnetic moment due to the electron spin is altered compared to the free electron case as a result of spin-orbit coupling. This can be described by substituting the scalar g-factor with the electronic g-tensor, a 3 × 3 matrix, which is in general anisotropic.[7] Electronic g-tensors contain information about the electronic structure of the molecule and are among the most important electron paramagnetic resonance (EPR) parameters.[7] Accurate quantum-chemical calculations of g-tensors are of great interest to experimentalists in the field of EPR spectroscopy as demonstrated by many studies that use theoretically predicted g-tensors to assist in the interpretation of the spectra.[8–23]

The quantum-chemical computation of g-tensors steadily progressed throughout the last decades. One important line of development was the accurate and efficient treatment of spin-orbit coupling effects in molecules. Both the one-electron and two-electron spin-orbit coupling need to be accounted for, of which the latter is computationally more challenging to handle. Due to this difficulty, several *ab initio* studies on g-tensors employed effective nuclear charges in the one-electron spin-orbit coupling contributions in order to account for the screening effects provided by the two-electron contributions.[24–26] In later years, several improved approximations to the two-electron spin-orbit coupling effects were developed. Schreckenbach and Ziegler[27] treated these contributions with the use of an effective potential in the first density functional theory (DFT) method for g-tensors. Furthermore, they also employed gauge-including atomic orbitals (GIAOs),[28–30] which ensure gauge-origin independence of the computed values.[27] Several other accurate approximations to the two-electron spin-orbit coupling contributions have been presented, which involve an effective mean-field operator. Heß *et al.*[31] introduced the spin-orbit mean-field (SOMF) operator. Schimmelpfennig *et al.*[32,33] used the related atomic mean-field integral (AMFI) approach, where in addition all multi-center integrals are neglected. More recently, Neese[34] presented several accurate approximations to the SOMF operator, which reduce the computational effort significantly.

Inclusion of electron correlation has been shown to be crucial for accurate predictions of electronic g-tensors.[35] Apart from DFT based methods,[26,27] also several wave function based methods were presented in the literature; important contributions have been made with configuration interaction theory.[36,37] Coupled-cluster methods for g-tensors were employed in Refs. 35 and 38. Static correlation has been accounted for in the multiconfiguration response theory from Vahtras *et al.*,[39] in several multi-reference configuration interactions approaches,[40–43] and in the recent density-matrix normalization group ansatz from Sayfutyarova and Chan.[44]

Important advances have also been made in the treatment of relativistic effects. van Lenthe *et al.*[45] developed a method based on the zero-order regular approximation. A self-consistent two-component method with a Douglas-Kroll Hamiltonian was introduced by Neyman *et al.*[46] Manninen *et al.*[47] presented a perturbational relativistic theory for g-tensors. Another two-component ansatz was presented by Malkin *et al.*,[48] and a four-component ansatz was introduced by Repiskỳ *et al.*[49]

While much progress has been made in the accurate computation of electronic g-tensors for small molecules, methods that enable to tackle large molecular systems are necessary in order to further increase the area of application for *ab initio* g-tensor calculations. Such methods have to be computationally efficient and display low scaling behavior with the molecule size; furthermore, they should be gauge-origin independent. In a recent benchmark study, we showed that significant errors can be introduced in g-tensor calculations on larger systems, if gauge-origin independence is not ensured by a suitable ansatz.[50]

This work builds on our previously established DFT based ansatz from Ref. 50, which employs the SOMF operator in order to accurately describe the two-electron spin-orbit coupling but does not suffer from gauge-origin dependence. Similar to the method from Schreckenbach and Ziegler[27] and a later coupled cluster approach by Gauss *et al.*,[35] the gauge-origin independence of our method is ensured by the use of GIAOs.[50] In this paper, we present an implementation of this method with improved computational efficiency and reduced scaling due to efficient integral screening. In addition, a local ansatz for treating the magnetic field perturbation is introduced, which can lead to sublinear scaling for molecules with localized spin density. These methods enable efficient and accurate g-tensor calculations at the DFT level for large molecules such as biological macro-molecules.

II. THEORY

A. General theory

The electronic g-tensor can be computed by evaluating the mixed second derivative of the electronic energy E with respect to the magnetic field $\vec{B}$ and the electron spin $\vec{s}$

$$g_{pq} = \frac{1}{\alpha \mathrm{S}} \frac{\mathrm{d}^2\mathrm{E}}{\mathrm{dB}_p \mathrm{d}s_q}\bigg|_{\vec{B}=\vec{0},\vec{s}=\vec{0}}, \qquad p,q \in \{x,y,z\}, \tag{1}$$

with the fine structure constant α.[27] S denotes the total spin of the electronic state ($\frac{1}{2}$ times the number of unpaired electrons).

Five contributions need to be included in order to compute the deviation of the electronic g-tensor from the free electron g-factor g_{el}[27]

$$g_{pq} = \delta_{pq} g_{\mathrm{el}} + g_{pq}^{\mathrm{rmc}} + g_{pq}^{\mathrm{dso1}} + g_{pq}^{\mathrm{dso2}} + g_{pq}^{\mathrm{pso1}} + g_{pq}^{\mathrm{pso2}}. \tag{2}$$

Here, $\mathbf{g}^{\mathrm{rmc}}$ is the relativistic mass correction, $\mathbf{g}^{\mathrm{dso1}}$ and $\mathbf{g}^{\mathrm{dso2}}$ are the diamagnetic one-electron and two-electron spin-orbit terms, respectively, and $\mathbf{g}^{\mathrm{pso1}}$ and $\mathbf{g}^{\mathrm{pso2}}$ denote the paramagnetic one-electron and two-electron spin-orbit terms, respectively. The theoretical approach for calculating these contributions is the same as in our recent work;[50] in our present work, we explore several methods for reducing the computational cost and scaling to linear or sublinear (Secs. II C and II D).

$\mathbf{g}^{\mathrm{rmc}}$ is computed from the following expression containing the kinetic energy integrals **T** and the spin density $\mathbf{P}^{\alpha-\beta}$:[26]

$$g_{pq}^{\mathrm{rmc}} = -\frac{\alpha^2 g_{\mathrm{el}}}{2\mathrm{S}} \delta_{pq} \sum_{\mu\nu} \mathrm{P}_{\mu\nu}^{\alpha-\beta} \mathrm{T}_{\mu\nu}, \tag{3}$$

where the spin density matrix is defined as the difference between α and β density matrices

$$\mathrm{P}_{\mu\nu}^{\alpha-\beta} \equiv \mathrm{P}_{\mu\nu}^{\alpha} - \mathrm{P}_{\mu\nu}^{\beta}. \tag{4}$$

We do not compute the two-electron contributions needed for $\mathbf{g}^{\mathrm{dso2}}$ explicitly, but instead use effective nuclear charges as determined by Koseki *et al.*[51] in the evaluation of $\mathbf{g}^{\mathrm{dso1}}$ in order to approximately account for this contribution

$$\begin{aligned} g_{pq}^{\mathrm{dso1}} + g_{pq}^{\mathrm{dso2}} \approx{}& \frac{\alpha^2 g'}{8\mathrm{S}} \sum_{\mathrm{A}} \tilde{\mathrm{Z}}_{\mathrm{A}} \sum_{\mu\nu} \mathrm{P}_{\mu\nu}^{\alpha-\beta} \\ &\times \Bigg[\int \chi_\mu^*(\vec{r}_1) \frac{\delta_{pq}(\vec{r}_1 - \vec{R}_{\mathrm{A}}) \cdot (\vec{r}_1 - \vec{R}_\nu)}{|\vec{r}_1 - \vec{R}_{\mathrm{A}}|^3} \chi_\nu(\vec{r}_1) d\vec{r}_1 \\ &- \int \chi_\mu^*(\vec{r}_1) \frac{(\vec{r}_1 - \vec{R}_{\mathrm{A}})_p (\vec{r}_1 - \vec{R}_\nu)_q}{|\vec{r}_1 - \vec{R}_{\mathrm{A}}|^3} \chi_\nu(\vec{r}_1) d\vec{r}_1 \Bigg], \end{aligned} \tag{5}$$

where $\tilde{Z}_A$ and $\vec{R}_A$ are the effective charge and position of nucleus A, respectively, and g' is the electronic spin-orbit g-factor.[7] $\vec{R}_\nu$ denotes the atomic center of the basis function χ_ν.

The paramagnetic contributions involve the SOMF operator $\hat{z}$[50]

$$g_{pq}^{\text{pso1}} + g_{pq}^{\text{pso2}} \approx \frac{1}{\alpha S} \sum_{\mu\nu} \left[\left. \frac{dP_{\mu\nu}^{\alpha-\beta}}{dB_p} \right|_{\vec{B}=\vec{0},\vec{s}=\vec{0}} z_{\mu\nu}^{q} + P_{\mu\nu}^{\alpha-\beta} z_{\mu\nu}^{q,B_p} \right]. \tag{6}$$

The matrix representation of the SOMF operator[31] in an atomic orbital basis is obtained from the following set of equations:

$$z_{\mu\nu}^{q} = h_{\mu\nu,q}^{\text{SOMF}} + J_{\mu\nu,q}^{\text{SOMF}} + K_{\mu\nu,q}^{\text{SOMF}}, \qquad q \in \{x, y, z\}, \tag{7}$$

$$h_{\mu\nu,q}^{\text{SOMF}} = \frac{\alpha^2 g'}{4} \sum_{A} Z_A \int \chi_\mu^*(\vec{r}_1) \frac{\left[(\vec{r}_1 - \vec{R}_A) \times \hat{p}\right]_q}{|\vec{r}_1 - \vec{R}_A|^3} \chi_\nu(\vec{r}_1) d\vec{r}_1, \tag{8}$$

$$J_{\mu\nu,q}^{\text{SOMF}} = -\frac{\alpha^2 g'}{4} \sum_{\lambda\kappa} P_{\lambda\kappa} g_{\mu\nu\lambda\kappa}^{\text{soc, q}}, \tag{9}$$

$$K_{\mu\nu,q}^{\text{SOMF}} = \frac{3}{8} \alpha^2 g' \sum_{\lambda\kappa} \left(P_{\lambda\kappa} g_{\mu\lambda\kappa\nu}^{\text{soc, q}} + P_{\lambda\kappa} g_{\kappa\nu\mu\lambda}^{\text{soc, q}} \right). \tag{10}$$

Here, $g_{\mu\nu\lambda\kappa}^{\text{soc, q}}$ denotes two-electron spin-orbit coupling integrals

$$g_{\mu\nu\lambda\kappa}^{\text{soc, q}} = \int\!\!\int \chi_\mu^*(\vec{r}_1) \hat{g}^{\text{soc},q}(\vec{r}_1, \vec{r}_2) \chi_\nu(\vec{r}_1) \chi_\lambda^*(\vec{r}_2) \chi_\kappa(\vec{r}_2) d\vec{r}_1 d\vec{r}_2, \tag{11}$$

with

$$\hat{g}^{\text{soc},q}(\vec{r}_1, \vec{r}_2) = \frac{\left[(\vec{r}_1 - \vec{r}_2) \times \hat{p}\right]_q}{|\vec{r}_1 - \vec{r}_2|^3}. \tag{12}$$

By partial integration, these integrals can be transformed to linear combinations of second derivatives of standard electron repulsion integrals as shown in Ref. 52. Finally, a contribution from the magnetic field derivative of the GIAO basis functions[28–30]

$$\chi_\mu(\vec{r}, \vec{B}) = e^{-\frac{i\alpha}{2}\left[\left(\vec{B}\times(\vec{R}_\mu - \vec{R}_0)\right)\cdot\vec{r}\right]} \chi_\mu(\vec{r}) \tag{13}$$

in the SOMF matrix arises[50]

$$z_{\mu\nu}^{q,B_p} = \frac{i\alpha}{2} \left(\vec{R}_\mu \times \vec{R}_\nu \right)_p z_{\mu\nu}^{q} + \frac{i\alpha}{2} \left[\left(\vec{R}_\mu - \vec{R}_\nu \right) \times z_{\mu\nu^+}^{q} \right]_p, \tag{14}$$

where χ_{ν^+} represents the basis function χ_ν with incremented angular momentum.[50] The direction, in which the angular momentum must be incremented, is determined by the cross product.

B. Coupled-perturbed SCF

For computing the paramagnetic contributions to the g-tensor as in Eq. (6), the derivative of the spin density with respect to the magnetic field is needed. The perturbed spin density $\mathbf{P}^{\alpha-\beta,B_p}$ is obtained from the difference between the perturbed α and β densities

$$\begin{aligned} \left. \frac{dP_{\mu\nu}^{\alpha-\beta}}{dB_p} \right|_{\vec{B}=\vec{0},\vec{s}=\vec{0}} &\equiv P_{\mu\nu}^{\alpha-\beta,B_p} \equiv P_{\mu\nu}^{\alpha,B_p} - P_{\mu\nu}^{\beta,B_p} \\ &\equiv \left. \frac{dP_{\mu\nu}^{\alpha}}{dB_p} \right|_{\vec{B}=\vec{0},\vec{s}=\vec{0}} - \left. \frac{dP_{\mu\nu}^{\beta}}{dB_p} \right|_{\vec{B}=\vec{0},\vec{s}=\vec{0}}. \end{aligned} \tag{15}$$

The perturbed α and β densities can be decomposed into their subspace projections resulting in occupied-occupied (oo), occupied-virtual (ov), virtual-occupied (vo), and virtual-virtual (vv) parts, of which the latter vanishes[53]

$$\begin{aligned} P_{\mu\nu}^{\sigma,B_p} &= \left[P_{\mu\nu}^{\sigma,B_p}\right]_{\text{oo}} + \left[P_{\mu\nu}^{\sigma,B_p}\right]_{\text{vo}} + \left[P_{\mu\nu}^{\sigma,B_p}\right]_{\text{ov}} + \left[P_{\mu\nu}^{\sigma,B_p}\right]_{\text{vv}} \\ &= \left[P_{\mu\nu}^{\sigma,B_p}\right]_{\text{oo}} + \left[P_{\mu\nu}^{\sigma,B_p}\right]_{\text{vo}} - \left[P_{\nu\mu}^{\sigma,B_p}\right]_{\text{vo}}, \end{aligned} \tag{16}$$

where σ denotes the spin (either α or β) and the negative sign in the last line results from the skew-symmetry of the perturbed densities. The occupied-occupied block of the perturbed density is obtained via the following expression, which contains the magnetic field derivative of the overlap matrix $\mathbf{S}$:[53]

$$\left[P_{\mu\nu}^{\sigma,B_p}\right]_{\text{oo}} = -\sum_{\lambda\kappa} P_{\mu\lambda}^{\sigma} \frac{dS_{\lambda\kappa}}{dB_p} P_{\kappa\nu}^{\sigma}. \tag{17}$$

The virtual-occupied block of the perturbed densities can be computed using the density-matrix based Laplace-transformed coupled-perturbed self-consistent field (DL-CPSCF) method[54]

$$\left[P_{\mu\nu}^{\sigma,B_p}\right]_{\text{vo}} = \sum_{\tau} \omega_\tau \sum_{\lambda\kappa} \overline{Q}_{\mu\lambda}^{\sigma,\tau} \left(-h_{\lambda\kappa}^{\sigma,B_p} + K_{\lambda\kappa}^{\sigma,[B_p]} \right) \underline{P}_{\kappa\nu}^{\sigma,\tau}. \tag{18}$$

Here, ω_τ is the weight corresponding to the Laplace quadrature point τ. $\underline{\mathbf{P}}^{\sigma,\tau}$ and $\overline{\mathbf{Q}}^{\sigma,\tau}$ are the occupied and virtual pseudo-density for Laplace point τ, respectively,

$$\underline{P}_{\mu\lambda}^{\sigma,\tau} = \sum_{i} C_{\mu i}^{\sigma *} e^{(\epsilon_i^\sigma - \epsilon_F^\sigma) t_\tau} C_{\lambda i}^{\sigma}, \tag{19}$$

$$\overline{Q}_{\kappa\nu}^{\sigma,\tau} = \sum_{a} C_{\kappa a}^{\sigma *} e^{-(\epsilon_a^\sigma - \epsilon_F^\sigma) t_\tau} C_{\nu a}^{\sigma}, \tag{20}$$

where $\mathbf{C}^\sigma$ are the molecular orbital coefficients and ϵ^σ are the orbital energies. t_τ is the exponent corresponding to Laplace point τ. The sums over i and a run over occupied and virtual molecular orbitals, respectively. The orbital energies are shifted by the Fermi level ϵ_F^σ, which is given by the average of the energies of HOMO and LUMO; this increases the numerical stability in the computation of the pseudo-densities.[55] $\mathbf{h}^{\sigma,B_p}$ includes the integrals of the orbital angular momentum operator $\hat{l}$, all magnetic field derivatives of the GIAO basis functions within the Fock matrix, a contribution from the derivative of the overlap matrix **S**, and an exchange contribution from $\left[\mathbf{P}^{\sigma,\mathrm{B}_p}\right]_{\mathrm{oo}}$[53]

$$\begin{aligned} h_{\mu\nu}^{\sigma,\mathrm{B}_p} = & \frac{\alpha}{2}\left\langle \mu \middle| (r\times\hat{p})_p \middle| \nu \right\rangle + \frac{d\mathrm{T}_{\mu\nu}}{d\mathrm{B}_p} + \frac{d\mathrm{V}_{\mu\nu}}{d\mathrm{B}_p} + \frac{df_{\mu\nu}^{\mathrm{xc},\sigma}}{d\mathrm{B}_p} \\ & + \sum_{\lambda\kappa}\frac{d(\mu\nu|\lambda\kappa)}{d\mathrm{B}_p}\left(\mathrm{P}_{\lambda\kappa}^{\alpha}+\mathrm{P}_{\lambda\kappa}^{\beta}\right) - a_x\sum_{\lambda\kappa}\frac{d(\mu\kappa|\lambda\nu)}{d\mathrm{B}_p}\mathrm{P}_{\lambda\kappa}^{\sigma} \\ & - \sum_{\kappa\lambda}\frac{d\mathrm{S}_{\mu\lambda}}{d\mathrm{B}_p}\mathrm{P}_{\lambda\kappa}^{\sigma}\mathrm{F}_{\kappa\nu}^{\sigma} - a_x\sum_{\kappa\lambda}(\mu\kappa|\lambda\nu)\left[\mathrm{P}_{\lambda\kappa}^{\sigma,\mathrm{B}_p}\right]_{\mathrm{oo}}, \end{aligned} \tag{21}$$

where $\frac{d\mathbf{f}^{\mathrm{xc},\sigma}}{d\mathrm{B}_p}$ denotes the derivative of the exchange-correlation potential, which depends on the density for α and β electrons. **T** and **V** are the kinetic and potential energy integrals, respectively. a_x is the fraction of exact exchange in the employed hybrid DFT functional. Mulliken notation is used for the two-electron integrals

$$(\mu\nu|\lambda\kappa) = \iint \chi_\mu^*(\vec{r}_1)\chi_\nu(\vec{r}_1)\frac{1}{|\vec{r}_1-\vec{r}_2|}\chi_\lambda^*(\vec{r}_2)\chi_\kappa(\vec{r}_2)d\vec{r}_1 d\vec{r}_2. \tag{22}$$

Finally, $\mathbf{K}^{\sigma,[\mathrm{B}_p]}$ is an exchange matrix built with the perturbed density

$$\mathrm{K}_{\mu\nu}^{\sigma,[\mathrm{B}_p]} = a_x\sum_{\lambda\kappa}(\mu\kappa|\lambda\nu)\left(\left[\mathrm{P}_{\kappa\lambda}^{\sigma,\mathrm{B}_p}\right]_{\mathrm{ov}} + \left[\mathrm{P}_{\kappa\lambda}^{\sigma,\mathrm{B}_p}\right]_{\mathrm{vo}}\right). \tag{23}$$

Therefore, the right-hand side of Eq. (18) depends on the perturbed density itself and the equation has to be solved iteratively. The exchange contribution from $\left[\mathbf{P}^{\sigma,\mathrm{B}_p}\right]_{\mathrm{oo}}$ has been added to $\mathbf{h}^{\sigma,\mathrm{B}_p}$ in Eq. (21), as it is constant throughout the CPSCF iterations and thus needs to be computed only once. Note that no Coulomb term appears in Eq. (23) as it vanishes due to the skew-symmetry of the perturbed densities for the imaginary magnetic field perturbation.[53]

C. Linear scaling approach using integral screening and sparse linear algebra

The relativistic mass contribution and the diamagnetic contributions to the g-tensor are evaluated non-iteratively and involve only one-electron integrals [see Eqs. (3) and (5)]. Therefore, these contributions are computationally inexpensive; if necessary, one could obtain them with sublinear scaling effort for systems with local spin densities by restricting the summations in Eqs. (3) and (5) to basis function pairs with significant entries in $\mathbf{P}^{\alpha-\beta}$. We will focus our attention on the computationally more demanding paramagnetic contributions, which involve the iterative solution of CPSCF equations and thus constitute a potential bottleneck. Computationally demanding steps include the computation of $\mathbf{h}^{\sigma,\mathrm{B}_p}$ [Eq. (21)], the matrix multiplications in Eq. (18), and the construction of $\mathbf{K}^{\sigma,[\mathrm{B}_p]}$ according to Eq. (23). With the perturbed densities from the converged CPSCF at hand, the contributions to the g-tensor from the SOMF-operator can be computed. The one-electron contribution from Eq. (8) is not computationally expensive in contrast to the Coulomb contribution from Eq. (9) and the exchange contribution from Eq. (10). In this section, we will explain how the computationally demanding steps can be implemented efficiently and in an asymptotically linear scaling fashion using integral screening methods and sparse linear algebra. We will denote the new method as "SOMF-GIAO" henceforth.

1. Linear scaling CPSCF

Usually, the most expensive step of the g-tensor calculation is the construction of $\mathbf{K}^{\sigma,[\mathrm{B}_p]}$ [Eq. (23)] in each iteration of the CPSCF. For this step, we employ LinK screening[56,57] which reduces the computational effort to $\mathcal{O}(\mathrm{N})$ for sparse perturbed density matrices.

In order to speed up the matrix multiplications needed in Eqs. (17) and (18), we save the matrices in the block-compressed sparse row (BCSR) format[53,58] such that the inherent block-sparsity of these matrices can be exploited.[53] The atoms and their shells of associated basis functions are sorted via the reverse Cuthill-McKee algorithm[59] in order to increase the block sparsity. For large systems with sufficiently local electronic structures, this allows scaling of the matrix multiplications to be reduced to $\mathcal{O}(\mathrm{N})$.

The most expensive steps in the computation of $\mathbf{h}^{\sigma,\mathrm{B}_p}$ [Eq. (21)] are the contributions involving the magnetic-field derivatives of the two-electron integrals and of the exchange-correlation potential matrix. Their efficient evaluation has also been discussed in Ref. 60. The Coulomb contribution

$$\mathrm{J}_{\mu\nu}^{\mathrm{B}_p} \equiv \sum_{\lambda\kappa}\frac{d(\mu\nu|\lambda\kappa)}{d\mathrm{B}_p}\left(\mathrm{P}_{\lambda\kappa}^{\alpha}+\mathrm{P}_{\lambda\kappa}^{\beta}\right) \tag{24}$$

can be calculated efficiently and in a linear scaling fashion using the continuous fast multipole method (CFMM).[60–62] For the near-field contribution, we use Schwarz-screening for the two-electron integrals.[63,64]

For the exchange term

$$\mathrm{K}_{\mu\nu}^{\sigma,\mathrm{B}_p} \equiv a_x\sum_{\lambda\kappa}\frac{d(\mu\kappa|\lambda\nu)}{d\mathrm{B}_p}\mathrm{P}_{\lambda\kappa}^{\sigma}, \tag{25}$$

we apply LinK-screening,[56,57] which takes the coupling of the bra- and ket-sides of the integrals by the density matrix into account. For systems with a sparse density matrix (which is generally the case for systems with a non-vanishing HOMO-LUMO gap), $\mathbf{K}^{\sigma,\mathrm{B}_p}$ can thus be obtained with linear scaling effort.

Finally, $\frac{d\mathbf{f}^{\mathrm{xc}}}{d\mathrm{B}_p}$ is computed for local batches of grid points. Each batch has an associated list of significant shell-pairs of basis functions, which includes all basis functions with significant values at the grid points in the batch. For each batch of grid points, only $\mathcal{O}(1)$ shell-pairs are significant. In this way, the derivative of the exchange-correlation potential can also be evaluated with linear scaling compute time.

2. Coulomb contribution to the SOMF operator

The SOMF matrices from Eq. (7) do not have to be explicitly constructed for the purpose of calculating the g-tensor. Instead, one can contract the spin-orbit integrals directly with the density matrices. For the Coulomb contribution in Eq. (9), this gives the following contributions to the g-tensor:

$$g_{pq}^{\mathrm{SOMF\text{-}J}} = -\frac{\alpha g'}{4S} \sum_{\mu\nu\lambda\kappa} \mathrm{P}_{\mu\nu}^{\alpha-\beta,\mathrm{B}_p} g_{\mu\nu\lambda\kappa}^{\mathrm{soc,\,q}} \mathrm{P}_{\lambda\kappa}, \tag{26}$$

$$\tilde{g}_{pq}^{\mathrm{SOMF\text{-}J}} = -\frac{\alpha g'}{4S} \sum_{\mu\nu\lambda\kappa} \frac{i\alpha}{2} \mathrm{P}_{\mu\nu}^{\alpha-\beta} \Big[(\vec{\mathrm{R}}_\mu \times \vec{\mathrm{R}}_\nu)_p \, g_{\mu\nu\lambda\kappa}^{\mathrm{soc,\,q}} + (\vec{\mathrm{R}}_{\mu\nu} \times g_{\mu\nu^+\lambda\kappa}^{\mathrm{soc,\,q}})_p \Big] \mathrm{P}_{\lambda\kappa}, \tag{27}$$

with

$$\vec{\mathrm{R}}_{\mu\nu} \equiv \vec{\mathrm{R}}_\mu - \vec{\mathrm{R}}_\nu. \tag{28}$$

The contributions from Eqs. (26) and (27) are computed together in our implementation such that common intermediates in the integral calculation can be used. The formal scaling of the Coulomb contributions is N^4. This can be reduced to $\mathcal{O}(\mathrm{N}^2)$ by neglecting shell-pairs $\mu\nu$ and $\lambda\sigma$ with vanishing overlap.[34,63] As discussed in Sec. IV A, the perturbed spin densities are local in many systems; if their locality is exploited, the scaling can be further reduced to linear because only $\mathcal{O}(1)$ bra shell-pairs $\mu\nu$ are significant.

An even more efficient screening protocol has to take the decay of the two-electron spin-orbit integrals with increasing bra-ket separation into account. These integrals show anisotropic decay behavior and decay in one direction like r_{12}^{-2} and in two directions like r_{12}^{-3}, which can be observed by inspecting the operator in Eq. (12). In order to exploit the fact that these integrals decay at least as r_{12}^{-2} with increasing bra-ket separation, we employ the distance-including screening technique called QQR[65] with an exponent of 2.0. We do not account for the anisotropic decay behavior. In our implementation, contributions from shell-quartets $\mu\nu\lambda\kappa$ are only computed if the following inequality holds:

$$\frac{\mathrm{P}_{\mu\nu}^{\alpha-\beta,\mathrm{max}} \mathrm{Q}_{\mu\nu} \mathrm{Q}_{\lambda\kappa} \mathrm{P}_{\lambda\kappa}^{\mathrm{max}}}{|\mathrm{R}_{\mu\nu-\lambda\kappa} - \mathrm{ext}'_{\mu\nu} - \mathrm{ext}'_{\lambda\kappa}|^2} < \vartheta_{\mathrm{QQR}}, \tag{29}$$

where ϑ_{QQR} is the screening threshold. **Q** is an intermediate that is commonly employed in integral screening based on the Cauchy-Schwarz inequality[63,64]

$$\mathrm{Q}_{\mu\nu} = |(\mu\nu|\mu\nu)|^{1/2}. \tag{30}$$

$\mathrm{Q}_{\mu\nu}\mathrm{Q}_{\lambda\kappa}$ is the Schwarz estimate for a standard electron repulsion integral (ERI) $(\mu\nu|\lambda\kappa)$.[63,64] We use this estimate to account for the exponential decay behavior of the integrals with separation between the basis functions in both bra and ket. $P_{\mu\nu}^{\alpha-\beta,\mathrm{max}}$ is the maximum value of the maximal values of spin density and perturbed spin densities for the shell-pair $\mu\nu$

$$P_{\mu\nu}^{\alpha-\beta,\mathrm{max}} \equiv \max_{\mu\nu} \Big[\Big(\frac{\alpha}{2} |\mathrm{P}_{\mu\nu}^{\alpha-\beta}| \Big), |\mathrm{P}_{\mu\nu}^{\alpha-\beta,\mathrm{B}_x}|, |\mathrm{P}_{\mu\nu}^{\alpha-\beta,\mathrm{B}_y}|, |\mathrm{P}_{\mu\nu}^{\alpha-\beta,\mathrm{B}_z}| \Big]. \tag{31}$$

By including $\mathrm{P}_{\mu\nu}^{\alpha-\beta,\mathrm{max}}$ in the screening from Eq. (29), calculation of integrals is required only for $\mathcal{O}(1)$ shell-pairs $\mu\nu$ for systems with local (perturbed) spin density. $\mathrm{P}_{\mu\nu}^{\mathrm{max}}$ is the maximal value of the ground state density matrix for this shell-pair

$$\mathrm{P}_{\mu\nu}^{\mathrm{max}} \equiv \max_{\mu\nu} |\mathrm{P}_{\mu\nu}|. \tag{32}$$

$R_{\mu\nu-\lambda\kappa}$ is the distance between the center of shell-pair $\mu\nu$ and the shell-pair $\lambda\kappa$.[65] The shell-pair extent $\mathrm{ext}'_{\mu\nu}$ is defined in Ref. 65.

3. Exchange contribution to the SOMF operator

The contribution from the exchange part of the SOMF operator to the g-tensor can also be evaluated by directly contracting the spin-orbit integrals with the density matrices leading to the contributions from Eqs. (33) and (34). In Eq. (33), the bra and ket sides of the two-electron spin-orbit integrals are coupled via the ground state density matrix **P** and the perturbed spin density matrices $\mathbf{P}^{\alpha-\beta,\mathrm{B}_p}$

$$g_{pq}^{\mathrm{SOMF\text{-}K}} = \frac{3\alpha g'}{8S} \sum_{\mu\nu\lambda\kappa} \Big[\mathrm{P}_{\mu\nu}^{\alpha-\beta,\mathrm{B}_p} \mathrm{P}_{\lambda\kappa} \Big(g_{\mu\lambda\kappa\nu}^{\mathrm{soc,\,q}} + g_{\kappa\nu\mu\lambda}^{\mathrm{soc,\,q}} \Big) \Big]. \tag{33}$$

In the corresponding GIAO contribution [see also Eq. (14)], the bra and ket sides of the integrals are coupled via **P** and the spin density matrix $\mathbf{P}^{\alpha-\beta}$

$$\tilde{g}_{pq}^{\mathrm{SOMF\text{-}K}} = \frac{3\alpha g'}{8S} \sum_{\mu\nu\lambda\kappa} \frac{i\alpha}{2} \mathrm{P}_{\mu\nu}^{\alpha-\beta} \mathrm{P}_{\lambda\kappa} \Big[(\vec{\mathrm{R}}_\mu \times \vec{\mathrm{R}}_\nu)_p \Big(g_{\mu\lambda\kappa\nu}^{\mathrm{soc,\,q}} + g_{\kappa\nu\mu\lambda}^{\mathrm{soc,\,q}} \Big) + \Big(\vec{\mathrm{R}}_{\mu\nu} \times \Big(g_{\mu\lambda\kappa\nu^+}^{\mathrm{soc,\,q}} + g_{\kappa\nu^+\mu\lambda}^{\mathrm{soc,\,q}} \Big) \Big)_p \Big]. \tag{34}$$

Similarly to the Coulomb contributions from Sec. II C 2, we evaluate both contributions from Eqs. (33) and (34) together in order to use intermediates in the integral calculation. For an efficient computation of the exchange contribution, it is crucial to exploit the sparsity and the locality in the density matrices in addition to the sparsity of the integrals. To this end, we use a screening that is based on Cauchy-Schwarz screening for the integrals but also takes the density matrices into account; the integrals are only computed for a shell quartet $\mu\nu\lambda\kappa$ if the following inequality is fulfilled:

$$\mathrm{P}_{\mu\nu}^{\alpha-\beta,\mathrm{max}} \mathrm{Q}_{\mu\lambda} \mathrm{Q}_{\nu\kappa} \mathrm{P}_{\lambda\kappa}^{\mathrm{max}} > \vartheta_{\mathrm{SOMF\text{-}K}}. \tag{35}$$

$\mathrm{P}_{\mu\nu}^{\alpha-\beta,\mathrm{max}}$ and $\mathrm{P}_{\lambda\kappa}^{\mathrm{max}}$ have been defined in Eqs. (31) and (32), respectively, and account for the coupling of bra and ket sides of the integrals by the density matrices.

D. Local perturbation approach for g-tensors

As described in Sec. IV A, the perturbed spin densities in Eq. (15) are, like the unperturbed spin density, local quantities for many larger molecules. For radicals with one or more localized spin centers, perturbed spin densities only occur in the immediate vicinity of the spin centers. In contrast, the magnetic field derivatives of the α and β density matrices are in general global quantities with significant entries for basis function pairs throughout the entire molecule. However, most of the contributions from α and β spin cancel, when the difference between the two spins is computed according to Eq. (15), resulting in a local perturbed spin density matrix. We will denote the space of basis functions that are part of AO pairs with a significant perturbed spin density by Ω_{P} henceforth.

It is highly desirable to altogether avoid the expensive computation of entries in the perturbed α and β densities that do not lead to significant contributions to the perturbed spin density because of cancellation during the evaluation of Eq. (15). In order to achieve this, we extend the SOMF-GIAO method from Sec. II C by employing a "locally projected perturbation" (LPP) approach. For this purpose, we rewrite Eq. (18) as

$$\left[P_{\mu\nu}^{\sigma,B_p}\right]_{\mathrm{vo}} = \left[\tilde{P}_{\mu\nu}^{\sigma,B_p}\right]_{\mathrm{vo}} + \sum_{\tau} \omega_\tau \sum_{\lambda\kappa} \overline{Q}_{\mu\lambda}^{\sigma,\tau} K_{\lambda\kappa}^{\sigma,[B_p]} \underline{P}_{\kappa\nu}^{\sigma,\tau}, \tag{36}$$

with the uncoupled-perturbed density $\left[\tilde{\mathbf{P}}^{\sigma,B_p}\right]_{\mathrm{vo}}$

$$\left[\tilde{P}_{\mu\nu}^{\sigma,B_p}\right]_{\mathrm{vo}} \equiv \sum_{\tau} \omega_\tau \sum_{\lambda\kappa} \overline{Q}_{\mu\lambda}^{\sigma,\tau} \left(-h_{\lambda\kappa}^{\sigma,B_p}\right) \underline{P}_{\kappa\nu}^{\sigma,\tau}. \tag{37}$$

$\left[\tilde{\mathbf{P}}^{\sigma,B_p}\right]_{\mathrm{vo}}$ can be computed prior to the CPSCF iterations, as it does not change during the iterations. In our local perturbation approach, we project $\mathbf{h}^{\sigma,B_p}$, $\left[\tilde{\mathbf{P}}^{\sigma,B_p}\right]_{\mathrm{vo}}$, or both quantities to the relevant subspaces of the full atomic orbital space; this has a similar effect as turning the magnetic field perturbation on only in certain parts of the molecule. The approach is intended to allow for a local computation of perturbed α and β densities in the relevant parts of the molecule without computing the perturbed densities globally. This method will be denoted as "LPP-SOMF-GIAO" henceforth.

In our first projection approach, the uncoupled-perturbed densities are projected onto $\tilde{\Omega}_{\mathrm{P}}$ by zeroing all entries for basis functions that are not contained in $\tilde{\Omega}_{\mathrm{P}}$, where $\tilde{\Omega}_{\mathrm{P}}$ shall denote an estimate of the space Ω_{P}:

$$\left[\tilde{\mathbf{P}}^{\sigma,B_p}\right]_{\mathrm{vo}} \Rightarrow \mathbf{o}^{\tilde{\Omega}_{\mathrm{P}}} \left[\tilde{\mathbf{P}}^{\sigma,B_p}\right]_{\mathrm{vo}} \mathbf{o}^{\tilde{\Omega}_{\mathrm{P}}}, \tag{38}$$

$$o_{\mu\nu}^{\tilde{\Omega}_{\mathrm{P}}} = \begin{cases} \delta_{\mu\nu} & \text{if } \mu \in \tilde{\Omega}_{\mathrm{P}} \\ 0 & \text{otherwise} \end{cases}. \tag{39}$$

Methods for obtaining $\tilde{\Omega}_{\mathrm{P}}$ will be presented in the following. When the projected uncoupled-perturbed density is inserted into Eq. (36), a modified DL-CPSCF equation is obtained

$$\left[P_{\mu\nu}^{\sigma,B_p}\right]_{\mathrm{vo}} = \sum_{\lambda\kappa} o_{\mu\lambda}^{\tilde{\Omega}_{\mathrm{P}}} \left[\tilde{P}_{\lambda\kappa}^{\sigma,B_p}\right]_{\mathrm{vo}} o_{\kappa\nu}^{\tilde{\Omega}_{\mathrm{P}}} + \sum_{\tau} \omega_\tau \sum_{\lambda\kappa} \overline{Q}_{\mu\lambda}^{\sigma,\tau} K_{\lambda\kappa}^{\sigma,[B_p]} \underline{P}_{\kappa\nu}^{\sigma,\tau}. \tag{40}$$

In cases where $\tilde{\Omega}_{\mathrm{P}}$ is significantly smaller than the full AO space, a *local* perturbed density is obtained as the solution of Eq. (40), as the projection of $\left[\tilde{\mathbf{P}}^{\sigma,B_p}\right]_{\mathrm{vo}}$ has a similar effect as turning the magnetic field perturbation on only in a small part of the molecule. For such a local perturbed density, $\mathbf{K}^{\sigma,[B_p]}$ can be built with $\mathcal{O}(1)$ computational cost if LinK screening is employed,[66] thus reducing the scaling of the rate-determining step to sublinear. Also the matrix multiplications of the local matrix $\mathbf{K}^{\sigma,[B_p]}$ with the pseudo-densities can be performed with sublinear effort if sparse linear algebra with the BCSR matrix format is used.[66]

The first method for obtaining an estimate $\tilde{\Omega}_{\mathrm{P}}$ involves the uncoupled-perturbed spin density $\tilde{\mathbf{P}}^{\alpha-\beta,B_p}$. It is obtained from the difference between the uncoupled-perturbed densities $\tilde{\mathbf{P}}^{\alpha,B_p}$ and $\tilde{\mathbf{P}}^{\beta,B_p}$, which can be viewed as approximate perturbed densities

$$\left[\tilde{P}_{\mu\nu}^{\alpha-\beta,B_p}\right]_{\mathrm{vo}} = \left[\tilde{P}_{\mu\nu}^{\alpha,B_p}\right]_{\mathrm{vo}} - \left[\tilde{P}_{\mu\nu}^{\beta,B_p}\right]_{\mathrm{vo}}. \tag{41}$$

In a similar way, Wolinski *et al.*[30] used the uncoupled-perturbed density in order to estimate the perturbed density prior to the CPSCF in the context of NMR shieldings.

Then, for each atom A, the following quantity $q_{\mathrm{A}}^{\mathrm{P1}}$ is formed:

$$\begin{aligned} q_{\mathrm{A}}^{\mathrm{P1}} &\equiv \sum_{\mu\in\{\mathrm{A}\}} \sum_{\nu} \max_{\{p\}} \left| \left[\tilde{P}_{\mu\nu}^{\alpha-\beta,B_p}\right]_{\mathrm{vo}} + \left[\tilde{P}_{\mu\nu}^{\alpha-\beta,B_p}\right]_{\mathrm{ov}} \right| \\ &= \sum_{\mu\in\{\mathrm{A}\}} \sum_{\nu} \max_{\{p\}} \left| \left[\tilde{P}_{\mu\nu}^{\alpha-\beta,B_p}\right]_{\mathrm{vo}} - \left[\tilde{P}_{\nu\mu}^{\alpha-\beta,B_p}\right]_{\mathrm{vo}} \right|, \end{aligned} \tag{42}$$

where $\{\mathrm{A}\}$ shall denote the set of basis functions centered on atom A. $q_{\mathrm{A}}^{\mathrm{P1}}$ is used as a measure for the importance of the basis functions on atom A for the computation of the perturbed densities. One could define $\tilde{\Omega}_{\mathrm{P}}$ as the set of all basis functions that are centered on atoms with q^{P1} values above a pre-defined threshold. However, this approach has certain drawbacks because the distribution of q^{P1} values can be rather different for different molecules depending on its electronic structure. Using a fixed threshold can be sub-optimal for some systems. For this reason, we use an adaptive, system-specific threshold and include only basis functions on atoms with a q^{P1} value larger than $\vartheta_{\text{proj-P1}}$ in the space $\tilde{\Omega}_{\mathrm{P}}$

$$\vartheta_{\text{proj-P1}} \equiv \vartheta_{\mathrm{P1}}^{90\%}\, \vartheta_{\text{proj-P1}}^{\text{user}}. \tag{43}$$

Here, $\vartheta_{\text{proj-P1}}^{\text{user}}$ is a threshold that is specified in the user input file. $\vartheta_{\mathrm{P1}}^{90\%}$ is the threshold for which 90% of $\sum_{\mathrm{A}} q_{\mathrm{A}}^{\mathrm{P1}}$ would be included and is dependent on the molecule. In order to obtain $\vartheta_{\mathrm{P1}}^{90\%}$, a list of q^{P1} values for all atoms is constructed and sorted by magnitude; starting with the atoms with largest q^{P1} values, more and more atoms are added until 90% of $\sum_{\mathrm{A}} q_{\mathrm{A}}^{\mathrm{P1}}$ is included. The q^{P1} value of the last atom needed to exceed 90% of $\sum_{\mathrm{A}} q_{\mathrm{A}}^{\mathrm{P1}}$ is then taken as $\vartheta_{\mathrm{P1}}^{90\%}$. The described projection approach based on the uncoupled-perturbed spin density will be denoted as proj-P1 henceforth.

An alternative way of obtaining an estimate for $\tilde{\Omega}_{\mathrm{P}}$ was implemented which uses the unperturbed spin density

$$q_{\mathrm{A}}^{\mathrm{P2}} \equiv \sum_{\mu\in\{\mathrm{A}\}} \sum_{\nu} |\mathrm{P}_{\mu\nu}^{\alpha-\beta}|. \tag{44}$$

The corresponding projection approach is in all other respects analogous to proj-P1 and will be denoted as proj-P2. Here, it is assumed that the spatial distribution of the unperturbed and perturbed spin densities is similar. In this context, we want to point out that the perturbed spin density as defined in Eq. (15) vanishes exactly for a closed-shell reference state where the α and β density matrices are identical (and the spin density is zero). This can be seen from the CPSCF equations in the DL-CPSCF formulation [Eq. (18)]. This equation can be written in a way such that only the spin-dependent quantities are the perturbed density matrices and the unperturbed density matrices. For this purpose, one needs to express the pseudo-density matrices as the following matrix exponentials:

$$\underline{\mathbf{P}}^{\sigma,\tau} = e^{\mathrm{t}_\tau \mathbf{P}^\sigma \mathbf{F}^\sigma} \mathbf{P}^\sigma, \tag{45}$$

$$\overline{\mathbf{Q}}^{\sigma,\tau} = e^{-\mathrm{t}_\tau \mathbf{Q}^\sigma \mathbf{F}^\sigma} \mathbf{Q}^\sigma \tag{46}$$

and use the definition of the Fock matrix

$$\mathrm{F}_{\mu\nu}^{\sigma} = h_{\mu\nu} + \sum_{\lambda\kappa} \left[\mathrm{P}_{\lambda\kappa}^{\alpha} + \mathrm{P}_{\lambda\kappa}^{\beta}\right](\mu\nu|\lambda\kappa) - a_x \sum_{\lambda\kappa} \mathrm{P}_{\lambda\kappa}^{\sigma}(\mu\kappa|\lambda\nu), \tag{47}$$

as well as the following relation between the occupied density matrix $\mathbf{P}$ and the virtual density matrix $\mathbf{Q}$:

$$\mathbf{Q}^\sigma = \mathbf{S}^{-1} - \mathbf{P}^\sigma. \tag{48}$$

Therefore, the DL-CPSCF equations describe a (rather complicated) relationship between the perturbed and unperturbed density matrices. This relationship is the same for both spin cases. If identical density matrices for α and β spin enter these equations for the case of a closed-shell state, the obtained perturbed α and β densities are therefore identical. Consequently, the perturbed spin density vanishes for a closed-shell state due to the symmetry between α and β spin. This implies that differences between α and β density matrices are a necessary prerequisite for a non-zero perturbed spin density as defined in Eq. (15).

We furthermore argue that the effects of local differences between α and β density are also local and lead to perturbed spin densities in the same spatial region. This is because all spin-dependent quantities in the CPSCF equations that can break the symmetry between α and β spin are sparse and short-ranged. These include the exchange matrices that occur in the Fock matrix [Eq. (47)] and the matrices $\mathbf{K}^{\mathrm{B}}$ in Eq. (25). The Coulomb-type matrices appearing in the CPSCF equations ($\mathbf{J}$ and $\mathbf{J}^{\mathrm{B}}$) depend on the total density ($\mathbf{P}^\alpha + \mathbf{P}^\beta$) and enter the equations for α and β spin identically; therefore these long-ranged contributions do not affect the spin symmetry. Consequently, one can expect the unperturbed and perturbed spin densities to share the same locality. This motivates the use of the unperturbed spin density as a criterion for determining the space $\tilde{\Omega}_{\mathrm{P}}$ in the approach proj-P2.

In a GIAO basis, the computation of $\mathbf{h}^{\sigma,\mathrm{B}_p}$ in Eq. (21) can become the bottleneck for large systems, if the scaling of the iterative part of the CPSCF is reduced to sublinear by the projection approaches proj-P1 or proj-P2. If the uncoupled-perturbed densities are only computed in a local subspace $\tilde{\Omega}_{\mathrm{P}}$, then it is possible to project also $\mathbf{h}^{\sigma,\mathrm{B}_p}$ onto a suitable local subspace $\tilde{\Omega}_h$ without introducing significant errors. The space $\tilde{\Omega}_h$ needs to be chosen large enough such that all significant effects of $\mathbf{h}^{\sigma,\mathrm{B}_p}$ on the perturbed densities in $\tilde{\Omega}_{\mathrm{P}}$ are accounted for. As one can see from Eq. (18), the contributions from $\mathbf{h}^{\sigma,\mathrm{B}_p}$ to the perturbed densities are obtained by matrix multiplications involving the occupied and virtual pseudo-density matrices. Entries in $\mathbf{h}^{\sigma,\mathrm{B}_p}$ from basis function pairs without significant coupling to the space $\tilde{\Omega}_{\mathrm{P}}$ through the pseudo-densities thus have no significant influence on the perturbed densities within $\tilde{\Omega}_{\mathrm{P}}$. In order to determine the space $\tilde{\Omega}_h$, we therefore employ the following auxiliary quantity $\mathbf{M}$:

$$\mathrm{M}_{\mu\nu}^{\sigma} \equiv \sum_{\tau} \omega_\tau \sum_{\lambda\kappa} \underline{\mathrm{P}}_{\mu\lambda}^{\sigma,\tau} |\mathrm{P}_{\lambda\kappa}^{\alpha-\beta}| \overline{\mathrm{Q}}_{\kappa\nu}^{\sigma,\tau}. \tag{49}$$

In Eq. (49), the spin density has been inserted because it can be used to determine $\tilde{\Omega}_{\mathrm{P}}$ as discussed in the context of projection approach proj-P2. We account for the coupling of the space $\tilde{\Omega}_{\mathrm{P}}$ to the surrounding environment by the matrix multiplications with the pseudo-densities. The quantity $\mathbf{M}$ is used for determining the space $\tilde{\Omega}_{\mathrm{h}}$. For each atom A, we form a quantity $q_{\mathrm{A}}^{\mathrm{h}}$

$$q_{\mathrm{A}}^{\mathrm{h}} = \sum_{\mu\in\mathrm{A}} \sum_{\nu} \max_{\{\alpha,\beta\}}\left(|\mathrm{M}_{\mu\nu}^{\sigma}| + |\mathrm{M}_{\nu\mu}^{\sigma}|\right). \tag{50}$$

In complete analogy to the projection approaches proj-P1 and proj-P2, a sorted list of $q_{\mathrm{A}}^{\mathrm{h}}$ for all atoms is constructed. The $q_{\mathrm{A}}^{\mathrm{h}}$ value needed for exceeding 90% of $\sum_{\mathrm{A}} q_{\mathrm{A}}^{\mathrm{h}}$ is taken as $\vartheta_{\mathrm{h}}^{90\%}$. The used system-specific threshold $\vartheta_{\text{proj-h}}$ is then obtained from

$$\vartheta_{\text{proj-h}} \equiv \vartheta_{\mathrm{h}}^{90\%}\, \vartheta_{\text{proj-h}}^{\text{user}}, \tag{51}$$

where $\vartheta_{\text{proj-h}}^{\text{user}}$ is pre-defined in the input. The space $\tilde{\Omega}_h$ then includes all basis functions on atoms with a q^{h} value larger than $\vartheta_{\text{proj-h}}$.

All matrix elements $h_{\mu\nu}^{\sigma,\mathrm{B}_p}$, where χ_μ and χ_ν do not both belong to $\tilde{\Omega}_h$, are set to zero and we will refer to this as the proj-h approach henceforth. In order to speed up the calculation of $\mathbf{h}^{\sigma,\mathrm{B}_p}$, one needs to avoid the computation of entries that would afterwards be removed by the projection—at least for the most expensive steps. These include the Coulomb contribution $\mathbf{J}^{\mathrm{B}_p}$, the exchange contributions from $\mathbf{K}^{\sigma,\mathrm{B}_p}$ in Eq. (25) and from $\left[\mathbf{P}^{\sigma,\mathrm{B}_p}\right]_{\mathrm{oo}}$ in Eq. (21), and finally also the evaluation of $\frac{\mathrm{d}\mathbf{f}^{\mathrm{xc}}}{\mathrm{dB}_p}$. In the computation of $\frac{\mathrm{d}\mathbf{f}^{\mathrm{xc}}}{\mathrm{dB}_p}$, a loop over batches of spatially close grid points is carried out and each batch has an associated list of significant shell-pairs. When applying proj-h, the grid batches are skipped if none of the shell-pairs containing basis functions from $\tilde{\Omega}_h$ are considered significant for them. In this way, $\mathbf{h}^{\sigma,\mathrm{B}_p}$ can be obtained in a sublinear scaling fashion.

The projection approaches proj-P2 and proj-P1 can also be combined with proj-h. The combined approaches allow for overall sublinear scaling of all computationally demanding steps and are denoted by proj-P1h and proj-P2h, respectively.

III. COMPUTATIONAL DETAILS

The theory described in Sec. II was implemented in a development version of the FermiONs++ program.[67–69] The reference state is obtained via unrestricted Kohn-Sham-DFT.[70] The root mean square deviation of ($\mathbf{FPS} - \mathbf{SPF}$) from zero is used as the SCF convergence criterion and converged to a threshold of 10^{-7}. Shell-pairs with a maximum basis function overlap of less than 10^{-12} were omitted from the calculation. Direct inversion of the iterative subspace (DIIS)[71] is used for accelerating and stabilizing the convergence of the CPSCF. In order to test for convergence, the paramagnetic one-electron contribution from Eq. (8) to the g-tensor is computed in each CPSCF iteration; the iterations are stopped if the relative change in this contribution falls below a predefined threshold; this threshold was taken to be 10^{-6} for the accuracy benchmark in Secs. IV B and IV C and 10^{-4} for the timings in Sec. IV D. For the BCSR matrices, we use a sparsity threshold of 10^{-7}.

In all calculations, the B3LYP functional[72] was employed. The molecular integration grids used for DFT are generated as a product of a spherical Lebedev/Laikov[73] grid and a Treutler-Ahlrichs "M4" grid.[74] Separate sets of Laplace quadrature points are used for α and β in Eq. (18), which are obtained from a minimax optimization[75,76] in the interval $\left[1, \frac{\Delta\epsilon_{\max}}{\Delta\epsilon_{\min}/2}\right]$ using a sloppy Remez algorithm.[75] Here, $\Delta\epsilon_{\min}$ denotes the HOMO-LUMO gap and $\Delta\epsilon_{\max}$ is the maximum difference between occupied and virtual orbital energies.

All molecular geometries used for benchmarking the accuracy of our method in Sec. IV C were optimized using the FermiONs++ program[67–69] and the PBE-H3c method[77] with geometric counterpoise-correction.[78] The spin-labeled DNA structures used in Sec. IV D were optimized with FermiONs++,[67–69] the HF-3c method,[79] and the DL-FIND optimizer.[80] All timings in Sec. IV D were performed on a single compute node with two Intel Xeon E5-2667 v4 processors with 3.2 GHz clock speed and 16 cores in total.

IV. RESULTS

A. Locality of the spin density and perturbed spin densities

The approaches presented in Sec. II involve an estimation procedure for the perturbed spin densities and then project $\tilde{\mathbf{P}}^{\sigma,\mathrm{B}_p}$ and $\mathbf{h}^{\sigma,\mathrm{B}_p}$ onto the relevant subspace of the full AO space. This reduces the asymptotic scaling from linear to sublinear by avoiding the computation of the perturbed α and β densities, which are in general global quantities. Instead it is assumed that the perturbed spin densities are often (not for all molecules) local quantities making a local and sublinear scaling computation possible. The approaches proj-P2 and proj-h further assume that the spatial distribution of the unperturbed spin density is a good estimate of the spatial distribution of the perturbed spin densities. In order to support these assumptions, the plots of the mentioned density matrices are shown in Fig. 1 for three exemplary molecules. In contrast to the completely delocalized perturbed α and

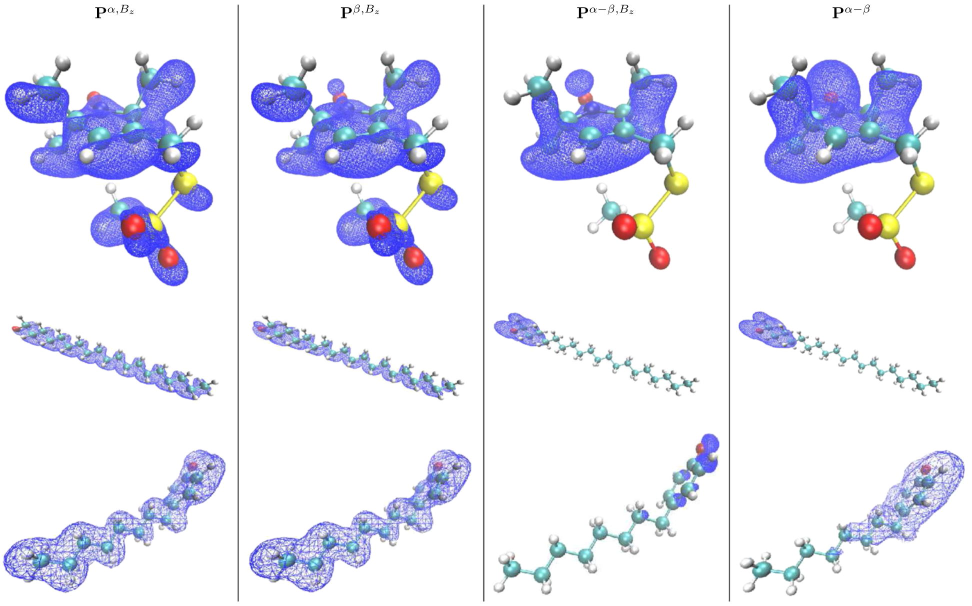


FIG. 1. Plots of $\sum_{\mu\nu} |X_{\mu\nu}|\chi^*_\mu(r)\chi_\nu(r)$, where **X** is the perturbed α density (first column), the perturbed β density (second column), the perturbed spin density (third column), or the unperturbed spin density (fourth column). Depicted are iso-surfaces, which include 90% of the respective density. For all perturbed densities, the magnetic field in the z-direction was taken as the perturbation. The employed molecules are MTSL (first row), ·O–[CH_2]$_{18}$–[CH_3] (middle row), and octylphenol radical (bottom row). The molecule images were created with the Visual Molecular Dynamics (VMD) program.[81,82]

β densities, the perturbed spin density has significant values only in the parts of the molecule that are close to the spin center. Both the perturbed spin densities and the unperturbed spin density have a similar spatial distribution in all three molecules. These results suggest on the one hand that a local approach for g-tensor computations must be feasible and on the other hand that the unperturbed spin density provides a good estimate of the locality of the perturbed spin densities.

B. Comparison to other methods

In Table I, we show benchmark calculations for small molecules with the SOMF-GIAO method (without integral screening) using the B3LYP functional. As a reference, we use accurate values from a recent study of Perera *et al.*,[38] which were obtained at the coupled-cluster singles and doubles (CCSD) level with full one-electron and two-electron spin-orbit operators. Overall, the SOMF-GIAO method agrees well with the CCSD values with an average deviation of 233 ppm. Further improvements might be possible by also including explicit two-electron contributions in the diamagnetic contributions as in Ref. 38.

Table I also contains g-shifts computed with an effective one-electron operator with empirically determined nuclear charges from Koseki *et al.*[51] as proposed by Neese.[26] Apart from small deviations, which are presumably due to the use of GIAOs and different integration grids, we were able to reproduce the B3LYP results with effective nuclear charges from Ref. 35 using our implementation in FermiONs++.[67–69] The results indicate that the ansatz with effective nuclear charges is less accurate than the SOMF-GIAO method and leads to an average deviation of 360 ppm to the CCSD reference.

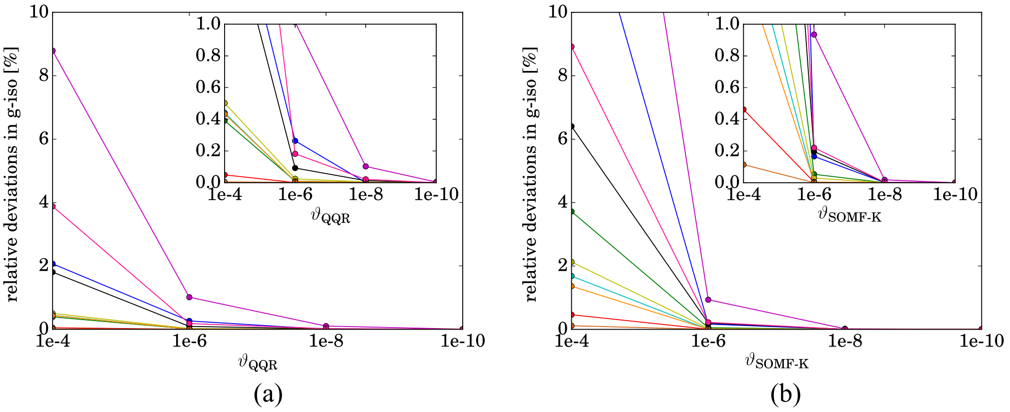
TABLE I. Comparison of the new SOMF-GIAO approach with the B3LYP functional to CCSD values from Ref. 38 and to values obtained with an effective one-electron spin-orbit operator with empirical nuclear charges ("Z_{eff}-GIAO"). All DFT calculations were carried out with FermiONs++[67–69] and the aug-cc-pVTZ basis set. The employed molecular geometries were taken from Ref. 35. The average deviations are computed with the CCSD values as reference. All values are given in ppm.

Molecule	Component	B3LYP		CCSD
		Z_{eff}-GIAO	SOMF-GIAO	
CN	Δg_{xx}	−2 173	−2 029	−2 033
	Δg_{zz}	−134	−134	−176
CO^+	Δg_{xx}	−2 620	−2 451	−2 436
	Δg_{zz}	−134	−134	−162
BO	Δg_{xx}	−1 840	−1 699	−1 734
	Δg_{zz}	−68	−68	−86
NH	Δg_{xx}	1 369	1 214	1 379
	Δg_{zz}	−106	−106	−163
OH^+	Δg_{xx}	3 722	3 389	3 979
	Δg_{zz}	−173	−173	−249
H_2O^+	Δg_{xx}	−189	−189	−250
	Δg_{yy}	13 654	12 300	15 156
	Δg_{zz}	4 701	4 152	4 353
CH_3	Δg_{xx}	−89	−90	−113
	Δg_{zz}	653	520	476
O_2	Δg_{xx}	2 686	2 498	2 516
	Δg_{zz}	−200	−200	−89
O_3^-	Δg_{xx}	−554	−526	−581
	Δg_{yy}	18 459	17 026	16 583
	Δg_{zz}	11 084	10 174	9 726
CO_2^-	Δg_{xx}	951	914	747
	Δg_{yy}	−5 142	−4 636	−4 674
	Δg_{zz}	−724	−677	−723
H_2CO^+	Δg_{xx}	5 927	5 458	5 668
	Δg_{yy}	86	80	233
	Δg_{zz}	249	212	608
NO_2	Δg_{xx}	3 643	3 444	3 327
	Δg_{yy}	−11 873	−10 851	−10 792
	Δg_{zz}	−697	−656	−630
NF_2	Δg_{xx}	−669	−619	−526
	Δg_{yy}	7 010	6 376	6 030
	Δg_{zz}	4 147	3 757	3 239
Average deviation		360	233	

C. Accuracy of the introduced approximations

The methods presented in Sec. II employ several thresholds and parameters, whose values determine accuracy and efficiency of the calculation. In this section, benchmark calculations are presented in which always one parameter was varied while all others were set to fixed values, which only introduce negligible errors. These values are zero for the projection approaches proj-P1, proj-P2, and proj-h (equivalent to performing no projection at all); 10^{-12} has been employed for all integral screening thresholds, and 15 Laplace points were used (it was confirmed by comparison to reference calculations without integral screening or projections that these settings introduce negligible errors, which were in all cases below 0.001%). In this manner, we analyze the influence of each parameter in isolation. In general, we consider relative errors on the order of 0.2% in the computed g-shifts as satisfying accuracy and try to determine threshold values that provide such high accuracy.

Our test set contains 10 medium-sized molecules; among them are several bio-molecules, such as short peptides [a cysteine radical bound to four glycines ("$\cdot$Cys-Gly$_4$"), a serine-lysine dipeptide ("Ser-Lys") with the radical center at the serine side-chain oxygen, and a tripeptide from the enzyme pyruvate formate lyase with the radical center at the residue 419 ("419," structure taken from Ref. 83)], a fatty acid radical (myristoleic acid), and all-trans-retinal. Furthermore, the test set contains a commonly employed spin label called MTSL,[84] an alcohol radical ("$\cdot$O-[CH$_2$]$_{18}$–CH$_3$"), and an alkane radical ("$\cdot$CH$_2$–[CH$_2$]$_{18}$–CH$_3$"). The two remaining molecules are a 4-octylphenol molecule with the radical center at the oxygen and triethylmethylamine with the radical center at the methyl group. The focus lies on molecules with a rather localized spin density; still, some systems have an at least partially delocalized spin density such as all-trans-retinal with its conjugated double-bonds or the octylphenol radical where the spin density is delocalized over the aromatic ring and the neighboring oxygen. All calculations were done with the def2-TZVP basis set.[85]



(a)

(b)

FIG. 2. Relative errors in the isotropic g-shift due to (a) the QQR screening for the Coulomb part of the SOMF operator from Sec. II C 2 and (b) the screening for the exchange part of the SOMF operator from Sec. II C 3 are shown as a function of the respective screening threshold. The insets provide a detailed view on relative errors in the range from 0.0% to 1.0%. For the color-coding of the molecules, see Fig. 5.

1. Integral screening

Figure 2 shows the errors introduced by the QQR screening in the Coulomb contribution to the SOMF operator and due to the screening for the contribution from the exchange part of the SOMF operator. For both contributions, thresholds of 10^{-8} are sufficient for achieving errors of less than 0.2%, even though the errors are slightly larger for the Coulomb contribution. Figure 3 displays the errors due to LinK screening in Eq. (23). Negligible errors below 10^{-3}% are introduced with a threshold of 10^{-6} for all molecules. A threshold of 10^{-4} is sufficient to reduce the errors below 3% in all cases. Finally, negligible errors are obtained, if integral screening with thresholds of 10^{-8} and 10^{-6} is used for $\mathbf{J}^{B_p}$ in Eq. (24) and $\mathbf{K}^{\sigma,B_p}$ in Eq. (25), respectively (see Fig. 3).

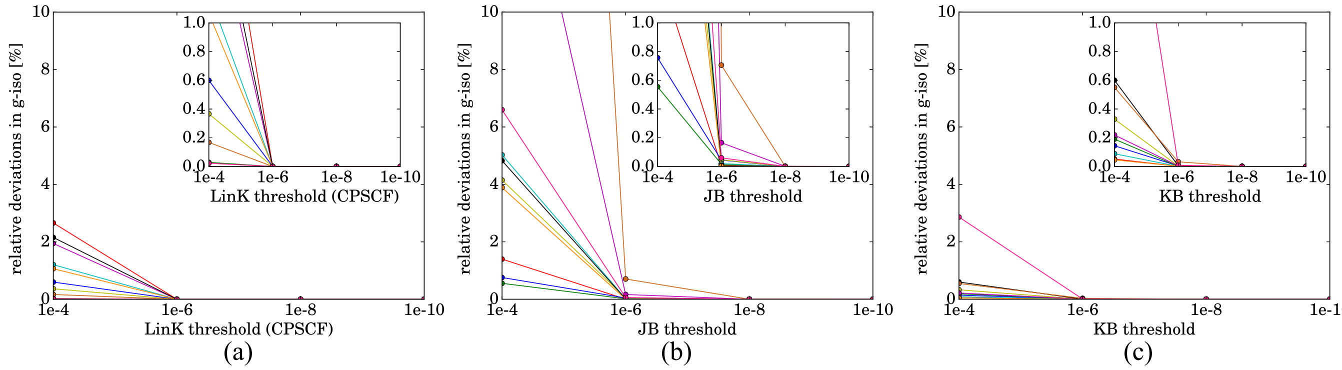
In summary, we observe fast convergence towards the reference values for all integral screening methods; the obtainable accuracy is fully controllable by adjusting the screening thresholds.

2. Laplace quadrature

Only a small number of Laplace points are needed for achieving negligible error due to the Laplace quadrature (see Fig. 4). For 7 quadrature points, errors below 0.1% are observed for all molecules; 10 quadrature points reduce the errors below 0.01% in all cases. As most steps of the g-tensor calculation are independent of the number of Laplace points [the exceptions are the matrix multiplications in Eq. (18) and the construction of the pseudo-densities, which are usually not time-dominating], a relatively high number of Laplace points can be used without significantly increasing the total compute time.

3. CPSCF with the "locally projected perturbation" approach

In Fig. 5, the relative errors in the isotropic g-shift caused by the subspace projection approaches described in Sec. II D are plotted against the corresponding thresholds. The results



(a)

(b)

(c)

FIG. 3. Relative errors in the isotropic g-shift as a function of the screening thresholds for the LinK screening in Eq. (23) (a), the Schwarz screening for the near-field part of $\mathbf{J}^{B_p}$ in Eq. (24) (b) and the LinK screening in the computation of $\mathbf{K}^{\sigma,B_p}$ in Eq. (25) (c). For the color-coding of the molecules, see Fig. 5.

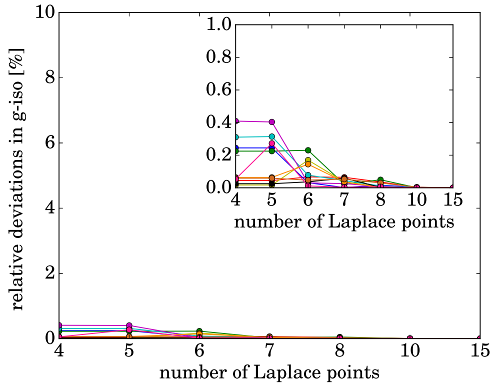


FIG. 4. Relative errors in the isotropic g-shift as a function of the number of Laplace integration points used in the DL-CPSCF. For the color-coding of the molecules, see Fig. 5.

indicate that all three projection approaches are robust and quickly converge to the result without projection for all analyzed systems. The approaches proj-P2 and proj-h are also used in Sec. IV D; suitable thresholds for obtaining errors of around 0.2% or less are 10^{-2} for $\vartheta_{\mathrm{h}}^{\mathrm{user}}$ and 10^{-1} for $\vartheta_{\mathrm{P2}}^{\mathrm{user}}$. In Fig. 5, it is also shown how many shell-pairs are removed by the projection approaches for the different thresholds. The fraction of shell-pairs that can be removed by the projection varies considerably among the molecules in the test set; this is expected due to the different degrees of locality of the spin densities. For molecules with local spin densities, large fractions of the AO space can be removed by the projection, while for molecules with less local spin densities, only small fractions can be removed.

When comparing the ratios between the errors and the sizes of the orbital spaces, the approaches proj-P1 and proj-P2 show similar performance. The ratios between the introduced errors and the sizes of the orbital spaces are significantly worse for proj-h; however, proj-h has a different purpose and is intended to allow for a sublinear computation of $\mathbf{h}^{\sigma,B_p}$, which has a smaller prefactor than the iterative part of the CPSCF.

D. Scaling behavior and efficiency

In this section, we present timings from calculations on two kinds of molecular systems: on the one hand, alkane radicals of varying length with a terminal spin center and, on the other hand, short DNA strands with 1-5 adenine-thymine (AT) base pairs and an attached nitroxide spin label. The alkanes are idealized systems with a high degree of sparsity in the density matrices and are well-suited to show the asymptotic scaling behavior for very large systems; the spin-labeled DNA strands with less sparse density matrices are more representative of molecules that might be considered in application studies.

First, we consider the SOMF-GIAO method introduced in Sec. II C and use integral screening thresholds that were found to allow for accurate computations in Sec. IV C: 10^{-6} for the LinK screening in the CPSCF, 10^{-8} for the exchange contribution to the SOMF operator, 10^{-10} for the QQR screening for the Coulomb part of the SOMF operator, 10^{-8} for the near-field part of $\mathbf{J}_p^{\mathrm{B}}$ [Eq. (24)], and 10^{-6} for the LinK screening used for $\mathbf{K}^{\sigma,B_p}$ [Eq. (25)].

The corresponding timings with the def2-SVP basis set[85] are shown in Fig. 6. For the alkanes, the observed scaling is clearly below quadratic and approaches linear scaling; for the two largest systems (alkane-160 and alkane-200), the scaling exponent amounts to 1.3 for the entire g-tensor calculation and 1.4 for the iterative part of the CPSCF, which is time-dominating. For the spin-labeled DNA, the effective scaling is higher [1.8 for the entire g-tensor calculation between $(\mathrm{AT})_4$ and $(\mathrm{AT})_5$], which is due to the more three-dimensional structure and lower sparsity in the density matrices. In order to observe a scaling closer to 1.0, one would need to consider larger DNA strands.

For both alkane radicals and spin-labeled DNA, the computation of $\mathbf{h}^{\sigma,B_p}$ requires a significant fraction of the total compute time for smaller systems, but displays a lower effective scaling than the iterative part and thus becomes less important for larger molecules. This is mainly due to the CFMM used for $\mathbf{J}^{\mathrm{B}}$ in Eq. (24) and the computation of $\frac{\mathrm{d}f^{\mathrm{xc},\sigma}}{\mathrm{d}B_p}$, both which scale close to linear. The compute time needed for the SOMF operator is negligible for both systems because the locality of spin density and perturbed spin densities can be efficiently exploited in the screening from Secs. II C 2 and II C 3.

The g-tensor calculations on both types of systems were comparably as fast as the linear scaling self-consistent field (SCF) method with LinK screening for the exchange matrix and CFMM for the Coulomb matrix (and integral screening thresholds of 10^{-10} for both; for the timings, see Fig. 8). The chosen SCF settings are tight, but this is advisable for practical calculations where g-tensors and potentially other molecular properties are computed based on the SCF density. For alkane-200, the SCF took 49 min and the g-tensor part 60 min. For the largest DNA system with 5 AT base pairs and the def2-SVP basis set, both the g-tensor calculation and the SCF took 212 min. This shows the high efficiency of the method. Also the scaling with system size is highly similar for the g-tensor calculation and the preceding SCF, for which the effective scaling between alkane-160 and alkane-200 amounts to 1.4, while it amounts to 1.7 between $(\mathrm{AT})_4$ and $(\mathrm{AT})_5$.

In Fig. 9, timings on $(\mathrm{AT})_1$ to $(\mathrm{AT})_3$ with the def2-TZVP basis set are shown. Due to CPSCF convergence problems in some of the calculations, we tightened the LinK threshold to 10^{-8}. Both computational cost and scaling with system size are higher than for the def2-SVP basis set. However, the effective scaling is decreasing and can be expected to further decrease for larger systems. Also for this basis set, scaling and wall time for the g-tensor calculation and the SCF are similar in magnitude.

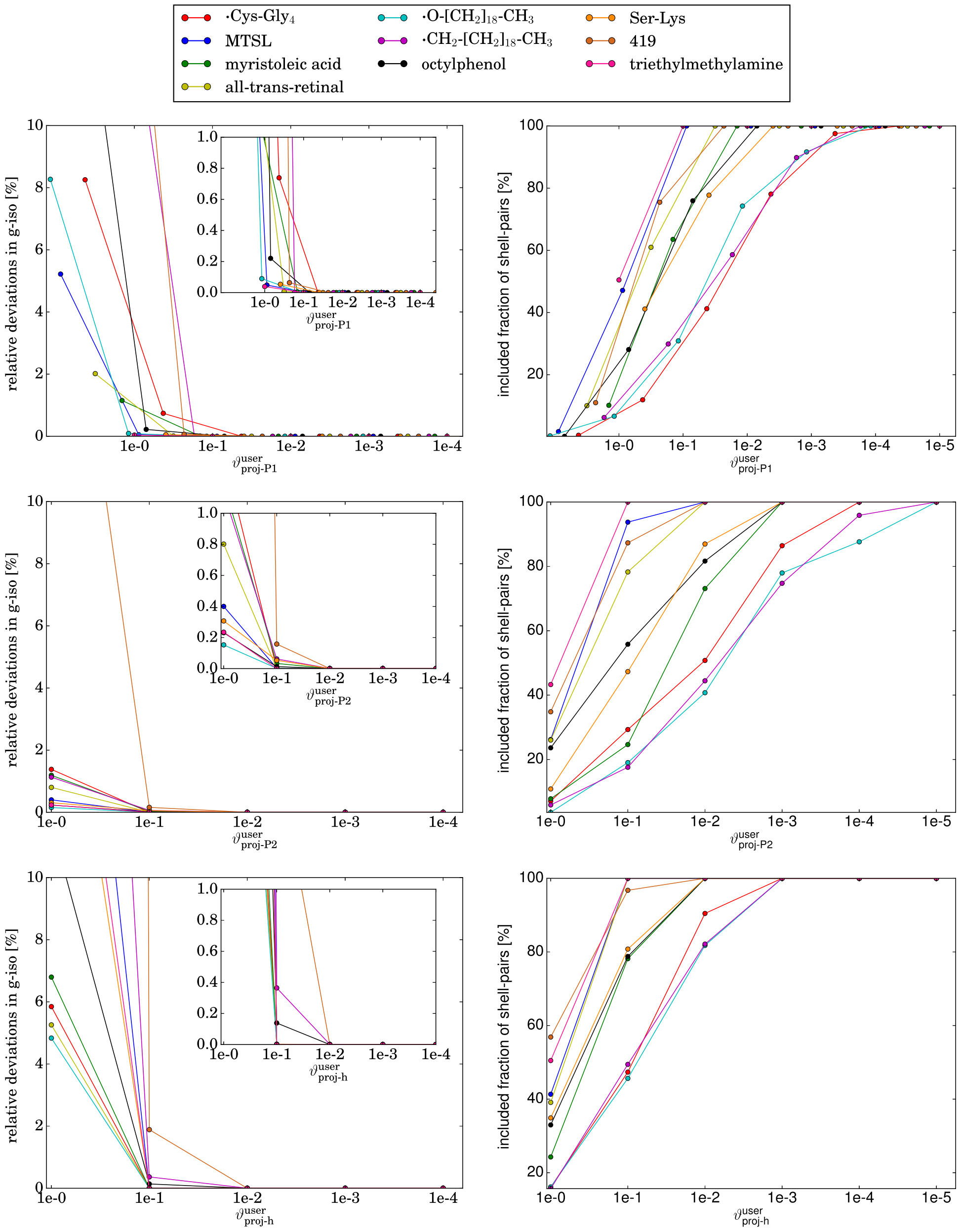
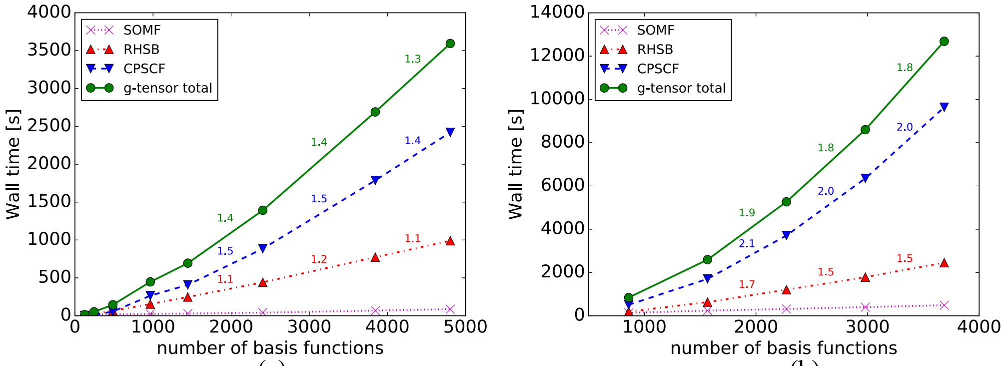


FIG. 5. Relative errors in the isotropic g-shifts due to the projection approaches proj-P1, proj-P2, and proj-h as a function of the corresponding projection threshold are shown in the left column. The right column shows the fractions of shell-pairs that need to be included in the spaces $\tilde{\Omega}_P$ and $\tilde{\Omega}_h$ as a function of the projection threshold.

In summary, the SOMF-GIAO method was shown to display near linear scaling for sufficiently large systems and a reasonably low prefactor and thus allows for calculations on large molecules.

In the calculations with the LPP-SOMF-GIAO method introduced in Sec. II D, we employed the projection approach P2h with thresholds of 10^{-1} for $\vartheta^{\text{user}}_{\text{P2}}$ and 10^{-2} for $\vartheta^{\text{user}}_{\text{h}}$; in addition, integral screening with the same thresholds as for the linear scaling ansatz was employed. In order to analyze the combined influence of the subspace projections proj-h and proj-P2 on the obtained accuracy, we show the relative deviations from the SOMF-GIAO method for all employed alkanes and spin-labeled DNA strands in Fig. 7. For both series of systems, the magnitude of the errors does not grow significantly if the molecule is enlarged. Small oscillations in the magnitude of the errors occur for the alkanes with a LinK screening threshold of 10^{-6}; these oscillations are due to numerical errors from the screening of integrals and disappear if the threshold is lowered to 10^{-8}. In Fig. 7, also the sizes of the orbital spaces $\tilde{\Omega}_h$ and $\tilde{\Omega}_P$ are shown. For both the alkanes and the spin-labeled DNA, $\tilde{\Omega}_h$ is significantly larger than $\tilde{\Omega}_P$. The

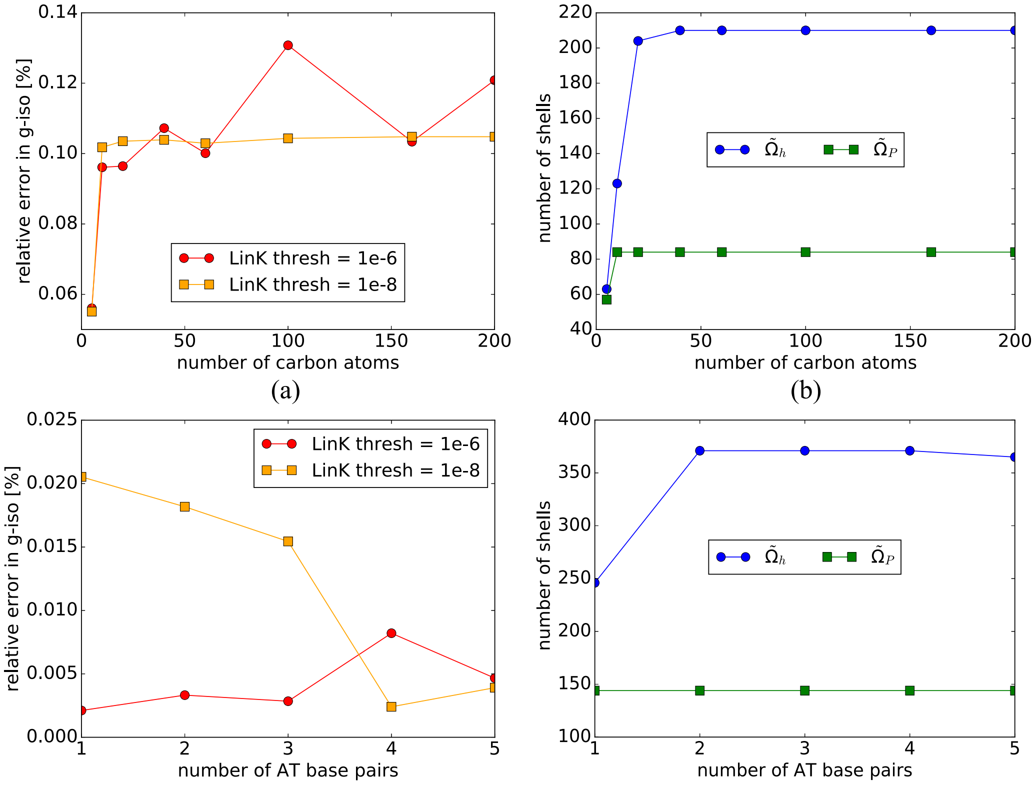


(a) (b)

FIG. 6. Timings for alkane radicals with 5 to 200 carbon atoms (a) and spin-labeled DNA strands with 1-5 AT base pairs (b). The def2-SVP basis set with the SOMF-GIAO method was employed in all calculations. "RHSB" denotes the time needed for the computation of all contributions to $\mathbf{h}^{\sigma,B_p}$ [see Eq. (21)]. "CPSCF" denotes the compute time needed for the iterative part of the CPSCF. The colored numbers indicate the effective scaling.

orbital spaces converge quickly with molecule size and stay essentially constant in size for systems larger than alkane-40 or spin-labeled (AT)$_2$.

In Fig. 8, timings with the LPP-SOMF-GIAO method and def2-SVP basis set are shown in comparison with the timings for the SOMF-GIAO approach. In most cases, the observed



(a) (b) (c) (d)

FIG. 7. Calculations with the LPP-SOMF-GIAO method and the def2-SVP basis set. [(a) and (b)] Calculations on alkane radicals with 5–200 carbon atoms. [(c) and (d)] calculations on spin-labeled DNA strands with 1-5 AT base pairs. [(a) and (c)] Relative errors in the isotropic g-shifts compared to the SOMF-GIAO method. Two different screening thresholds for the LinK screening have been used (10^{-6} and 10^{-8}). [(b) and (d)] Number of shells of basis functions that need to be included in the spaces $\tilde{\Omega}_P$ and $\tilde{\Omega}_h$.

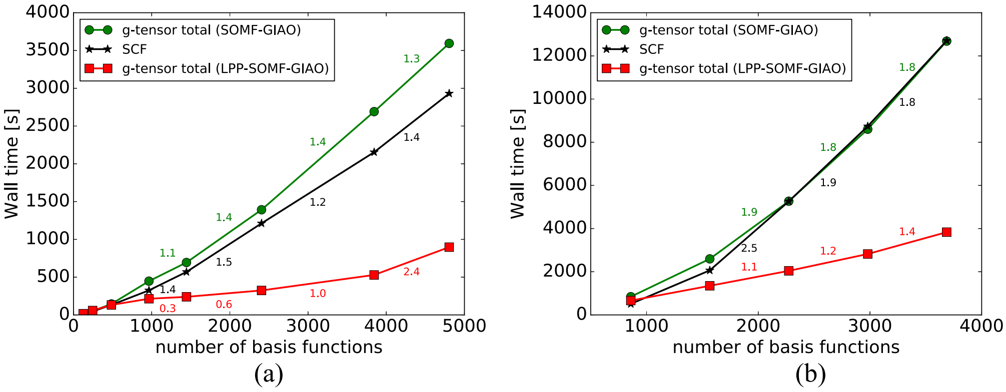


FIG. 8. Timings on alkane radicals with 5–200 carbon atoms (a) and spin-labeled DNA strands with 1-5 AT base pairs (b). The SOMF-GIAO and LPP-SOMF-GIAO methods are compared. In addition, the wall time needed for the SCF is shown. The def2-SVP basis set was used in all calculations. The colored numbers indicate the effective scaling.

effective scaling is significantly lower than the scaling of the SOMF-GIAO method. Between alkane-40 and alkane-100, effective sublinear scaling is observed with a minimal scaling of 0.3 between alkane-40 and alkane-60. Between $(AT)_2$ and $(AT)_3$, an effective scaling of 1.1 was obtained. For the largest considered systems, the scaling increases again. Several factors may contribute to this such as overhead due to the integral screening (which could be reduced by resorting to a more coarse-grained screening, i.e., screening blocks of shell-pairs instead of individual shell-pairs) or less efficient cache utilization. For the largest alkanes, the increase in scaling is also partly due to a higher number of required CPSCF iterations (7 and 10 iterations for alkane-160 and alkane-200 compared to 6 iterations for alkane-40 to alkane-100).

Irrespective of the discussed complications in the analysis of the scaling behavior, the LPP-SOMF-GIAO method is computationally highly efficient and leads to considerable speedups compared to the SOMF-GIAO method while only introducing negligible errors. The speedup amounts to 4.0 for alkane-200 (15 min compared to 60 min) and 3.3 for $(AT)_5$ with the def2-SVP basis set (64 min compared to 211 min). For $(AT)_3$ in a def2-TZVP basis, a speedup of 2.1 is observed (1005 min compared to 2131 min, see Fig. 9). For these systems, the g-tensor calculations were significantly faster than the SCF calculations with the employed settings (which took 49 min for alkane-200, 212 min for $(AT)_5$ in a def2-SVP basis, and 1973 min for $(AT)_3$ in a def2-TZVP basis). As the method requires an SCF reference state, the SCF itself becomes the bottleneck for large molecules with local spin density. Any further substantial speed-up would thus require a more efficient SCF.

While all calculations shown in this work were carried with our ansatz employing the SOMF operator and GIAOs, the presented strategies for reducing the scaling are without modification also applicable to other DFT-based methods for g-tensors, which differ, e.g., in the treatment of spin-orbit effects. In fact, some of these methods are implemented in our program code, including the methods from Neese using either

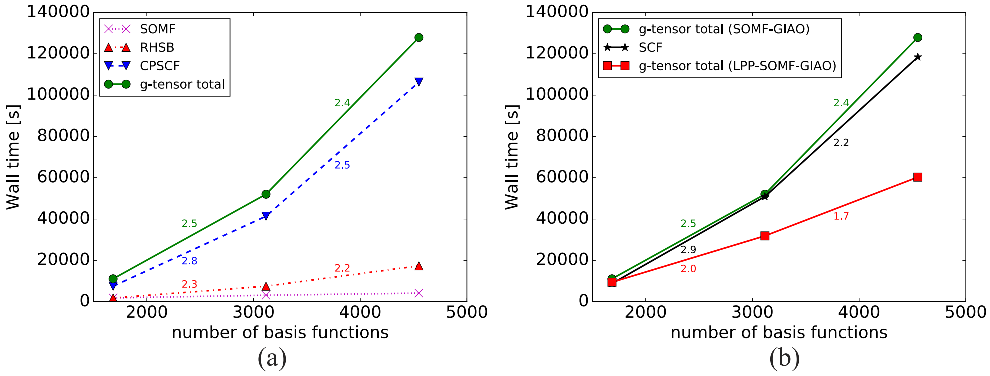


FIG. 9. Timings for spin-labeled DNA strands with 1-3 AT base pairs. The def2-TZVP basis set was employed in all calculations. A LinK screening threshold of 10^{-8} was used in the CPSCF. (a) shows timings with the SOMF-GIAO method and the individual contributions. For an explanation of the legend entries, see Fig. 6. (b) presents wall times and scaling of the SOMF-GIAO method, the LPP-SOMF-GIAO method, and the SCF. The colored numbers indicate the effective scaling.

the SOMF operator[34] or effective nuclear charges.[26] Our program also features a method using both GIAOs and effective nuclear charges, which has been used in the literature.[86–88] No computational differences between these approaches occur in the usually time-dominating iterative part of the CPSCF; therefore, one can expect that the application of the outlined strategies would lead to very similar scaling and efficiency for all these methods.

V. CONCLUSION

An efficient and accurate method for electronic g-tensor calculations at the DFT level has been presented. It employs the spin-orbit mean-field operator for an accurate computation of both one- and two-electron spin-orbit coupling contributions and makes use of gauge-including atomic orbitals for ensuring gauge-origin independence.

Asymptotically linear scaling with molecule size is achieved by integral screening methods and sparse linear algebra. In addition, a local perturbation approach for solving the CPSCF equations was introduced, which allows the scaling to be further reduced to sublinear for molecules with localized spin density. These methods make g-tensor computations on large molecular systems feasible and allow us to compute the g-tensor of an alkane-200 radical with 4805 basis functions in 15 min on a single node. For such molecules, the g-tensor calculation with the local perturbation approach can even be significantly faster than the SCF.

ACKNOWLEDGMENTS

The authors acknowledge financial support by the “Deutsche Forschungsgemeinschaft” (DFG) funding (Proposal Nos. Oc35/4-1 and SFB749) and by the Cluster of Excellence EXC 114 “Center for Integrated Protein Science Munich” (CIPSM). Furthermore, the authors thank Daniel Graf and Travis H. Thompson (LMU Munich) for helpful discussions. S.V. thanks the Studienstiftung des Deutschen Volkes for a graduate fellowship. C.O. acknowledges additional financial support as a Max-Planck-Fellow at MPI-FKF Stuttgart.

REFERENCES

[1] B. Odom, D. Hanneke, B. d’Urso, and G. Gabrielse, Phys. Rev. Lett. **97**, 030801 (2006).

[2] D. Hanneke, S. F. Hoogerheide, and G. Gabrielse, Phys. Rev. A **83**, 052122 (2011).

[3] J. Schwinger, Phys. Rev. **73**, 416 (1948).

[4] S. Laporta and E. Remiddi, Phys. Lett. B **379**, 283 (1996).

[5] T. Aoyama, M. Hayakawa, T. Kinoshita, and M. Nio, Phys. Rev. Lett. **109**, 111807 (2012).

[6] T. Aoyama, M. Hayakawa, T. Kinoshita, and M. Nio, Phys. Rev. D **91**, 033006 (2015).

[7] J. E. Harriman, *Theoretical Foundations of Electron Spin Resonance: Physical Chemistry: A Series of Monographs* (Academic Press, 2013).

[8] S. Frantz, H. Hartmann, N. Doslik, M. Wanner, W. Kaim, H.-J. Kümmerer, G. Denninger, A.-L. Barra, C. Duboc-Toia, J. Fiedler *et al.*, J. Am. Chem. Soc. **124**, 10563 (2002).

[9] J. R. Asher, N. L. Doltsinis, and M. Kaupp, J. Am. Chem. Soc. **126**, 9854 (2004).

[10] L. Benisvy, R. Bittl, E. Bothe, C. D. Garner, J. McMaster, S. Ross, C. Teutloff, and F. Neese, Angew. Chem., Int. Ed. **44**, 5314 (2005).

[11] T. Yamaji, I. S. Saiful, M. Baba, S. Yamauchi, and J. Yamauchi, J. Phys. Chem. A **111**, 4612 (2007).

[12] J. Niklas, M. Westwood, K. L. Mardis, T. L. Brown, A. M. Pitts-McCoy, M. D. Hopkins, and O. G. Poluektov, Inorg. Chem. **54**, 6226 (2015).

[13] M. R. Fuchs, E. Schleicher, A. Schnegg, C. W. Kay, J. T. Törring, R. Bittl, A. Bacher, G. Richter, K. Möbius, and S. Weber, J. Phys. Chem. B **106**, 8885 (2002).

[14] R. Owenius, M. Engström, M. Lindgren, and M. Huber, J. Phys. Chem. A **105**, 10967 (2001).

[15] M. Witwicki, A. R. Jaszewski, J. Jezierska, M. Jerzykiewicz, and A. Jezierski, Chem. Phys. Lett. **462**, 300 (2008).

[16] M. Fehr, A. Schnegg, B. Rech, K. Lips, O. Astakhov, F. Finger, G. Pfanner, C. Freysoldt, J. Neugebauer, R. Bittl, and C. Teutloff, Phys. Rev. B **84**, 245203 (2011).

[17] C. Bährle, T. U. Nick, M. Bennati, G. Jeschke, and F. Vogel, J. Phys. Chem. A **119**, 6475 (2015).

[18] C. Gionco, M. C. Paganini, E. Giamello, R. Burgess, C. Di Valentin, and G. Pacchioni, Chem. Mater. **25**, 2243 (2013).

[19] J. Niklas, K. L. Mardis, R. R. Rakhimov, K. L. Mulfort, D. M. Tiede, and O. G. Poluektov, J. Phys. Chem. B **116**, 2943 (2012).

[20] W. Sinha, M. G. Sommer, N. Deibel, F. Ehret, M. Bauer, B. Sarkar, and S. Kar, Angew. Chem., Int. Ed. **54**, 13769 (2015).

[21] 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).

[22] H. Malissa, R. Miller, D. Baird, S. Jamali, G. Joshi, M. Bursch, S. Grimme, J. van Tol, J. Lupton, and C. Boehme, Phys. Rev. B **97**, 161201 (2018).

[23] J. Niklas, K. L. Mardis, and O. G. Poluektov, J. Phys. Chem. Lett. **9**, 3915 (2018).

[24] W. Moores and R. McWeeny, Proc. R. Soc. London, Ser. A **332**, 365 (1973).

[25] R. Angstl, Chem. Phys. **132**, 435 (1989).

[26] F. Neese, J. Chem. Phys. **115**, 11080 (2001).

[27] G. Schreckenbach and T. Ziegler, J. Phys. Chem. A **101**, 3388 (1997).

[28] F. London, J. Phys. Radium **8**, 397 (1937).

[29] R. Ditchfield, J. Chem. Phys. **65**, 3123 (1976).

[30] K. Wolinski, J. F. Hinton, and P. Pulay, J. Am. Chem. Soc. **112**, 8251 (1990).

[31] B. A. Heß, C. M. Marian, U. Wahlgren, and O. Gropen, Chem. Phys. Lett. **251**, 365 (1996).

[32] B. Schimmelpfennig, *AMFI–An Atomic Mean Field Integral Program* (Stockholm University, 1996).

[33] B. Schimmelpfennig, L. Maron, U. Wahlgren, C. Teichteil, H. Fagerli, and O. Gropen, Chem. Phys. Lett. **286**, 267 (1998).

[34] F. Neese, J. Chem. Phys. **122**, 034107 (2005).

[35] J. Gauss, M. Kállay, and F. Neese, J. Phys. Chem. A **113**, 11541 (2009).

[36] F. Neese, Int. J. Quantum Chem. **83**, 104 (2001).

[37] F. Neese, Chem. Phys. Lett. **380**, 721 (2003).

[38] A. Perera, J. Gauss, P. Verma, and J. A. Morales, J. Chem. Phys. **146**, 164104 (2017).

[39] O. Vahtras, B. Minaev, and H. Ågren, Chem. Phys. Lett. **281**, 186 (1997).

[40] G. Lushington and F. Grein, J. Chem. Phys. **106**, 3292 (1997).

[41] G. H. Lushington, J. Phys. Chem. A **104**, 2969 (2000).

[42] S. Brownridge, F. Grein, J. Tatchen, M. Kleinschmidt, and C. M. Marian, J. Chem. Phys. **118**, 9552 (2003).

[43] F. Neese, Mol. Phys. **105**, 2507 (2007).

[44] E. R. Sayfutyarova and G. K.-L. Chan, J. Chem. Phys. **148**, 184103 (2018).

[45] E. van Lenthe, P. E. Wormer, and A. van der Avoird, J. Chem. Phys. **107**, 2488 (1997).

[46] K. M. Neyman, D. I. Ganyushin, A. V. Matveev, and V. A. Nasluzov, J. Phys. Chem. A **106**, 5022 (2002).

[47] P. Manninen, J. Vaara, and K. Ruud, J. Chem. Phys. **121**, 1258 (2004).

[48] I. Malkin, O. L. Malkina, V. G. Malkin, and M. Kaupp, J. Chem. Phys. **123**, 244103 (2005).

[49] M. Repiskỳ, S. Komorovskỳ, E. Malkin, O. L. Malkina, and V. G. Malkin, Chem. Phys. Lett. **488**, 94 (2010).

[50] M. Glasbrenner, S. Vogler, and C. Ochsenfeld, J. Chem. Phys. **148**, 214101 (2018).

[51] S. Koseki, M. W. Schmidt, and M. S. Gordon, J. Phys. Chem. **96**, 10768 (1992).

[52] H. F. King and T. R. Furlani, J. Comput. Chem. **9**, 771 (1988).

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

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

[55] P. Y. Ayala and G. E. Scuseria, J. Chem. Phys. **110**, 3660 (1999).

[56] C. Ochsenfeld, C. A. White, and M. Head-Gordon, J. Chem. Phys. **109**, 1663 (1998).

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

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

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

[60] M. Beer, J. Kussmann, and C. Ochsenfeld, J. Chem. Phys. **134**, 074102 (2011).

[61] C. A. White, B. G. Johnson, P. M. Gill, and M. Head-Gordon, Chem. Phys. Lett. **230**, 8 (1994).

[62] C. A. White, B. G. Johnson, P. M. Gill, and M. Head-Gordon, Chem. Phys. Lett. **253**, 268 (1996).

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

[64] M. Häser, Theor. Chim. Acta **87**, 147 (1993).

[65] S. A. Maurer, D. S. Lambrecht, D. Flaig, and C. Ochsenfeld, J. Chem. Phys. **136**, 144107 (2012).

[66] A. Luenser, J. Kussmann, and C. Ochsenfeld, J. Chem. Phys. **145**, 124103 (2016).

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

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

[69] J. Kussmann and C. Ochsenfeld, J. Chem. Theory Comput. **13**, 3153 (2017).

[70] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[71] P. Pulay, J. Comput. Chem. **3**, 556 (1982).

[72] K. Kim and K. Jordan, J. Phys. Chem. **98**, 10089 (1994).

[73] V. I. Lebedev, USSR Comput. Math. Math. Phys. **16**, 10 (1976).

[74] O. Treutler and R. Ahlrichs, J. Chem. Phys. **102**, 346 (1995).

[75] M. Kaltak, J. Klimeš, and G. Kresse, J. Chem. Theory Comput. **10**, 2498 (2014).

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

[77] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, J. Chem. Phys. **132**, 154104 (2010).

[78] H. Kruse and S. Grimme, J. Chem. Phys. **136**, 154101 (2012).

[79] R. Sure and S. Grimme, J. Comput. Chem. **34**, 1672 (2013).

[80] J. Kästner, J. M. Carr, T. W. Keal, W. Thiel, A. Wander, and P. Sherwood, J. Phys. Chem. A **113**, 11856 (2009).

[81] W. Humphrey, A. Dalke, and K. Schulten, J. Mol. Graphics **14**, 33 (1996).

[82] J. Stone, "An efficient library for parallel ray tracing and animation," M.S. thesis, Computer Science Department, University of Missouri-Rolla, 1998.

[83] S. Vogler, G. Savasci, M. Ludwig, and C. Ochsenfeld, J. Chem. Theory Comput. **14**, 3014 (2018).

[84] G. E. Fanucci and D. S. Cafiso, Curr. Opin. Struct. Biol. **16**, 644 (2006).

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

[86] E. Pauwels, J. Asher, M. Kaupp, and M. Waroquier, Phys. Chem. Chem. Phys. **13**, 18638 (2011).

[87] D. M. Ramo, A. Shluger, J. Gavartin, and G. Bersuker, Phys. Rev. Lett. **99**, 155504 (2007).

[88] B. Brogioni, D. Biglino, A. Sinicropi, E. J. Reijerse, P. Giardina, G. Sannia, W. Lubitz, R. Basosi, and R. Pogni, Phys. Chem. Chem. Phys. **10**, 7284 (2008).

Chapter 4

Conclusion and Outlook

In this thesis, several new quantum chemical methods for efficiently and accurately computing correlation energies, NMR shieldings, and electronic g-tensors were described.

In **Publication V**, the gauge-origin dependence of the electronic g-tensor is analyzed in an extensive benchmark study using a DFT approach. It is shown that simple, gauge-origin dependent methods using a single gauge-origin can give large errors for some molecules. These errors can be eliminated completely by employing distributed gauge-origin methods like GIAOs,[157–160] which are standard for NMR shielding calculations. In contrast, most implementations of g-tensors previously described in the literature employ a common gauge-origin.[59–76] The most used gauge-origin in the literature is the electronic charge centroid (ECC),[163] which corresponds to the center of mass of the electron cloud. As shown in **Publication V**, this choice of gauge-origin can lead to substantial errors especially for large molecules with spin centers far away from the ECC. An alternative choice of gauge-origin called the spin density center (SDC) is proposed, which is by construction located near the center of mass of the unsigned spin density. We showed that for molecules with a single well-localized spin center, much more accurate g-tensors are obtained with the SDC than with the ECC. The data from **Publication V** furthermore suggests that also the SDC approach is insufficient for computing accurate g-tensors if the spin density is delocalized or if multiple spin centers are present. In such situations the use of a distributed gauge-origin approach like GIAOs is mandatory for obtaining reliable results. The findings of **Publication V** might have a significant impact on the way the gauge-origin problem is tackled in future theoretical methods for computing g-tensors and application studies.

The first linear- and sublinear-scaling methods for computing g-tensors with unrestricted DFT or HF are presented in **Publication VI**. The implementation allows to treat the spin-orbit coupling effects either with an effective one-electron operator and effective nuclear charges,[187] or with the more accurate spin-orbit mean field (SOMF)[135] approach. GIAOs[157–160] are employed for ensuring gauge-origin independence, whose importance was shown in **Publication V**. Asymptotically linear scaling is obtained by employing integral screening and sparse linear algebra approaches. It is shown that for molecules with a well-localized spin density distribution, asymptotic $\mathcal{O}(1)$ scaling is possible. This is enabled by exploiting the locality of the spin density matrix in the computation of the spin-orbit contributions and by using a modified CPSCF approach, which allows to include only contributions from parts of the molecule close to the spin center in the computation of the perturbed densities. Even higher performance could likely be obtained by using RI for the Coulomb part of the SOMF matrix and a semi-numerical approach[188] for the exchange. The outlined strategies for efficiently computing the SOMF operator and for exploiting the locality of the spin density are likely also transferable to other methods for computing g-tensors at the MP2 or coupled cluster level of theory.

In **Publication I**, the efficiency of the RI-CDD-MP2 method for correlation energies from Maurer *et al.*[146] is improved further by introducing an erfc-attenuated Coulomb metric for the RI approximation. In contrast to the method of Maurer *et al.*,[146] where distance-dependent QQR screening[44] is used and Coulomb and exchange energy contributions are computed together, a separate treatment of these contributions using specialized algorithms is proposed in order to exploit the benefits of the local RI metric. The Coulomb contributions and the transformations of the three-center integrals are accelerated using a modified form of the natural blocking approach, which was originally introduced by Jung *et al.*[47,127] The fast exponential decay of the exchange contributions is exploited using a Schwarz screening[50] approach. Timings on alkanes and DNA strands show that the new method called ω-RI-CDD-MP2 is significantly more efficient than the original RI-CDD-MP2. The additional sparsity in the three-center integrals provided by the local RI metric allows to also lower memory and disk space requirements compared to RI-CDD-MP2 and achieve asymptotically linear scaling for these quantities. With the new MP2 method from **Publication I**, larger systems than previously possible can thus be computed at the MP2 level of theory. The largest tackled system was a DNA strand with 16 adenine-thymine base pairs and 11230 basis functions. The applicability to very large systems may be further improved by an adaptation to graphics processing units (GPUs), or parallelization over multiple nodes. Furthermore, implementing various analytical derivatives of the new method could allow to efficiently compute properties of large molecules with MP2 or performing geometry optimizations and molecular dynamics simulations at this level of theory. So far, only NMR shieldings were implemented using a mixed second derivative as described in **Publication III**.

A different path to efficient MP2 energies is explored in **Publication II** using the tensor hypercontraction (THC)[51,52] approach from Martínez and co-workers. In contrast to the RI approximation, THC allows to decompose the fourth order ERI tensor into two-index quantities and can reduce the formal scaling of many correlation methods. Both obtaining the THC factorization of the ERIs and computing the MP2 energy from the THC factors formally scales as $\mathcal{O}(N^4)$.[51] In **Publication II**, we describe how the THC decomposition can be done efficiently for large molecules with reduced scaling. For this, local MOs obtained from a Cholesky decomposition of the occupied and virtual density matrix are used in combination with an attenuated Coulomb metric and natural blocking. Furthermore, the explicit inversion of the grid-overlap metric, which was reported to be a computational bottleneck,[52,137,138] is circumvented by solving a linear equation system. It is also shown how the Coulomb and exchange contributions to the MP2 energy can be computed efficiently using THC and Cholesky-decomposed pseudo-densities. While the performance of the method is similar to the ω-RI-CDD-MP2 from **Publication III** at the cost of a small additional error due to the THC factorization, especially the very low memory requirements make the method attractive. It is also shown in **Publication II** that the computation of the Coulomb contributions to the MP2 energy is extremely fast. Significant speed-ups compared to ω-RI-CDD-MP2 are only prevented by the computational cost of obtaining the THC factorization of the ERIs. As the computational cost for decomposing the ERIs does not depend on the specific method used, one can expect that a very good performance would be obtained if the developed THC approach would be applied to other methods that are similar to SOS-MP2, but more computationally expensive, like, e.g., direct RPA. For the same reasons, also an excellent computational efficiency could be expected for SOS-MP2 properties, which require only unperturbed ERIs and no ERI derivatives. Some examples are hyperfine coupling constants (HFCCs), NMR spin-spin coupling constants, electric dipole moments, polarizabilities, and hyper-polarizabilities. Work on computation of HFCCs at the MP2 level of theory with the THC approach is currently in progress. For other molecular properties, which do require to compute perturbed ERIs, additional work on the efficient THC decomposition of perturbed ERIs is needed.

An efficient method for computing NMR shieldings at the MP2 level of theory is presented in **Publication III**. It is related to the method from Maurer *et al.*[53] for computing NMR shieldings with AO-MP2, but additionally uses Cholesky-decomposed pseudo-densities and the RI approximation for the two-electron integrals in order to lower the computational cost. A nested Z-vector approach is used for minimizing the number of CPSCF equations that need to be solved for obtaining perturbed density matrices. In contrast to the *selected-nuclei* approach[180] used by Maurer *et al.*,[53] an *all-nuclei* approach is employed, which allows to more efficiently compute the entire set of NMR shieldings for all atoms in a given molecule. This is achieved by formulating the nested Z-vector approach[54,55] in such a way that the explicit computation of the nuclear spin derivative of the density matrices and pseudo-density matrices is circumvented. Calculations on glycine chains show that quadratic scaling of the computation time and linear scaling of the required disk space can be achieved for large molecules. Also computations on DNA strands with up to 194 atoms and 2039 basis functions were shown in **Publication III**, which illustrates the potential of the method for applications on large biomolecules. A further speed-up and reduction of the memory requirements may be possible by combining the method with THC and applying techniques from **Publication II**. For this, the THC decomposition approach needs to be extended to magnetic field perturbed ERIs. Also, a suitable matrix decomposition of the perturbed pseudo-densities, which preserves locality and at the same time exploits rank-deficiency, could improve the efficiency. In contrast to the unperturbed pseudo-densities, Cholesky decomposition is not applicable, because the perturbed pseudo-densities are not positive semi-definite. Furthermore, the method could be combined with a QM/MM approach[189] in order to enable calculations on even larger molecules.

The first method for computing NMR shieldings at the post-Kohn–Sham RPA level of theory is introduced in **Publication IV**. In the presented implementation, RPA NMR shieldings are computed as a non-trivial numerical derivative of the RPA energy using a finite-difference scheme. Benchmark calculations on a test set of Teale *et al.*[190] are performed and compared to high-quality coupled cluster reference values. The results show that the accuracy of RPA shieldings strongly depends on the employed orbitals and orbital energies. A clear correlation between the amount of exact HF exchange in the reference functional and the accuracy is found. The most accurate shieldings are obtained with HF orbitals. An analysis of the basis set convergence shows that at least triple-zeta basis sets are needed for computing accurate RPA NMR shieldings. Comparison with other methods shows that RPA NMR shieldings with a HF reference are significantly more accurate than MP2 shieldings and are close to CCSD quality. The good observed accuracy makes RPA a promising method for NMR shielding calculations on large molecules considering also its comparatively low $\mathcal{O}(N^4)$ scaling. The developments from **Publication IV** only present a first step towards RPA shielding calculations for large molecules. The development of efficient analytical derivative methods for RPA NMR shieldings involving also a Z-vector approach is highly desirable. Due to the similarities of direct RPA and SOS-MP2, many of the developments for (SOS)-MP2 shieldings from **Publication III** are likely transferable to analytical RPA NMR shieldings.

Bibliography

[1] L. Pauling, *The Nature of the Chemical Bond and the Structure of Molecules and Crystals: An Introduction to Modern Struct. Chem..* (Oxford University Press, London, 1940).

[2] W. Ritz, J. Reine Angew. Math. **1909**, 1 (1909).

[3] H. D. R., Proc. Cambridge Philos. Soc. **24**, 89 (1928).

[4] J. C. Slater, Phys. Rev. **35**, 210 (1930).

[5] V. Fock, Z. Phys. **61**, 126 (1930).

[6] P.-O. Löwdin, Phys. Rev. **97**, 1509 (1955).

[7] B. O. Roos, P. R. Taylor, and P. E. M. Siegbahn, Chem. Phys. **48**, 157 (1980).

[8] S. R. White, Phys. Rev. Lett. **69**, 2863 (1992).

[9] S. R. White, Phys. Rev. B **48**, 10345 (1993).

[10] S. R. White and R. L. Martin, J. Chem. Phys. **110**, 4127 (1999).

[11] P.-O. Löwdin, Phys. Rev. **97**, 1474 (1955).

[12] C. D. Sherrill and H. F. Schaefer III, Adv. Quantum Chem. **34**, 143 (1999).

[13] T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular electronic-structure theory* (John Wiley & Sons, New York, 2014).

[14] J. Čížek, J. Chem. Phys. **45**, 4256 (1966).

[15] J. Čížek, J. Paldus, and L. Šroubková, Int. J. Quantum Chem. **3**, 149 (1969).

[16] R. J. Bartlett and J. Noga, Chem. Phys. Lett. **150**, 29 (1988).

[17] E. Schrödinger, Ann. Phys. **385**, 437 (1926).

[18] C. Møller and M. S. Plesset, Phys. Rev. **46**, 618 (1934).

[19] G. D. Purvis III and R. J. Bartlett, J. Chem. Phys. **76**, 1910 (1982).

[20] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).

[21] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[22] D. C. Langreth and J. P. Perdew, Solid State Commun. **17**, 1425 (1975).

[23] D. C. Langreth and J. P. Perdew, Phys. Rev. B **15**, 2884 (1977).

[24] J. G. Ángyán, R.-F. Liu, J. Toulouse, and G. Jansen, J. Chem. Theory Comput. **7**, 3116 (2011).

[25] F. Furche, Phys. Rev. B **64**, 195120 (2001).

[26] F. Furche, J. Chem. Phys. **129**, 114105 (2008).

[27] D. Bohm and D. Pines, Phys. Rev. **92**, 609 (1953).

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

[29] W. Kohn, Phys. Rev. Lett. **76**, 3168 (1996).

[30] E. Prodan and W. Kohn, Proc. Natl. Acad. Sci. USA **102**, 11635 (2005).

[31] C. Ochsenfeld, J. Kussmann, and D. S. Lambrecht, Rev. Comput. Chem. **23**, 1 (2007).

[32] P. Y. Ayala and G. E. Scuseria, J. Chem. Phys. **110**, 3660 (1999).

[33] D. S. Lambrecht, B. Doser, and C. Ochsenfeld, J. Chem. Phys. **123**, 184102 (2005).

[34] D. S. Lambrecht, B. Doser, and C. Ochsenfeld, J. Chem. Phys. **136**, 149902 (2012).

[35] B. Doser, D. S. Lambrecht, and C. Ochsenfeld, Phys. Chem. Chem. Phys. **10**, 3335 (2008).

[36] B. Doser, D. S. Lambrecht, J. Kussmann, and C. Ochsenfeld, J. Chem. Phys. **130**, 064107 (2009).

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

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

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

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

[41] J. Zienau, L. Clin, B. Doser, and C. Ochsenfeld, J. Chem. Phys. **130**, 204112 (2009).

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

[43] S. A. Maurer, D. S. Lambrecht, J. Kussmann, and C. Ochsenfeld, J. Chem. Phys. **138**, 014101 (2013).

[44] S. A. Maurer, D. S. Lambrecht, D. Flaig, and C. Ochsenfeld, J. Chem. Phys. **136**, 144107 (2012).

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

[46] Y. Jung, A. Sodt, P. M. Gill, and M. Head-Gordon, Proc. Natl. Acad. Sci. USA **102**, 6692 (2005).

[47] Y. Jung, Y. Shao, and M. Head-Gordon, J. Comput. Chem. **28**, 1953 (2007).

[48] S. Reine, E. Tellgren, A. Krapp, T. Kjærgaard, T. Helgaker, B. Jansik, S. Høst, and P. Salek, J. Chem. Phys. **129**, 104101 (2008).

[49] A. Luenser, H. F. Schurkus, and C. Ochsenfeld, J. Chem. Theory Comput. **13**, 1647 (2017).

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

[51] E. G. Hohenstein, R. M. Parrish, and T. J. Martínez, J. Chem. Phys. **137**, 044103 (2012).

[52] R. M. Parrish, E. G. Hohenstein, T. J. Martínez, and C. D. Sherrill, J. Chem. Phys. **137**, 224106 (2012).

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

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

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

[56] J. Gerratt and I. M. Mills, J. Chem. Phys. **49**, 1719 (1968).

[57] E. Trushin, A. Thierbach, and A. Görling, J. Chem. Phys. **154**, 014104 (2021).

[58] 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) pp. 541–592.

[59] I. Ciofini, R. Reviakine, A. Arbuznikov, and M. Kaupp, Theor. Chem. Acc. **111**, 132 (2004).

[60] M. Witwicki, A. R. Jaszewski, J. Jezierska, M. Jerzykiewicz, and A. Jezierski, Chem. Phys. Lett. **462**, 300 (2008).

[61] B. Frecus, Z. Rinkevicius, and H. Ågren, Phys. Chem. Chem. Phys. **15**, 10466 (2013).

[62] M. Witwicki and J. Jezierska, J. Phys. Chem. B **115**, 3172 (2011).

[63] 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).

[64] S. M. Mattar, J. Phys. Chem. B **108**, 9449 (2004).

[65] M. Engström, R. Owenius, and O. Vahtras, Chem. Phys. Lett. **338**, 407 (2001).

[66] Q. Daniel, P. Huang, T. Fan, Y. Wang, L. Duan, L. Wang, F. Li, Z. Rinkevicius, F. Mamedov, M. S. Ahlquist, *et al.*, Coord. Chem. Rev. **346**, 206 (2017).

[67] G. Lushington and F. Grein, Theor. Chim. Acta **93**, 259 (1996).

[68] D. Jayatilaka, J. Chem. Phys. **108**, 7587 (1998).

[69] G. Lushington and F. Grein, J. Chem. Phys. **106**, 3292 (1997).

[70] O. Vahtras, B. Minaev, and H. Ågren, Chem. Phys. Lett. **281**, 186 (1997).

[71] P. J. Bruna, G. H. Lushington, and F. Grein, Chem. Phys. **225**, 1 (1997).

[72] P. J. Bruna and F. Grein, Int. J. Quantum Chem. **77**, 324 (2000).

[73] M. Engström, O. Vahtras, and H. Ågren, Chem. Phys. **243**, 263 (1999).

[74] M. Engström, B. Minaev, O. Vahtras, and H. Ågren, Chem. Phys. **237**, 149 (1998).

[75] P. Bruna, G. Lushington, and F. Grein, J. Mol. Struct. THEOCHEM **527**, 139 (2000).

[76] S. Brownridge, F. Grein, J. Tatchen, M. Kleinschmidt, and C. M. Marian, J. Chem. Phys. **118**, 9552 (2003).

[77] E. Schrödinger, Phys. Rev. **28**, 1049 (1926).

[78] M. Born and R. Oppenheimer, Ann. Phys. **389**, 457 (1927).

[79] J. C. Slater, Phys. Rev. **34**, 1293 (1929).

[80] W. Pauli, Z. Phys. **31**, 765 (1925).

[81] S. F. Boys, Proc. R. Soc. London, Ser. A **200**, 542 (1950).

[82] C. C. J. Roothaan, Rev. Mod. Phys. **23**, 69 (1951).

[83] G. G. Hall, Proc. R. Soc. London, Ser. A **205**, 541 (1951).

[84] P.-O. Löwdin, J. Chem. Phys. **18**, 365 (1950).

[85] J. A. Pople and R. K. Nesbet, J. Chem. Phys. **22**, 571 (1954).

[86] C. Roothaan, Rev. Mod. Phys. **32**, 179 (1960).

[87] J. S. Andrews, D. Jayatilaka, R. G. Bone, N. C. Handy, and R. D. Amos, Chem. Phys. Lett. **183**, 423 (1991).

[88] P. Pulay, Chem. Phys. Lett. **100**, 151 (1983).

[89] P. Pulay and S. Saebø, Theor. Chim. Acta **69**, 357 (1986).

[90] M. Schütz, G. Hetzer, and H.-J. Werner, J. Chem. Phys. **111**, 5691 (1999).

[91] G. Hetzer, M. Schütz, H. Stoll, and H.-J. Werner, J. Chem. Phys. **113**, 9443 (2000).

[92] K. Kristensen, I.-M. Høyvik, B. Jansik, P. Jørgensen, T. Kjærgaard, S. Reine, and J. Jakowski, Phys. Chem. Chem. Phys. **14**, 15706 (2012).

[93] P. Baudin, P. Ettenhuber, S. Reine, K. Kristensen, and T. Kjærgaard, J. Chem. Phys. **144**, 054102 (2016).

[94] C. Edmiston and K. Ruedenberg, Rev. Mod. Phys. **35**, 457 (1963).

[95] N. J. Russ and T. D. Crawford, J. Chem. Phys. **121**, 691 (2004).

[96] J. E. Subotnik and M. Head-Gordon, J. Chem. Phys. **123**, 064108 (2005).

[97] J. E. Subotnik, A. Sodt, and M. Head-Gordon, J. Chem. Phys. **128**, 034103 (2008).

[98] D. G. Fedorov and K. Kitaura, J. Chem. Phys. **121**, 2483 (2004).

[99] Y. Mochizuki, K. Yamashita, T. Murase, T. Nakano, K. Fukuzawa, K. Takematsu, H. Watanabe, and S. Tanaka, Chem. Phys. Lett. **457**, 396 (2008).

[100] M. Kobayashi, T. Akama, and H. Nakai, J. Chem. Phys. **125**, 204106 (2006).

[101] J. Almlöf, Chem. Phys. Lett. **181**, 319 (1991).

[102] M. Häser and J. Almlöf, J. Chem. Phys. **96**, 489 (1992).

[103] M. Häser, Theor. Chim. Acta **87**, 147 (1993).

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

[105] Y. Jung, R. C. Lochan, A. D. Dutoi, and M. Head-Gordon, J. Chem. Phys. **121**, 9793 (2004).

[106] L. Goerigk and S. Grimme, WIREs Comput Mol Sci **4**, 576 (2014).

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

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

[109] J. P. Perdew and K. Schmidt, AIP Conf Proc **577**, 1 (2001).

[110] H. Chen and A. Zhou, Numer. Math. Theor. Meth. Appl. **1**, 1 (2008).

[111] A. D. Becke, J. Chem. Phys. **98**, 1372 (1993).

[112] J. Gauss, M. Kállay, and F. Neese, J. Phys. Chem. A **113**, 11541 (2009).

[113] O. Gunnarsson and B. I. Lundqvist, Phys. Rev. B **13**, 4274 (1976).

[114] E. Gross and W. Kohn, Phys. Rev. Lett. **55**, 2850 (1985).

[115] A. Görling, Phys. Rev. B **99**, 235120 (2019).

[116] M. Petersilka, U. J. Gossmann, and E. K. U. Gross, Phys. Rev. Lett. **76**, 1212 (1996).

[117] B. Mussard, D. Rocca, G. Jansen, and J. G. Ángyán, J. Chem. Theory Comput. **12**, 2191 (2016).

[118] H. Eshuis, J. Yarkony, and F. Furche, J. Chem. Phys. **132**, 234114 (2010).

[119] M. Fuchs, Y.-M. Niquet, X. Gonze, and K. Burke, J. Chem. Phys. **122**, 094116 (2005).

[120] J. Harl and G. Kresse, Phys. Rev. B **77**, 045136 (2008).

[121] J. Harl and G. Kresse, Phys. Rev. Lett. **103**, 056401 (2009).

[122] A. M. Burow, J. E. Bates, F. Furche, and H. Eshuis, J. Chem. Theory Comput. **10**, 180 (2014).

[123] D. L. Freeman, Phys. Rev. B **15**, 5512 (1977).

[124] A. Grüneis, M. Marsman, J. Harl, L. Schimka, and G. Kresse, J. Chem. Phys. **131**, 154115 (2009).

[125] J. E. Bates and F. Furche, J. Chem. Phys. **139**, 171103 (2013).

[126] E. Cuthill and J. McKee, in *Proceedings of the 1969 24th national conference* (ACM, New York, 1969) pp. 157–172.

[127] Y. Jung and M. Head-Gordon, Phys. Chem. Chem. Phys. **8**, 2831 (2006).

[128] A. L. Cauchy, Paris: Imprim. Royale (1821).

[129] M. Beuerle, D. Graf, H. F. Schurkus, and C. Ochsenfeld, J. Chem. Phys. **148**, 204104 (2018).

[130] C. Ochsenfeld, C. A. White, and M. Head-Gordon, J. Chem. Phys. **109**, 1663 (1998).

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

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

[133] P. E. Maslen, C. Ochsenfeld, C. A. White, M. S. Lee, and M. Head-Gordon, J. Phys. Chem. A **102**, 2215 (1998).

[134] S. Goedecker, Rev. Mod. Phys. **71**, 1085 (1999).

[135] B. A. Heß, C. M. Marian, U. Wahlgren, and O. Gropen, Chem. Phys. Lett. **251**, 365 (1996).

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

[137] C. Song and T. J. Martínez, J. Chem. Phys. **144**, 174111 (2016).

[138] C. Song and T. J. Martínez, J. Chem. Phys. **146**, 034104 (2017).

[139] E. G. Hohenstein, R. M. Parrish, C. D. Sherrill, and T. J. Martínez, J. Chem. Phys. **137**, 221101 (2012).

[140] R. M. Parrish, C. D. Sherrill, E. G. Hohenstein, S. I. Kokkila, and T. J. Martínez, J. Chem. Phys. **140**, 181102 (2014).

[141] E. G. Hohenstein, S. I. Kokkila, R. M. Parrish, and T. J. Martínez, J. Chem. Phys. **138**, 124111 (2013).

[142] E. G. Hohenstein, S. I. Kokkila, R. M. Parrish, and T. J. Martínez, J. Phys. Chem. B **117**, 12972 (2013).

[143] C. Song and T. J. Martínez, J. Chem. Phys. **149**, 044108 (2018).

[144] J. D. Carroll and J.-J. Chang, Psychometrika **35**, 283 (1970).

[145] S. I. Kokkila Schumacher, E. G. Hohenstein, R. M. Parrish, L.-P. Wang, and T. J. Martínez, J. Chem. Theory Comput. **11**, 3042 (2015).

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

[147] N. H. Beebe and J. Linderberg, Int. J. Quantum Chem. **12**, 683 (1977).

[148] I. Røeggen and E. Wisløff-Nilssen, Chem. Phys. Lett. **132**, 154 (1986).

[149] H. Koch and A. Sánchez de Merás, J. Chem. Phys. **113**, 508 (2000).

[150] D. Graf, M. Beuerle, H. F. Schurkus, A. Luenser, G. Savasci, and C. Ochsenfeld, J. Chem. Theory Comput. **14**, 2505 (2018).

[151] F. Aquilante, T. Bondo Pedersen, A. Sánchez de Merás, and H. Koch, J. Chem. Phys. **125**, 174101 (2006).

[152] S. F. Boys, Rev. Mod. Phys. **32**, 296 (1960).

[153] J. Pipek and P. G. Mezey, J. Chem. Phys. **90**, 4916 (1989).

[154] B. Jansík, S. Høst, K. Kristensen, and P. Jørgensen, J. Chem. Phys. **134**, 194104 (2011).

[155] I.-M. Høyvik, B. Jansik, and P. Jørgensen, J. Chem. Phys. **137**, 224114 (2012).

[156] D. P. Craig and T. Thirunamachandran, *Molecular Quantum Electrodynamics* (Dover Publications, New York, 1984).

[157] F. London, J. Phys. Radium **8**, 397 (1937).

[158] H. Hameka, Mol. Phys. **1**, 203 (1958).

[159] H. Hameka, Rev. Mod. Phys. **34**, 87 (1962).

[160] R. Ditchfield, J. Chem. Phys. **56**, 5688 (1972).

[161] W. Kutzelnigg, Isr. J. Chem. **19**, 193 (1980).

[162] A. E. Hansen and T. D. Bouman, J. Chem. Phys. **82**, 5035 (1985).

[163] A. Luzanov, E. Babich, and V. Ivanov, J. Mol. Struct. **311**, 211 (1994).

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

[165] T. Helgaker, M. Jaszunski, and K. Ruud, Chem. Rev. **99**, 293 (1999).

[166] R. Ditchfield, Mol. Phys. **27**, 789 (1974).

[167] K. Wolinski, J. F. Hinton, and P. Pulay, J. Am. Chem. Soc. **112**, 8251 (1990).

[168] M. Häser, R. Ahlrichs, H. Baron, P. Weis, and H. Horn, Theor. Chim. Acta **83**, 455 (1992).

[169] R. Stevens, R. Pitzer, and W. Lipscomb, J. Chem. Phys. **38**, 550 (1963).

[170] G. Schreckenbach and T. Ziegler, J. Phys. Chem. **99**, 606 (1995).

[171] J. R. Cheeseman, G. W. Trucks, T. A. Keith, and M. J. Frisch, J. Chem. Phys. **104**, 5497 (1996).

[172] G. Rauhut, S. Puyear, K. Wolinski, and P. Pulay, J. Phys. Chem. **100**, 6310 (1996).

[173] J. Gauss, Chem. Phys. Lett. **191**, 614 (1992).

[174] J. Gauss and J. F. Stanton, Advances in Chem. Phys. **123**, 355 (2002).

[175] J. Gauss and J. F. Stanton, J. Chem. Phys. **102**, 251 (1995).

[176] J. Gauss and J. F. Stanton, J. Chem. Phys. **104**, 2574 (1996).

[177] M. Kállay and J. Gauss, J. Chem. Phys. **120**, 6841 (2004).

[178] C. Ochsenfeld, J. Kussmann, and F. Koziol, Angew. Chem. **116**, 4585 (2004).

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

[180] M. Beer, J. Kussmann, and C. Ochsenfeld, J. Chem. Phys. **134**, 074102 (2011).

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

[182] J. Gauss, J. Chem. Phys. **99**, 3629 (1993).

[183] M. Bühl and T. van Mourik, WIREs Comput Mol Sci **1**, 634 (2011).

[184] J. Gauss and H.-J. Werner, Phys. Chem. Chem. Phys. **2**, 2083 (2000).

[185] S. Loibl and M. Schütz, J. Chem. Phys. **137**, 084107 (2012).

[186] G. L. Stoychev, A. A. Auer, J. Gauss, and F. Neese, J. Chem. Phys. **154**, 164110 (2021).

[187] S. Koseki, M. W. Schmidt, and M. S. Gordon, J. Phys. Chem. **96**, 10768 (1992).

[188] H. Laqua, T. H. Thompson, J. Kussmann, and C. Ochsenfeld, J. Chem. Theory Comput. **16**, 1456 (2020).

[189] D. Flaig, M. Beer, and C. Ochsenfeld, J. Chem. Theory Comput. **8**, 2260 (2012).

[190] A. M. Teale, O. B. Lutnæs, T. Helgaker, D. J. Tozer, and J. Gauss, J. Chem. Phys. **138**, 024111 (2013).