NMR Chemical Shift Computations at Second-Order Møller-Plesset Perturbation Theory Using Gauge-Including Atomic Orbitals and Cholesky-Decomposed Two-Electron Integrals

Sophia Burger, Filippo Lipparini, Jürgen Gauss, and Stella Stopkowicz

1) Department Chemie, Johannes Gutenberg-Universität Mainz, Duesbergweg 10-14, D-55128 Mainz, Germany
2) Dipartimento di Chimica e Chimica Industriale, Università di Pisa, Via G. Moruzzi 13, I-56124 Pisa, Italy

(Dated: 10 February 2022)

We report on a formulation and implementation of a scheme to compute NMR shieldings at second-order Møller-Plesset (MP2) perturbation theory using gauge-including atomic orbitals (GIAOs) to ensure gauge-origin independence and Cholesky decomposition (CD) to handle unperturbed as well as perturbed two-electron integrals. We investigate the accuracy of the CD for the derivatives of the two-electron integrals with respect to an external magnetic field as well as for the computed NMR shieldings, before we illustrate the applicability of our CD based GIAO-MP2 scheme in calculations involving up to about one hundred atoms and more than one thousand basis functions.

I. INTRODUCTION

The computation of NMR chemical shifts is an important application of quantum chemistry. For the accurate computation of NMR shieldings it has been amply shown that the consideration of electron-correlation effects is essential. In this respect, second-order Møller-Plesset (MP2) perturbation theory has been shown to be very useful, even though highly accurate predictions of NMR shieldings, in particular when aiming at absolute shieldings, require coupled-cluster (CC) treatments. Nevertheless, as the calculation of relative NMR chemical shifts benefits from some error cancellation, MP2 computations can provide a very useful and reliable tool. However, MP2 computations of NMR shieldings are quite costly in comparison to corresponding Hartree-Fock (HF) and density-functional theory (DFT) treatments and thus DFT computations are currently the first choice for NMR shielding computations, despite certain deficiencies of standard DFT in treating magnetic properties. Efforts to speed up MP2 computations of NMR shieldings and thus to increase the applicability of MP2 have a long history. Integral-direct schemes together with an efficient exploitation of point-group symmetry and coarse-grain parallelization have significantly extended the applicability of MP2. The use of local-correlation treatments has further enhanced the applicability of MP2, as shown by the work of Loibl and Schütz. These authors also use density fitting to avoid the computation of the perturbed two-electron integrals when using gauge-including atomic orbitals (GIAOs). Maurer and Ochsenfeld reported on a Laplace-based GIAO-MP2 formulation and implementation for the computation of NMR chemical shifts. This scheme, together with an efficient implementation, should in principle allow to achieve linear scaling. Stoychev et al. recently described the implementation of a scheme to compute GIAO-MP2 NMR chemical shifts within the domain-based local pair natural orbital (DLPNO) framework. The applications reported by these authors involved cases with up to 4700 basis functions and convincingly demonstrate the efficiency of this implementation which is part of the ORCA package. While the work by Stoychev et al. achieves computational efficiency by the use of a local-correlation treatment, of pair natural orbitals, and of density fitting together with an efficient exploitation of sparsity, there are more possibilities to reduce the overall cost of the calculation. One option is here to apply a Cholesky decomposition (CD) to the two-electron integrals. This option has in particular proven useful for medium-sized systems as CD alone does not have the potential to reach linear scaling. The advantage of CD over density fitting is that no auxiliary basis sets are required and that the error can be controlled in a rigorous manner. While there is an extensive literature on the use of CD within energy computations, less work has been reported for property computations. Recently, CD has been applied to accelerate the computation of nuclear gradients, but nothing has been so far reported concerning the use of CD for the computation of magnetic properties.

In this paper, we describe how CD can be used for the perturbed two-electron integrals that appear in GIAO computations of NMR shieldings and demonstrate the efficiency of a corresponding CD treatment for GIAO-MP2 computations on medium-sized systems with more than one thousand basis functions. In the following, after a brief review of standard GIAO-MP2 theory and CD, we discuss how CD is applied to the perturbed two-electron integrals and describe how CD is applied to the perturbed two-electron integrals and demonstrate the efficiency of a corresponding CD treatment for GIAO-MP2 computations on medium-sized systems with more than one thousand basis functions. The theory section is followed by a description of our implementation within the CFOUR program package. We then discuss the accuracy
of CD in GIAO-MP2 computations, before demonstrating the computational efficiency of our CD based GIAO-MP2 scheme in calculations.

II. THEORY

A. Standard GIAO-MP2 theory

We start by recapitulating the standard theory for the computations of NMR shieldings at the MP2 level when using GIAO-As. We follow here closely Ref. 21 in which spin-adapted expressions have been given.

The NMR shielding tensor \( \sigma^N \) of the \( N \)th nucleus in a molecule is defined as the second derivative of the energy with respect to the external magnetic field \( B \) and the magnetic moment \( m_N \) of the \( N \)th nucleus and most conveniently evaluated using the following density-matrix based expression:

\[
\sigma^N_{ij} = \sum_{\mu \nu} D_{\mu \nu} \frac{\partial^2 h_{\mu \nu}}{\partial B_i \partial m_N j} + \sum_{\mu \nu} \frac{\partial D_{\mu \nu}}{\partial B_i} \frac{\partial h_{\mu \nu}}{\partial m_N j}.
\]  

(1)

In Eq. (1), \( D_{\mu \nu} \) refers to the one-particle density matrix, with the Greek indices \( \mu \) and \( \nu \) labeling atomic orbitals (AOs), and \( h_{\mu \nu} \) denotes the matrix elements of the one-electron Hamiltonian. Expression for the derivatives of \( h_{\mu \nu} \) with respect to \( B_i \) and/or \( m_N \) can be, for example, found in Ref. 13. The MP2 contribution to the density matrix \( D_{\mu \nu} \) is usually defined in the corresponding molecular-orbital (MO) representation:

\[
D_{\mu \nu} = \sum_{pq} c_{\mu \rho}^* D_{pq} c_{\nu \sigma},
\]  

(2)

with \( c_{\mu \rho} \) specifying the MO coefficients obtained by solving the HF self-consistent-field (HF-SCF) equations. Indices \( i, j, \ldots \) label in the following occupied spatial orbitals, indices \( a, b, \ldots \) denote virtual spatial orbitals, and indices \( p, q, \ldots \) generic MOs that are either occupied or unoccupied. In the MO representation, the occupied-occupied and virtual-virtual block of the MP2 density matrix are given by

\[
D_{ij} = -2 \sum_{m \epsilon f} \varphi_{i m}^\epsilon \varphi_{j m}^f
\]  

(3)

and

\[
D_{ab} = \sum_{m \epsilon} \varphi_{a m}^\epsilon \varphi_{b m}^\epsilon.
\]  

(4)

Note that the given expression for the occupied-occupied block does not include the HF contribution to the density matrix. The MP2 amplitudes in Eqs. (3) and (4) are defined as

\[
\epsilon_{ij} = \frac{(ai|bj)}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b}
\]  

(5)

with \( (pq|rs) \) as the MO two-electron integrals in Mulliken notation, \( \epsilon_p \) as the orbital energy of the \( p \)th orbital, and \( \pi_{ij} \) as the corresponding “spin-adapted” amplitudes

\[
\pi_{ij} = 2t_{ij} - t_{ji}.
\]  

(6)

The virtual-occupied block of the MP2 density matrix is obtained by solving the Z-vector equations

\[
\sum_{m \epsilon} \sum_{e f} D_{m e} \left[ \delta(\epsilon_m|\epsilon_0) - \delta(\epsilon_m|\epsilon_1) - \delta(\epsilon_m|\epsilon_a) + \delta(\epsilon_m|\epsilon_b) \right] + \delta_{\epsilon m} \delta_{\epsilon e} (\epsilon_a - \epsilon_i) = -2X_{ai}
\]  

(7)

with the intermediate \( X_{ai} \) given by

\[
X_{ai} = \sum_{m \epsilon f} \sum_{e f m} \frac{(ea|f m)}{\epsilon_a - \epsilon_f} \delta_{\epsilon m} \delta_{\epsilon e} (\epsilon_a - \epsilon_i)
\]  

(8)

Expressions for the perturbed MP2 density matrix can be obtained by straightforward differentiation of Eqs. (3), (4), (7), and (8) with respect to the components \( B_i \) of the external magnetic field. The corresponding expressions for the occupied-occupied and virtual-virtual block of the perturbed MP2 density matrix are

\[
\frac{\partial D_{ij}}{\partial B_i} = -2 \sum_{m \epsilon f} \left\{ \frac{\partial \epsilon_{im}}{\partial B_i} \varphi_{i m}^\epsilon \varphi_{j m}^f + \frac{\partial \epsilon_{jm}}{\partial B_i} \varphi_{i m}^\epsilon \varphi_{j m}^f \right\}
\]  

(9)

and

\[
\frac{\partial D_{ab}}{\partial B_i} = 2 \sum_{m \epsilon} \left\{ \frac{\partial \epsilon_{a m}}{\partial B_i} \varphi_{a m}^\epsilon \varphi_{b m}^\epsilon + \frac{\partial \epsilon_{b m}}{\partial B_i} \varphi_{a m}^\epsilon \varphi_{b m}^\epsilon \right\}
\]  

(10)

while the perturbed virtual-occupied block is obtained as the solution to the perturbed Z-vector equation

\[
\sum_{m \epsilon f} \sum_{e f m} \frac{\partial D_{m e}}{\partial B_i} \left\{ \delta(\epsilon_m|\epsilon_0) - \delta(\epsilon_m|\epsilon_1) - \delta(\epsilon_m|\epsilon_a) + \delta(\epsilon_m|\epsilon_b) \right\}
\]  

(11)

with \( \partial X_{ai}/\partial B_i \) defined by

\[
\frac{\partial X_{ai}}{\partial B_i} = \sum_{m \epsilon f} \sum_{e f m} \left\{ \frac{\partial \epsilon_{im}}{\partial B_i} \varphi_{i m}^\epsilon \varphi_{f m}^e + \frac{\partial \epsilon_{jm}}{\partial B_i} \varphi_{i m}^\epsilon \varphi_{f m}^e \right\}
\]  

(12)
In the given equations $\frac{\partial f_{pq}}{\partial B_i}$ denotes the perturbed Fock matrix, $\frac{\partial (pq|rs)}{\partial B_i}$ the perturbed integrals
\[
\frac{\partial (pq|rs)}{\partial B_i} = + \sum_{\mu \nu \rho \sigma} \partial c_{\mu \rho}^{*} \partial c_{\nu \sigma}^{*} \partial c_{\rho \sigma}^{*} c_{\mu \sigma} (\mu \nu|\sigma \rho)
+ \sum_{\mu \nu \rho \sigma} c_{\mu \rho} \frac{\partial c_{\nu \sigma}}{\partial B_i} c_{\rho \sigma}^{*} c_{\mu \sigma} (\mu \nu|\sigma \rho)
+ \sum_{\mu \nu \rho \sigma} c_{\mu \rho}^{*} c_{\nu \sigma} \frac{\partial c_{\rho \sigma}}{\partial B_i} c_{\mu \sigma} (\mu \nu|\sigma \rho)
+ \sum_{\mu \nu \rho \sigma} c_{\mu \rho}^{*} c_{\nu \sigma} c_{\rho \sigma} \frac{\partial c_{\mu \sigma}}{\partial B_i} (\mu \nu|\sigma \rho),
\] with the updated diagonal elements of the two-electron integral matrix given by
\[
(\mu \nu|\mu \nu) = (\mu \nu|\nu \mu) - \sum_{R=1}^{p-1} L_{\mu \nu}^{R} L_{\mu \nu}^{*}. \tag{19}
\] Note that the CD follows a (partial) pivotal procedure\textsuperscript{40} in which in each iteration a new CD vector (with index $P$) is assigned to the largest of all updated diagonal elements of the two-electron integral matrix with indices $\mu$ and $\nu$. The decomposition is continued until the largest updated diagonal element is smaller than a predefined Cholesky threshold $10^{-6}$. This threshold also determines the accuracy of the decomposition, as it can be shown via the Cauchy-Schwarz inequality that the error of the two-electron integrals, approximated via Eq. (17), is in absolute terms always smaller than $10^{-6}$.

It has been amply shown\textsuperscript{41,42} that the storage requirements for the CVs are substantially lower than for the two-electron integrals so that even for quite large calculations (with more than one thousand basis functions) the whole set of CVs can be kept in core memory unlike the two-electron integrals which either have to be stored on disk or handled using integral-direct algorithms.

For the two-electron integrals in the MO representation, the CVs are transformed from the AO into the MO representation
\[
L_{pq}^{P} = \sum_{\sigma \rho} c_{\sigma \rho}^{*} L_{\sigma \rho}^{P} c_{pq}
\] such that the MO two-electron integrals are given by
\[
(pq|rs) \approx \sum_{P=1}^{M} L_{pq}^{P} L_{rs}^{*}. \tag{21}
\] As for the AO CVs, it is also for the MO CVs usually possible to keep all of them in memory.

Note also that due to the eightfold permutational symmetry of the two-electron integrals the CVs are symmetric with respect an interchange of the two AO or MO indices.

### C. Cholesky decomposition of the magnetic two-electron integral derivatives

For derivatives of the two-electron integrals, a CD scheme can be derived by differentiating Eqs. (17) and (18) with respect to the corresponding perturbation.\textsuperscript{50} In the case of a magnetic field $B$ as perturbation, this yields
\[
\frac{\partial (\sigma \rho|\nu \mu)}{\partial B_i} \approx \sum_{P=1}^{M} \left\{ \frac{\partial L_{\sigma \rho}^{P}}{\partial B_i} L_{\mu \nu}^{P} - L_{\sigma \rho}^{P} \frac{\partial L_{\mu \nu}^{P}}{\partial B_i} \right\}, \tag{22}
\]
and
\[
\frac{\partial L_{\sigma \rho}^{P}}{\partial B_i} = (\nu \mu|\nu \mu) \left\{ \frac{\partial (\sigma \rho|\nu \mu)}{\partial B_i} \right\} - \sum_{R=1}^{p-1} \left\{ \frac{\partial L_{\sigma \rho}^{R}}{\partial B_i} L_{\mu \nu}^{R} - L_{\sigma \rho}^{R} \frac{\partial L_{\mu \nu}^{R}}{\partial B_i} \right\}. \tag{23}
\]
Note that there are some differences to the corresponding equations given in Ref. 50 for nuclear coordinates as perturbation. For a magnetic field the derivative two-electron integrals (and so the derivatives of the CVs) are purely imaginary. As a consequence, the derivatives of the updated diagonal element of the two-electron integral matrix vanishes and Eq. (22) consists of the difference instead of the sum of two terms. A similar observation holds for the the correction term to the integral derivative in Eq. (23).

Furthermore, the perturbed two-electron integrals no longer exhibit the full eightfold permutational symmetry. However, one can split the perturbed two-electron integrals according to
\begin{equation}
\frac{\partial (\sigma \rho | v \mu)}{\partial B_i} = \left( \frac{\partial (\sigma \rho | v \mu)}{\partial B_i} + (\sigma \rho | \frac{\partial v \mu}{\partial B_i}) \right)
\end{equation}

and equate the first (second) term on the right hand side of Eq. (24) with the first (second) term on the right hand side of Eq. (22). As these partial derivatives in Eq. (24) exhibit the full permutational symmetry albeit with an additional consideration of a sign change, one can choose the corresponding perturbed CVs antisymmetric with respect to an AO index change. Eq. (23) can be then recast in the following form
\begin{equation}
\frac{\partial L_{\sigma \rho}^{\mu \nu}}{\partial B_i} = \left\{ \frac{1}{2} \frac{\partial (\sigma \rho | v \mu)}{\partial B_i} + \frac{\partial (\sigma \rho | v \mu)}{\partial B_i} \right\} - \sum_{k=1}^{P-1} \frac{\partial L_{\sigma \rho}^{\mu \nu}}{\partial B_i} \frac{\partial L_{\sigma \rho}^{\mu \nu}}{\partial B_k}.
\end{equation}

We emphasize that Eq. (25) is not an independent CD of the perturbed integrals, as the Cholesky basis, which is completely determined by the unperturbed two-electron integral matrix, is already defined. In other words, with Eq. (25) we are building a representation of the perturbed integrals in a given Cholesky basis.

The perturbed MO two-electron integrals comprise not only the contribution due to the perturbed AO two-electron integrals, but also contributions due to the perturbed MO coefficients (see Eq. (13)). It is thus advantageous to define the perturbed CVs in the MO representation as
\begin{equation}
\frac{\partial L_{\sigma \rho}^{\mu \nu}}{\partial B_i} = \sum_{\sigma \rho} \left\{ (\sigma \rho | \frac{\partial v \mu}{\partial B_i}) c_{\sigma \rho} - \frac{\partial (\sigma \rho | v \mu)}{\partial B_i} c_{\sigma \rho} + c_{\sigma \rho} L_{\sigma \rho} \frac{\partial c_{\sigma \rho}}{\partial B_i} \right\}
\end{equation}

such that the perturbed MO two-electron integrals can be approximated by
\begin{equation}
\frac{\partial (pq|rs)}{\partial B_i} \approx \sum_{P=1}^{M} \left\{ \frac{\partial L_{pq}^{fr}}{\partial B_i} L_{pq}^{sr} - \frac{\partial L_{pq}^{fr}}{\partial B_i} L_{pq}^{sr} \right\} \approx \sum_{P=1}^{M} \left\{ \frac{\partial L_{pq}^{fr}}{\partial B_i} L_{pq}^{sr} + L_{pq}^{sr} \frac{\partial L_{pq}^{fr}}{\partial B_i} \right\}.
\end{equation}

D. GIAO-MP2 theory with Cholesky-decomposed two-electron integrals

A CD based implementation of GIAO-MP2 can be carried out by replacing in the expressions given in section II A all the two-electron integrals by their CD equivalents. In the following, we give the corresponding equations by explicitly replacing the (derivative) integrals with their CD only when this allows for an alternative evaluation of the term; otherwise we will just use $(pq|rs)$ and $\partial (pq|rs)$ CD/\partial B$_i$ to indicate that the term is evaluated with (derivative) two-electron integrals reconstructed from the CD. For the terms in which we explicitly insert the CD, we indicate a possible way for their evaluation by setting appropriate parentheses.

Before giving the detailed equations, we note that we decided to use in the case of the $(ij|ka)$ as well in most cases for the $(ai|bj)$ integrals (and as well also for the corresponding integrals) the reconstructed integrals. These integrals can be computed from the (derivative) integrals directly after the transformation of the integrals from the AO into the MO representation. From the $(ai|bj)$ CD integrals or from the corresponding derivative integrals $\partial (ai|bj)$ CD/\partial B$_i$, it is possible to obtain the (perturbed) MP2 amplitudes
\begin{equation}
\text{CD}_{ij}^{ab} = \frac{(ai|bj)\text{CD}}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b}
\end{equation}

and
\begin{equation}
\frac{\partial \text{CD}_{ij}^{ab}}{\partial B_i} = \left\{ \frac{\partial (ai|bj)\text{CD}}{\partial B_i} + \sum_e \left[ \frac{\partial f_{ae}}{\partial B_i} \text{CD}_{ij}^{ab} + \frac{\partial f_{be}}{\partial B_i} \text{CD}_{ij}^{ae} \right] \right\} \frac{1}{(\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b)}.
\end{equation}

Using these amplitudes, one can compute the occupied-occupied and virtual-virtual blocks of the (perturbed) MP2 density matrix according to Eqs. (3), (4), (9), and (10). For the virtual-occupied block of the (perturbed) density matrix, insertion of the CD yields
\begin{equation}
4 \sum_p L_{ia}^{pe} \left\{ \sum_{m} \sum_{e} D_{em} L_{em}^{pe} - \sum_{p} L_{ia}^{pe} \left\{ \sum_{e} D_{em} L_{em}^{pe} \right\} \right\} - \sum_{p} \sum_{m} L_{im}^{pe} \left\{ \sum_{e} D_{em} L_{em}^{pe} \right\} - \sum_{p} L_{ia}^{pe} \left\{ \sum_{e} D_{em} L_{em}^{pe} \right\} + D_{ai}(\epsilon_a - \epsilon_i) = -2 \text{CD} X_{ai},
\end{equation}

with
\begin{align}
\text{CD} X_{ai} &= \sum_{p} \sum_{m} \sum_{f} L_{jm}^{pi} \left\{ \sum_{e} L_{ea}^{pf} \text{CD} \sum_{e} \frac{1}{2} \left( m | n \right)^{\text{CD}} D_{mn} \right\} \\
&- \sum_{mn} \sum_{e} (mn|ea)^{\text{CD}} \text{CD} \sum_{e} D_{mn} \\
&+ \sum_{mn} \sum_{e} (mn|ia)^{\text{CD}} - \frac{1}{2} (ma|in)^{\text{CD}} \\
&+ \sum_{p} L_{ia}^{pe} \left\{ \sum_{e} D_{ef} L_{ef}^{pe} \right\} - \frac{1}{2} \sum_{p} \sum_{e} L_{ea}^{pe} \left\{ \sum_{e} L_{ef}^{pe} D_{ef} \right\}.
\end{align}
The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.

The im-
plemented within the CFOUR program package.
A. Accuracy of the Cholesky decomposition of the derivatives of the two-electron integrals with respect to the components of an external magnetic field

In Table I, we compare the accuracy of the CD of the unperturbed and perturbed two-electron integrals for different Cholesky thresholds $10^{-\delta}$. The computations have been carried out for water ($\text{H}_2\text{O}$, $r(\text{OH})=1.0$ Å, $\langle\text{HOH}\rangle=100.0^\circ$) and hydrogen peroxide ($\text{H}_2\text{O}_2$, $r(\text{OO})=1.5$ Å, $r(\text{OH})=1.0$ Å, $\langle\text{OOH}\rangle=100.0^\circ$, $\tau(\text{HOOH})=120.0^\circ$) using Dunning’s cc-pVXZ basis sets with $X = \text{D}$, $\text{T}$, and $\text{Q}$. While for the unperturbed integrals the Cholesky threshold provides a rigorous upper bound for the error in the CD, this is not the case for the perturbed integrals. It is seen that the errors in the perturbed integrals are somewhat larger (by a factor of about 10 to 200) than for the unperturbed integrals and that the error is somewhat larger for calculations with larger basis sets. The error due to the CD is furthermore found to be of similar magnitude in the case of $\text{H}_2\text{O}$ and $\text{H}_2\text{O}_2$, thus indicating that the Cholesky threshold will be a useful for judging the accuracy of the unperturbed and perturbed integrals also for larger molecules.

Overall, we conclude that the Cholesky threshold $\delta$ is a useful measure for the accuracy of the CD in case of the perturbed integrals.

B. Accuracy of NMR shieldings computed at the GIAO-MP2 level with Cholesky-decomposed two-electron integrals

For the three organic molecules acetaldehyde ($\text{H}_3\text{C}=-\text{CHO}$), vinyl alcohol ($\text{H}_2\text{C}=\text{CHOH}$), and ethylene oxide ($\text{C}_2\text{H}_4\text{O}$) we compare in Table II the results from CD-based GIAO-HF and GIAO-MP2 computations with those from standard GIAO-HF and GIAO-MP2 treatments, again for different Cholesky thresholds $\delta$ and for different basis sets from Dunning’s cc-pVXZ hierarchy with $X = \text{D}$, $\text{T}$, and $\text{Q}$. The geometries of these molecules have been determined at the same level as the NMR shieldings and are given for completeness in the supplementary material. Table II reports the corresponding maximum absolute errors in the isotropic shieldings obtained in the CD-based computations. One can conclude from the data in Table II that the corresponding errors are for all Cholesky thresholds small and of no relevance for actual computations. The error amounts to several hundredths ppm in the case of $\delta=4$, to about a few thousandths ppm in the case of $\delta=5$, and to less than a thousandth ppm in the case of $\delta=6$. The comparison suggests that a Cholesky threshold of $\delta=5$ is sufficient to guarantee converged values for the shieldings, though the results obtained with $\delta=4$ already exhibit an accuracy that is fully acceptable for chemical applications.

At this point, it is interesting to compare the accuracy of the present CD based scheme with the accuracy of corresponding
TABLE I. Maximum errors (in a.u.) in the computed unperturbed and perturbed two-electron integrals for different CD thresholds $10^{-\delta}$ in calculations for water ($\text{H}_2\text{O}$) and hydrogen peroxide ($\text{H}_2\text{O}_2$) using the cc-pVXZ basis sets with $X = D, T, \text{and } Q$.

| basis set | two-electron integrals | Cholesky threshold $\delta$ |
|-----------|-------------------------|-----------------------------|
|           |                         | 4   | 5   | 6   | 7   | 8   | 9   |
| a) $\text{H}_2\text{O}$ |                          |     |     |     |     |     |     |
| cc-pVDZ  | unperturbed              | $8.1 \times 10^{-5}$ | $9.5 \times 10^{-6}$ | $1.0 \times 10^{-6}$ | $9.4 \times 10^{-8}$ | $6.4 \times 10^{-9}$ | $8.7 \times 10^{-10}$ |
|          | perturbed                | $3.4 \times 10^{-4}$ | $1.0 \times 10^{-4}$ | $8.6 \times 10^{-6}$ | $2.3 \times 10^{-6}$ | $8.1 \times 10^{-7}$ | $7.4 \times 10^{-8}$ |
| cc-pVTZ  | unperturbed              | $9.5 \times 10^{-5}$ | $9.8 \times 10^{-6}$ | $8.9 \times 10^{-7}$ | $8.4 \times 10^{-8}$ | $8.7 \times 10^{-9}$ | $1.0 \times 10^{-9}$ |
|          | perturbed                | $3.4 \times 10^{-4}$ | $6.3 \times 10^{-5}$ | $8.4 \times 10^{-6}$ | $1.9 \times 10^{-6}$ | $3.1 \times 10^{-7}$ | $9.1 \times 10^{-8}$ |
| cc-pVQZ  | unperturbed              | $9.6 \times 10^{-5}$ | $9.5 \times 10^{-6}$ | $9.6 \times 10^{-7}$ | $9.9 \times 10^{-8}$ | $9.9 \times 10^{-9}$ | $9.6 \times 10^{-10}$ |
|          | perturbed                | $4.9 \times 10^{-4}$ | $4.7 \times 10^{-5}$ | $1.7 \times 10^{-5}$ | $3.3 \times 10^{-6}$ | $9.2 \times 10^{-7}$ | $2.0 \times 10^{-7}$ |
| b) $\text{H}_2\text{O}_2$ |                          |     |     |     |     |     |     |
| cc-pVDZ  | unperturbed              | $9.4 \times 10^{-5}$ | $9.5 \times 10^{-6}$ | $9.3 \times 10^{-7}$ | $7.9 \times 10^{-8}$ | $8.0 \times 10^{-9}$ | $8.8 \times 10^{-10}$ |
|          | perturbed                | $5.6 \times 10^{-4}$ | $4.4 \times 10^{-4}$ | $6.8 \times 10^{-6}$ | $8.9 \times 10^{-7}$ | $2.9 \times 10^{-7}$ | $5.9 \times 10^{-8}$ |
| cc-pVTZ  | unperturbed              | $9.4 \times 10^{-5}$ | $9.8 \times 10^{-6}$ | $9.9 \times 10^{-7}$ | $9.9 \times 10^{-8}$ | $9.4 \times 10^{-9}$ | $9.8 \times 10^{-10}$ |
|          | perturbed                | $4.9 \times 10^{-4}$ | $4.5 \times 10^{-5}$ | $5.6 \times 10^{-6}$ | $1.9 \times 10^{-6}$ | $4.4 \times 10^{-7}$ | $9.5 \times 10^{-8}$ |
| cc-pVQZ  | unperturbed              | $9.5 \times 10^{-5}$ | $9.9 \times 10^{-6}$ | $9.4 \times 10^{-7}$ | $9.5 \times 10^{-8}$ | $9.9 \times 10^{-9}$ | $9.9 \times 10^{-10}$ |
|          | perturbed                | $5.0 \times 10^{-4}$ | $9.7 \times 10^{-5}$ | $2.2 \times 10^{-5}$ | $6.6 \times 10^{-6}$ | $1.1 \times 10^{-6}$ | $2.3 \times 10^{-7}$ |

TABLE II. Maximum absolute errors (given in ppm) in the computed isotropic shieldings (with respect to a standard GIAO-HF and GIAO-MP2 computation) for acetaldehyde, ethylene oxide, and vinyl alcohol obtained in CD based GIAO-HF and GIAO-MP2 computations for different Cholesky thresholds $\delta$ using the cc-pVXZ basis sets with $X = D, T, \text{and } Q$.

| nucleus | $\delta=4$ | $\delta=5$ |
|---------|------------|------------|
|         | cc-pVDZ    | cc-pVTZ    | cc-pVQZ    | cc-pVDZ | cc-pVTZ | cc-pVQZ |
| a) HF treatment |             |             |             |         |         |         |
| $^{13}\text{C}$ | $0.013$   | $0.002$   | $0.008$   | $0.002$  | $0.001$  | $0.001$  |
| $^{17}\text{O}$ | $0.063$   | $0.007$   | $0.021$   | $0.001$  | $0.004$  | $0.001$  |
| $^{1}\text{H}$  | $0.001$   | $0.001$   | $0.001$   | $0.000$  | $0.000$  | $0.001$  |
| b) MP2 treatment |             |             |             |         |         |         |
| $^{13}\text{C}$ | $0.013$   | $0.002$   | $0.081$   | $0.001$  | $0.002$  | $0.002$  |
| $^{17}\text{O}$ | $0.049$   | $0.015$   | $0.037$   | $0.003$  | $0.002$  | $0.003$  |
| $^{1}\text{H}$  | $0.001$   | $0.001$   | $0.001$   | $0.001$  | $0.001$  | $0.000$  |

One might also consider a mixed scheme that combines a rigorous HF treatment (without CD) with a CD based MP2 part. In this way, one may hope to reduce the error when using a rather loose Cholesky threshold (i.e., $\delta=4$). However, based on results from some exploratory calculations, it is concluded that such a scheme offers no significant improvement for the computation of NMR shieldings.

C. Representative Applications

To illustrate the applicability of our CD based GIAO-MP2 computations, we report in the following results from corresponding calculations for several large molecules with up to close to 100 atoms and more than 1000 basis functions. The chosen examples comprise coronene ($C_{24}H_{12}$), hexabenzo coronene ($C_{42}H_{18}$), tetrakis(t-butyl)tetraborane(4) ($B_4C_{16}H_{36}$), tetrameric cyclopentadienyl aluminum(I) ($Al_4C_{20}H_{20}$), the buckminsterfullerene ($C_{60}$), and a tweezer host-guest complex ($C_{34}N_2H_{36}$). The structures of these molecules are depicted in Figure 2 with the corresponding Cartesian coordinates given in the supplementary material.

The geometries for the calculations have been taken from Ref. 33 in the case of coronene and the tweezer host-guest complex, from Ref. 62 in the case of hexabenzo coronene, from Ref. 22 in the case of $B_4t$-$B_4u$ as well as $Al_4Cp_4$, and from Ref. 63 in the case of $C_{60}$. For a detailed account why NMR chemical-shift computations for these molecules are important, we refer in the case of hexabenzo coronene to Ref. 62, in the case of the tetrahedral boron compound to Ref. 2, in the case of the tetrameric aluminum(I) compound to Ref. 64, and for the tweezer host-guest complex to Ref. 65.
The calculations have been performed with the dzp (11s7p1d/6s4p1d for Al, 8s4p1d/4s2p1d for B, C, O as well as N, and 4s1p/2s1p for H) and tz2p (12s9p2d/7s5p2d for Al, 9s5p2d/5s3p2d for B, C, O, as well as N, and 5s2p/3s2p for H) versions of the Karlsruhe basis set and polarization functions from Ref. 13; only for the rather large tweezer host-guest complex the larger calculation had to be restricted to tzp (9s5p1d/5s3p1d for C as well as N and 5s1p/3s1p for H) instead of tz2p. For completeness, the basis sets as well as the results of the calculations, i.e., the obtained shieldings are given in the supplementary material, while the focus in the following is on the size of these computations and the computational resources required to perform the computations.

In Table III we report the number of basis functions ($N_{bf}$), the number of CVs ($N_{CV}$), and the CD compression rate (ratio of maximum number and actual number of CVs) for the computations with two different Cholesky thresholds ($\delta = 4$ and $\delta = 5$). We also report the timings of the calculations together with the memory requirements. As seen from the table, the required resources are still rather modest for calculations with a couple of hundred basis functions, while they are quite demanding for the computations with more than 1000 basis functions. This finding indicates that the application range of our CD-GIAO-MP2 approach comprises cases with several hundred up to one thousand basis functions, while calculations with significantly more than 1000 basis function require additional means to exploit sparsity and the local nature of electron correlation and thus are better treated using, for example, the recently reported DLPNO-MP2 approach for computing NMR shieldings.

A closer look at Table III reveals that the CD very effectively reduces the memory requirements for the two-electron integrals. The CD compression rates are in all cases high (i.e., above 28) and reach for the most demanding computations on $\text{C}_{60}$ and the tweezer host-guest complex in the case of $\delta = 5$ values of 94 and 98, respectively. The actual number of CVs is in all cases merely several thousand with only for $\text{C}_{60}$ when using the tz2p basis and a threshold of 5 reaching a value above 10000. For a threshold of 4, the number of CVs is typically by about 20 to 25% lower and the compression rates are higher by the same percentage. For $\text{C}_{60}$, this means that for $\delta = 4$ only 8776 CVs are required instead of the 11022 CVs for $\delta = 5$.

For the calculation on $\text{C}_{60}$ with the tz2p basis and $\delta = 5$, the
memory that is required to keep the whole set of CVs in core is about 86 GB for the unperturbed two-electron integrals and about 256 GB for the perturbed two-electron integrals, i.e., amounts of memory that nowadays are easily affordable. In the case of \( \delta = 4 \), these values reduce to about 68 GB for the unperturbed and to about 204 GB for the perturbed integrals. Loosening the Cholesky threshold does not only reduce the memory requirements but also speeds up the calculation (at least those steps whose cost scales with the number of CVs). For the large calculation on \( C_{60} \) with the tz2p basis, this means that the memory requirement is reduced from 1089 GB (\( \delta = 5 \)) to 1023 GB (\( \delta = 4 \)) and the calculation only required 3 days, 9 hours, and 49 minutes instead of 4 days, 11 hours, and 31 minutes (see table III). The computational results differed by at most 0.02 ppm which is negligible for chemical applications.

The limiting factor of GIAO-MP2 computations using the currently implemented CD based scheme is the necessity to store for the MP2 treatment two vectors of length \( N_{\text{virt}}^2 N_{\text{occ}}^2 \) in memory (one vector requires about 384 GB in the case of \( C_{60} \) and the tz2p basis). For that reason the largest computations (on the tweezer host-guest complex and \( C_{60} \), both with more than 1000 basis functions) needed 814 or even 1089 GB core memory, respectively. However, the actual computational times (we report wall-clock times) are in all cases acceptable with only the largest computations requiring more than a day. The tz2p computation for \( C_{60} \) required about 4 days and 12 hours when using 20 cores with \( \delta = 5 \) and 3 days and 10 hours with \( \delta = 4 \), while the tzp calculation for the tweezer host-guest complex has been completed, again using 20 cores, in slightly less than 2 days and 10 hours. We note that parallelization is essential for ensuring reasonable wall-clock timings, even though the parallelization efficiency is unfortunately not very high. For example, the dzp computation on the tweezer host-guest complex required about 5 days when running on one CPU and only 1 day and 5 hours when running on 20 CPUs. The speed up of about 4.1 is not too convincing, but we note that at this stage no specific effort was put into achieving an efficient parallelization, as we only used threaded matrix-matrix multiplication routines from an appropriate BLAS library. We are convinced that with additional effort a higher parallelization speed up can be reached, but this is an issue beyond the scope of the present paper.

In Table IV, we present a detailed analysis of the computational cost for the dzp and tzp computations on the tweezer host-guest complex. As it is seen, the costs are dominated by the following steps: (a) the CD of the (perturbed) two-electron integrals, all SCF related steps (solution of the HF-SCF, CPHF, as well as unperturbed and perturbed Z-vector equations), the construction of unperturbed and perturbed \( (ai|bj) \) integrals, and the unperturbed and perturbed \( (ab|ci) \) contributions to the unperturbed and perturbed \( X_{ai} \) intermediate. These findings are expected, as the costs of these steps are those that in a formal analysis appear the most costly. The cost for the CD scales as \( N_{\text{occ}}^2 N_{\text{virt}}^2 \), the SCF related steps exhibit a scaling of \( N_{\text{occ}}^2 N_{\text{occ}} N_{\text{CV}} \) (SCF) or \( N_{\text{virt}}^2 N_{\text{occ}} N_{\text{CV}} \) (CPHF, Z-vector equations), respectively, due to the exchange contributions, the formation of the \( (ai|bj) \) integrals scales as \( N_{\text{CV}} N_{\text{virt}}^2 N_{\text{occ}}^2 \), and the cost for the \( (ab|ci) \) contributions to the \( X_{ai} \) intermediates are of the same order. The efficiency of a CD based calculation thus depends in a rather crucial manner on the number of Cholesky vectors used for the representation of the unperturbed and perturbed two-electron integrals.

Finally, we note that in case of corone and the dzp basis (420 basis functions) the CD based GIAO-MP2 scheme outperforms the standard GIAO-MP2 approach based on the regular two-electron integrals. The CD based calculation (with \( \delta=4 \)) runs in 2 hours and 5 minutes, while the standard GIAO-MP2 calculations requires 9 hours and 42 minutes (both calculations have been carried out on 1 CPU of an Intel Xeon(R) Gold 5215M node running at 2.5 GHz). This clearly shows that CD based GIAO-MP2 computations with several hundred basis functions are already beyond the break-even point at which the CD based schemes surpasses the standard approach in terms of computational efficiency.

V. CONCLUSIONS AND OUTLOOK

In this paper, we report on an MP2 scheme for the computation of NMR shieldings that uses a Cholesky decomposition for the handling of the unperturbed and perturbed two-electron integrals. The latter arise when using GIAOs for ensuring gauge-origin independence. The storage and handling of these integrals is the main bottleneck of traditional electron-correlated NMR chemical-shift computations and has significantly limited their applicability in the past. As shown, the CD for the derivatives of the two-electron integrals with respect to the components of an external magnetic field can be achieved by using a recipe obtained by straightforward differentiation of the relevant equations for the Cholesky decomposition of the corresponding unperturbed integrals. Such a scheme has already been successfully applied for geometrical derivatives of the two-electron integrals and is used here for the first time for an external magnetic field as perturbation. As shown, the CD of the perturbed two-electron integrals leads to a very compact representation that allows to maintain the whole set of unperturbed and perturbed Cholesky vectors in core memory, even for very large calculations. We have derived the required formulae for a CD-based GIAO-MP2 approach and have reported on its first implementation. Calculations on systems consisting of close to one hundred atoms and with more than one thousand basis functions demonstrate the applicability of our scheme. However, for even larger systems it becomes mandatory to couple the present CD based scheme with techniques that enable a local treatment of correlation and that allow to exploit sparsity.

The present paper represents a first step along the lines of formulating and implementing electron-correlated approaches for the computation of magnetic properties for large molecules. Further work can be envisioned and is planned in two directions. The first is to move to a multiscale description for computing NMR chemical shifts in which the present CD based GIAO-MP2 scheme provides the engine for the treatment of the QM region. MP2 fails in the accurate prediction of absolute shieldings (it has a strong
### TABLE III. Computational requirements for the CD based GIAO-MP2 computations (with Cholesky thresholds $\delta=4$ and $\delta=5$) for the molecules given in Fig. 2. The number of basis functions is denoted by $N_{bf}$, $N_{el}$ is the number of electrons, $N_{CPU}$ specifies the number of CPUs used, and $N_{CV}$ the number of Cholesky vectors in the corresponding calculation. The compression is defined as the ratio of the theoretically maximum number of Cholesky vectors and the actual number $N_{CV}$. The required memory is given in GB and the wall-clock time $t_{wall}$ in terms of days, hours, and minutes. If not otherwise noted, calculations have been carried out on an Intel Xeon(R) E5-2643 node running at 3.4 GHz.

| molecule                  | basis | $N_{bf}$ | $N_{el}$ | $N_{CPU}$ | $N_{CV}$ compression | memory [GB] | $t_{wall}$ [d:h:min] | $N_{CV}$ | compression | memory [GB] | $t_{wall}$ [d:h:min] |
|---------------------------|-------|----------|----------|-----------|-----------------------|-------------|----------------------|----------|-------------|-------------|----------------------|
| coronene                  | dzp   | 420      | 156      | 8         | 2493                  | 35.46       | 17 1:05              |          |             |             | 3097 28.55           | 19 1:20       |
|                           | tz2p  | 684      | 156      | 8         | 3917                  | 59.81       | 60 3:50              |          |             |             | 4959 47.24           | 67 4:50       |
| hexabenzocoronene         | dzp   | 720      | 270      | 8         | 4310                  | 60.22       | 124 5:38             |          |             |             | 5331 48.69           | 132 8:28      |
|                           | tz2p  | 1170     | 270      | 8         | 6710                  | 102.09      | 422 10:12            |          |             |             | 8546 80.16           | 457 1:07:35   |
| $B_{4tBu4}$               | dzp   | 480      | 152      | 8         | 2627                  | 43.94       | 23 1:15              |          |             |             | 3402 33.93           | 25 1:22       |
|                           | tz2p  | 804      | 152      | 8         | 4281                  | 75.59       | 86 4:51              |          |             |             | 5511 58.72           | 97 5:57       |
| $Al_{4}Cp_{4}$            | dzp   | 492      | 192      | 8         | 2738                  | 44.29       | 31 1:35              |          |             |             | 3494 34.71           | 34 1:58       |
|                           | tz2p  | 788      | 192      | 8         | 4414                  | 70.43       | 105 5:37             |          |             |             | 5524 56.28           | 115 6:56      |
| $C_{60}$                  | dzp   | 900      | 360      | 8         | 5502                  | 73.69       | 312 19:07            |          |             |             | 6722 60.32           | 325 23:20     |
|                           | tz2p  | 1440     | 360      | 8         | 8776                  | 118.22      | 1023 3:09:49         |          |             |             | 11022 94.13          | 1089 4:11:31  |
| tweezer host-guest complex| dzp   | 1020     | 374      | 20        | 7400                  | 70.37       | 649 1:06:11          |          |             |             | 7400 70.37           | 649 1:07:35   |
|                           | tzp   | 1280     | 374      | 20        | 8344                  | 98.26       | 814 2:09:26          |          |             |             | 8344 98.26           | 814 2:09:26   |

$^a$ calculation has been carried out on an Intel Xeon(R) Gold 5215M node running at 2.5 GHz.

### TABLE IV. Breakup of the computational timings (in minutes) for the CD based GIAO-MP2 computations (with $\delta=5$) for the tweezer host-guest complex using a dzp and tzp basis, respectively. Calculations have been carried out on an Intel Xeon(R) Gold 5215M node running at 2.5 GHz.

| computational step | dzp basis$^a$ | dzp basis$^b$ | tzp basis$^b$ |
|--------------------|---------------|---------------|---------------|
| CD of two-electron integrals | 163           | 33            | 60            |
| solution of HF-SCF equations | 221           | 57            | 208           |
| transformation of unperturbed CVs into MO basis | 33            | 7             | 19            |
| construction of integrals $(ij|ka)$ | 34            | 4             | 6             |
| construction of integrals $(ai|bj)$ | 261           | 35            | 67            |
| construction of $t$ amplitudes | 11            | 1             | 1             |
| computation of occ.-occ. block of density matrix | 11            | 1             | 2             |
| computation of vrt.-vrt. block of density matrix | 51            | 5             | 12            |
| evaluation of the contribution due to $(ij|ka)$ to $X_{ai}$ | 35            | 4             | 8             |
| evaluation of the contribution due to $(ab|ci)$ to $X_{ai}$ | 268           | 51            | 99            |
| solution of Z-vector equation | 52            | 17            | 37            |
| CD of perturbed two-electron integrals | 805           | 125           | 230           |
| solution of CPHF equations | 101           | 24            | 46            |
| transformation of perturbed CVs with unperturbed MOs | 3 x 33        | 3 x 7         | 3 x 21        |
| transformation of unperturbed CVs with perturbed MOs | 3 x 64        | 3 x 12        | 3 x 32        |
| construction of integrals $\partial(ij|ka)/\partial B_i$ | 3 x 91        | 3 x 12        | 3 x 18        |
| construction of integrals $\partial(ai|bj)/\partial B_i$ | 3 x 514       | 3 x 69        | 3 x 134       |
| evaluation of the contribution due to $\partial(ij|ka)/\partial B_i$ to $\partial X_{ai}/\partial B_i$ | 3 x 25        | 3 x 4         | 3 x 8.3        |
| evaluation of the contribution due to $\partial(ab|ci)/\partial B_i$ to $\partial X_{ai}/\partial B_i$ | 3 x 523       | 3 x 98        | 3 x 191        |
| construction of perturbed $t$ amplitudes | 3 x 68        | 3 x 7         | 3 x 15        |
| computation of occ.-occ.block of perturbed density matrix | 3 x 12        | 3 x 1         | 3 x 3         |
| computation of vrt.-vrt. block of perturbed density matrix | 3 x 55        | 3 x 10        | 3 x 25        |
| evaluation of the contribution due to $(ij|ka)$ to $\partial X_{ai}/\partial B_i$ | 3 x 58        | 3 x 4         | 3 x 8         |
| evaluation of the contribution due to $(ab|ci)$ to $\partial X_{ai}/\partial B_i$ | 3 x 264       | 3 x 49        | 3 x 97         |
| solution of perturbed Z-vector equations | 3 x 27        | 3 x 2         | 3 x 17        |
| total wall-clock time | 7215          | 1740          | 3446          |

$^a$ calculation has been performed using 1 CPU

$^b$ calculation has been performed using 20 CPUs; CPU time is given per node
tendency to overshoot\textsuperscript{17} and its performance deteriorates in cases where electron-correlation effects are large (see, for example, Ref. 4). Therefore, the second direction is an extension of the present scheme towards coupled-cluster methods,\textsuperscript{15–17} as it has been amply shown that highly accurate results are only achieved using these methods.\textsuperscript{5}

VI. DATA AVAILABILITY STATEMENT

The data that supports the findings of this study are available within the article and its supplementary material.

VII. SUPPLEMENTARY MATERIAL

See supplementary material for the details of the reported calculations (geometries in Cartesian Coordinates and basis sets) as well as the CD based GIAO-MP2 results.

ACKNOWLEDGMENTS

One of the authors (J.G.) thanks Professor Anna I. Krylov (University of Southern California, Los Angeles, USA) for introducing him to the Cholesky decomposition as well as for encouragement to implement Cholesky decomposition into the CFOUR program package and to pursue research along these lines. S.S. acknowledges support from the Deutsche Forschungsgemeinschaft via grant STO 1239/1-1.

1M. Kaupp, M. Bühl, and V. G. Malkin, eds., Calculation of NMR and EPR Parameters: Theory and Applications (Wiley-VCH, Weinheim, 2004).

2M. Bühl, J. Gauss, and P. v. R. Schleyer, “Decision electron correlation effects on computed \textsuperscript{14}B and \textsuperscript{13}C NMR chemical shifts. application of the GIAO-MP2 method to boranes and carboranes,” J. Am. Chem. Soc. 115, 12385–12390 (1993).

3J. Gauss, “Accurate calculation of NMR chemical shifts,” Ber. Bunsenges. Phys. Chem. 99, 1001–1008 (1995).

4J. F. Stanton, J. Gauss, and H. U. Siehl, “CCSD(T) calculation of NMR chemical shifts,” Adv. Chem. Phys. 123, 355–422 (2002).

5M. E. Harding, M. Lenhart, A. A. Auer, and J. Gauss, “Quantitative prediction of gas-phase \textsuperscript{19}F nuclear magnetic shielding constants,” J. Chem. Phys. 128, 244111 (2008).

6M. E. Harding, J. Gauss, and P. v. R. Schleyer, “Why benchmark-quality computations are needed to reproduce l-adamantyl cation NMR chemical shifts accurately,” J. Phys. Chem. A 115, 2340–2344 (2011).

7A. M. Teale, O. B. Luttrell, T. Helgaker, D. J. Tozer, and J. Gauss, “Benchmarking density-functional-theory calculations of NMR shielding constants and spin-rotation constants using accurate coupled-cluster calculations,” J. Chem. Phys. 138, 024111 (2013).

8M. Jaszunski, S. P. A. Sauer, R. Faber, and D. J. D. Wilson, “NMR parameters of FNNF as a test for coupled-cluster methods: CCSDT shielding and C3C spin–spin coupling,” Phys. Chem. Chem. Phys. 22, 21350–21359 (2020).

9C. Möller and M. S. Plesset, “Note on an approximation treatment for many-electron systems,” Phys. Rev. 46, 618–622 (1934).

10D. Cremer, “Møller–Plesset perturbation theory: From small molecule methods to methods for thousands of atoms,” WIREs Comp. Mol. Sci. 1, 509–530 (2000).

11J. Gauss, “Calculation of NMR chemical shifts at second-order many-body perturbation theory using gauge-including atomic orbitals,” Chem. Phys. Lett. 191, 614–620 (1992).

12J. Gauss, “Effects of electron correlation in the calculation of nuclear magnetic resonance chemical shifts,” J. Chem. Phys. 99, 3629–3643 (1993).

13S. Sieber, J. Gauss, and P. v. R. Schleyer, “The accurate C22, phenonium and benzenonium ion structures confirmed by correlated GIAO-MP2 NMR chemical shift calculations,” J. Am. Chem. Soc. 115, 6987–6988 (1993).

14J. Gauss and J. F. Stanton, “Gauge-invariant calculation of nuclear magnetic shielding constants at the coupled–cluster singles and doubles level,” J. Chem. Phys. 102, 251–253 (1995).

15J. Gauss and J. F. Stanton, “Coupled-cluster calculations of nuclear magnetic resonance chemical shifts,” J. Chem. Phys. 103, 3561–3578 (1995).

16J. Gauss and J. F. Stanton, “Perturbative treatment of triple excitations in coupled-cluster calculations of nuclear magnetic shielding constants,” J. Chem. Phys. 104, 2574–2583 (1996).

17D. Sundholm, A. Schäfer, and J. Gauss, “Rovibrationally averaged nuclear magnetic shielding tensors calculated at the coupled-cluster level,” J. Chem. Phys. 105, 11051–11059 (1996).

18C. Puzzarini, G. Cazzoli, M. E. Harding, J. Vázquez, and J. Gauss, “A new experimental absolute nuclear magnetic shielding scale for oxygen based on the rotational hyperfine structure of H$_2$O,” J. Chem. Phys. 131, 234304 (2009).

19I. Shawitt and R. J. Bartlett, Many-Body Methods in Chemistry and Physics: MBPT and Coupled-Cluster Theory (Cambridge University Press, 2009).

20M. Kollwitz and J. Gauss, “A direct implementation of the GIAO-MBPT(2) method for calculating NMR chemical shifts. Application to the naphthalenium and anthracenium ions,” Chem. Phys. Lett. 260, 639–646 (1996).

21M. Kollwitz, M. Häser, and J. Gauss, “Non-Abelian point group symmetry in direct second-order many-body perturbation theory calculations of NMR chemical shifts,” J. Chem. Phys. 108, 8295–8301 (1998).

22S. Loibl and M. Schütz, “NMR shielding tensors for density fitted local second-order Møller-Plesset perturbation theory using gauge including atomic orbitals,” J. Chem. Phys. 137, 084107 (2012).

23J. L. Whitten, “Coulombic potential energy integrals and approximations,” J. Chem. Phys. 58, 6496–6501 (1973).

24B. I. Dunlap, J. W. D. Connolly, and J. R. Sabín, “On some approximations in applications of X$_2$Q theory,” J. Chem. Phys. 71, 3396–3402 (1979).

25K. Eichkorn, O. Treutler, H. Öhm, M. Häser, and R. Ahlrichs, “Auxiliary basis sets to approximate Coulomb potentials,” Chem. Phys. Lett. 240, 283–290 (1995).

26F. London, “Théorie quantique des courants interatomiques dans les combinaisons aromatiques,” J. Phys. Radiois 8, 397–409 (1937).

27H. F. Hameka, “On the nuclear magnetic shielding in the hydrogen molecule,” Mol. Phys. 1, 203–215 (1958).

28R. Ditchfield, “Molecular orbital theory of magnetic shielding and magnetic susceptibility,” J. Chem. Phys. 56, 5688–5691 (1972).

29K. Wolinski, J. F. Hinton, and P. Puly, “Efficient implementation of the gauge-independent atomic orbital method for NMR chemical shift calculations,” J. Am. Chem. Soc. 112, 8251–8260 (1990).

30T. Helgaker and P. Jorgensen, “An electronic Hamiltonian for origin independent calculations of magnetic properties,” J. Chem. Phys. 95, 2595–2601 (1991).

31M. Maurer and C. Ochenseifeld, “A linear- and sub-linear-scaling method for calculating NMR shieldings in atomic orbital-based second-order Møller-Plesset perturbation theory,” J. Chem. Phys. 127, 174104 (2010).

32G. L. Stoychev, A. A. Auer, J. Gauss, and F. Neese, “DLPNO-PM2 second derivatives for the computation of polarizabilities and NMR shieldings,” J. Chem. Phys. 154, 164110 (2021).

33F. Neese, F. Wennmohs, U. Becker, and C. Riplinger, “The ORCA quantum chemistry program package,” J. Chem. Phys. 152, 224108 (2020).

34F. Neese, F. Wennmohs, and A. Hansen, “Efficient and accurate local approximations to coupled-electron pair approaches: An attempt to revive the pair natural orbital method,” J. Chem. Phys. 130, 114108 (2009).

35G. L. Stoychev, A. A. Auer, R. Izsák, and F. Neese, “Self-consistent field calculation of nuclear magnetic resonance chemical shielding constants using gauge-including atomic orbitals and approximate two-electron integrals,” J. Chem. Theory Comput. 14, 619–637 (2018).

36G. L. Stoychev, A. A. Auer, and F. Neese, “Efficient and accurate prediction of nuclear magnetic shielding tensors with double-hybrid density
an efficient and simple linear scaling local MP2 method that uses an inter-

temocratic infrastructure for reduced-scaling electronic structure methods. i.

approximation with tunable accuracy and efficiency,” J. Chem. Phys.

F. Aquilante, L. Gagliardi, T. B. Pedersen, and R. Lindh, “Atomic Cholesky

decomposition: Coulomb and exchange energies,” J. Chem. Phys.

L. Boman, H. Koch, and A. Sánchez de Merás, “Method specific Cholesky

Cholesky decompositions,” Chem. Phys. Lett.

duced opposite spin second-order Møller–Plesset correlation energy using

F. Aquilante and T. B. Pedersen, “Quartic scaling evaluation of canonical

basis sets from Cholesky decompositions,” Theor. Chem. Acc.

J. Boström, V. Veryazov, F. Aquilante, T. B. Pedersen, and R. Lindh, “Ana-

Cholesky decomposition of electron repulsion integrals,” J. Chem. Phys.

T. Nottoli, J. Gauss, and F. Lipparini, “A second-order CASSCF algorithm

with the Cholesky decomposition of the two-electron repulsion integrals,”

J. Chem. Comput. xx, submitted (2021).

F. Aquilante, R. Lindh, and T. B. Pedersen, “Unbiased auxiliary basis sets for

accurate two-electron integral approximations,” J. Chem. Phys. 127,

114107 (2007).

F. Aquilante and T. B. Pedersen, “Quartic scaling evaluation of canonical

scaled opposite spin second-order Møller–Plesset correlation energy using

Cholesky decompositions,” Chem. Phys. Lett. 449, 354–357 (2007).

L. Boman, H. Koch, and A. Sánchez de Merás, “Method specific Cholesky

decomposition: Coulomb and exchange energies,” J. Chem. Phys. 129,

134107 (2008).

F. Aquilante, T. B. Pedersen, and R. Lindh, “Density fitting with auxiliary

basis sets from Cholesky decompositions,” Theor. Chem. Acc. 124, 1–10

(2009).

F. Aquilante, L. Gagliardi, T. B. Pedersen, and R. Lindh, “Atomic Cholesky

decompositions: A route to unbiased auxiliary basis sets for density fitting

approximation with tunable accuracy and efficiency,” J. Chem. Phys. 130,

154107 (2009).

S. D. Folkestad, E. F. Kjønstad, and H. Koch, “An efficient algorithm for

Cholesky decomposition of electron repulsion integrals,” J. Chem. Phys.

150, 194112 (2019).

J. Boström, V. Veryazov, F. Aquilante, T. B. Pedersen, and R. Lindh, “Ana-

lytical gradients of the second-order Møller–Plesset energy using Cholesky

decompositions,” Int. J. Quantum Chem. 114, 321–327 (2014).

X. Feng, E. Epifanovsky, J. Gauss, and A. I. Krylov, “Implementation of

analytic gradients for CCSD and EOM-CCSD using Cholesky decompo-

sition of the electron-repulsion integrals and their derivatives: Theory and

benchmarks,” J. Chem. Phys. 151, 014110 (2019).

J. F. Stanton, J. Gauss, L. Cheng, M. E. Harding, D. A. Matthews, and P. G.

Szalay, “CFOUR, Coupled-Cluster techniques for Computational Chem-

istry, a quantum-chemical program package,” With contributions from A.A.

Auer, R.J. Bartlett, U. Benedikt, C. Berger, D.E. Bernholdt, S. Blaschke,

Y.I. Bonhle, S. Burger, O. Christiansen, D. Datta, F. Engel, R. Faber, J.

Greiner, M. Heckert, O. Heun, M. Hilgenberg, C. Huber, T.-C. Jagau, D.

Jonson, J. Jusélius, T. Kirsch, K. Klein, G.M. Kopper, W.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. Schöflimann, W. Schwalbach, C.

Simmons, S. Stopkowicz, A. Tajti, J. Vázquez, F. Wang, J.D. Watts and the

integral packages MOLECULE (J. Almlof 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 Wullen. For the current

version, see http://www.cfour.de.

D. A. Matthews, L. Cheng, M. E. Harding, F. Lipparini, S. Stopkowicz,

T. Jagau, P. G. Szalay, J. Gauss, and J. F. Stanton, “Coupled-cluster tech-

niques for computational chemistry: The CFOUR program package,” J.

Chem. Phys. 152, 214108 (2020).

N. C. Handy and H. F. Schaefer III, “On the evaluation of analytic energy
derivatives for correlated wave functions,” J. Chem. Phys. 81, 5031 (1984).

J. Gerratt and I. M. Mills, “Force constants and dipole-moment derivatives

of molecules from perturbed Hartree–Fock calculations. I.” J. Chem. Phys.

49, 1719 (1968).

A. Pople, R. Krishnan, H. B. Schlegel, and J. S. Binkley, “Derivative

studies in Hartree-Fock and Møller-Plesset theories,” Int. J. Quantum Chem.

Symp. 13, 225–241 (1979).

Note that we define the two-electron integral matrix with the indices for

electron 2 interchanged in order to ensure positive semi-definiteness of the

matrix in the presence of an external magnetic field, see Ref. ?.

OpenMP Architecture Review Board, “OpenMP application program inter-

face version 4.5.,” (2015), last accessed 30 March 2021.

J. Gauss, F. Lipparini, S. Burger, S. Blaschke, M.-P. Kitsaras, and S. Stop-

kowicz, (2021), Johannes Gutenberg-Universität Mainz, unpublished.

L. E. McMurchie and E. R. Davidson, “One- and two-electron integrals

over Cartesian Gaussian functions,” J. Comput. Phys. 26, 218–231 (1978).

T. H. Dunning, Jr., “Gaussian basis sets for use in correlated molecular

calculations. I. The atoms boron through neon and hydrogen,” J. Chem.

Phys. 90, 1007–1023 (1989).

J. Gauss, Z. W. Schneider, and H. Schnöckel, “?Al NMR spectroscopic investigation of aluminum(III) compounds: Ab initio calculations and experiment,” J. Am. Chem. Soc. 115, 2402–2408 (1993).

P. Brown, T. Schaller, U. P. Saalbach, F. Koziol, C. Ochsenfeld, F.-G.

Krämer, and H. W. Spiess, “Structure assignment in the solid state by the coupling of quantum chemical calculations with NMR experiments: A columnar hexabenzocoronene derivative,” J. Am. Chem. Soc. 123, 2597–2606 (2001).

M. Häsär, J. Almlof, and G. E. Scuseria, “The equilibrium geometry of CsO as predicted by 2nd-order (MP2) perturbation-theory,” Chem. Phys. Lett. 181, 497–500 (1991).

J. Gauss, U. Schneider, R. Ahlricks, C. Dohmene, and H. Schnöckel, “2Al NMR spectroscopic investigation of aluminum(III) compounds: Ab initio calculations and experiment,” J. Am. Chem. Soc. 115, 2571–2577 (1992).

J. Gauss and H.-J. Werner, “NMR chemical shift calculations within local correlation methods: the GIAO-LMP2 approach,” Phys. Chem. Chem. Phys. 2, 2083–2090 (2000).

P. B. Karadakov and K. Morokuma, “ONIOM as an efficient tool for calcu-

lating NMR chemical shielding constants in large molecules,” Chem. Phys.

Lett. 317, 589–596 (2000).

A. Schäfer, H. Horn, and R. Ahlricks, “Fully optimized contracted Gaussian

basis sets for atoms Li to Kr,” J. Chem. Phys. 97, 2571–2577 (1992).

J. Gauss and H.-J. Werner, “NMR chemical shift calculations within local correlation methods: The GIAO-LMP2 approach,” Phys. Chem. Chem. Phys. 2, 2083–2090 (2000).

F. Lipparini, C. Cappelli, and V. Barone, “A gauge invariant multiscale ap-

proach to magnetic spectroscopies in condensed phase: General three-layer

model, computational implementation and pilot applications,” J. Chem.

Phys. 138, 234108 (2013).

L. W. Chung, W. M. C. Sameera, R. Ramozzi, A. J. Page, M. Hatanaka, G. P. Petrova, T. V. Harris, X. Li, Z. K., F. Liu, H.-B. Li, L. Ding, and K. Morokuma, “The ONIOM method and its applications,” Chem. Rev. 115, 5678–5796 (2015).

X. Jin, T. Zhu, J. Z. H. Zhang, and X. He, “A systematic study on RNA

NMR chemical shift calculation based on the automated fragmentation

QM/MM approach,” RSC Adv. 6, 108890–108602 (2016).

S. Hashemi, L. Cupellini, F. Lipparini, and B. Mennucci, “A polarisable

QM/MM description of NMR chemical shifts of a photoreceptor protein,”

Mol. Phys. 118, e1771449 (2020).