

# Tensor-Hypercontracted MP2 First Derivatives: Runtime and Memory Efficient Computation of Hyperfine Coupling Constants

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



Cite This: *J. Chem. Theory Comput.* 2022, 18, 5233–5245



Read Online

ACCESS |



Metrics & More



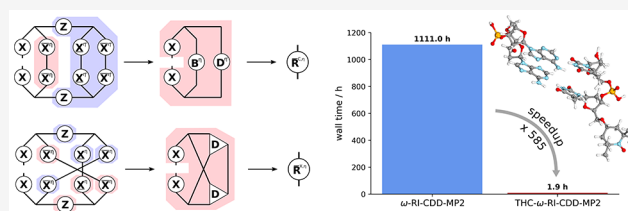
Article Recommendations



Supporting Information

**ABSTRACT:** We employ our recently introduced tensor-hypercontracted (THC) second-order Møller–Plesset perturbation theory (MP2) method [Bangerter, F. H., Glasbrenner, M., Ochsenfeld, C. *J. Chem. Theory Comput.* 2021, 17, 211–221] for the computation of hyperfine coupling constants (HFCCs). The implementation leverages the tensor structure of the THC factorized electron repulsion integrals for an efficient formation of the integral-based intermediates. The computational complexity

of the most expensive and formally quintic scaling exchange-like contribution is reduced to effectively subquadratic, by making use of the intrinsic, exponentially decaying coupling between tensor indices through screening based on natural blocking. Overall, this yields an effective subquadratic scaling with a low prefactor for the presented THC-based AO-MP2 method for the computation of isotropic HFCCs on DNA fragments with up to 500 atoms and 5000 basis functions. Furthermore, the implementation achieves considerable speedups with up to a factor of roughly 600–1000 compared to previous implementations [Vogler, S., Ludwig, M., Maurer, M., Ochsenfeld, C. *J. Chem. Phys.* 2017, 147, 024101] for medium-sized organic radicals, while also significantly reducing storage requirements.



## 1. INTRODUCTION

Ever since the advent of modern computers in the 1990s, Møller–Plesset perturbation theory (MP $n$ )<sup>1</sup> has been a good compromise in the family of quantum chemical methods, being sufficiently accurate for many applications while still being computationally affordable.<sup>2</sup> As opposed to coupled cluster theory (CC), MP $n$  lacks infinite-order corrections present in the cluster operator expansion of the CC models, which generally makes MP $n$  less accurate.<sup>2</sup> However, when going from energies to gradients and molecular properties, MP $n$ , especially second-order MP $n$  (MP2), was shown to yield accurate hyperfine coupling constants (HFCCs)<sup>3–7</sup> and relative nuclear magnetic resonance (NMR) shifts.<sup>8–12</sup> However, MP2 is sensitive to spin-contamination in the Hartree–Fock wave function, which can be improved upon when used in its orbital-optimized variant<sup>4</sup> or as part of double-hybrid density functionals.<sup>13</sup>

Furthermore, when comparing MP $n$  and CC at the same expansion orders, for example, MP2 and singles and doubles CC (CCSD), MP $n$  comes with a scaling advantage, both in the prefactor and the scaling exponent. Nonetheless, canonical MP2, as well as the associated first and second derivative,<sup>6,11</sup> still scale with the fifth power of the molecule size, thereby severely restricting the accessible chemical space. To alleviate this limitation several formulations of the MP2 derivatives have been proposed.

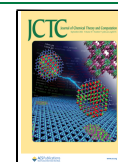
Early work from Pulay and Sæbø<sup>14,15</sup> on local correlation was applied to the computation of MP2 gradients.<sup>16,17</sup> Following up on this, in recent years the domain-based local pair natural orbital (DLPNO) formulation of MP2 by Neese and co-

workers<sup>18–20</sup> was extended to first and second derivatives. Conceptually related is the divide-expand-consolidate (DEC) formulation of MP2 by Jørgensen and co-workers,<sup>21,22</sup> which likewise was extended to the computation of molecular gradients in a linear-scaling and massively parallel manner.<sup>23,24</sup> Instead of exploiting locality in the correlation space, the equations of the MP2 first<sup>6,25</sup> and second derivative<sup>11</sup> can be reformulated entirely in terms of atomic orbitals (AOs). However, for an efficient implementation and the reduction of the scaling prefactor, an orbital localization by pivoted Cholesky decomposition (PCD) of the associated pseudodensities is essential.<sup>6</sup>

Besides ensuring low-scaling and efficiency, for derivative calculations of electron correlation methods in general, it is pivotal to efficiently manage the available memory and disk space. Compared to the energy equations, the associated gradients and higher derivatives often not only include electron repulsion integrals (ERIs), in either atomic or molecular orbital basis, but can also include partially transformed ERIs and derivatives thereof, for example, with respect to the magnetic field in NMR calculations. Since canonical ERIs are fourth-order tensors, their memory requirements prohibitively scale with the

Received: February 3, 2022

Published: August 9, 2022



forth power of the number of basis functions and saving multiple such tensors quickly becomes unfeasible. To reduce the memory footprint of the ERIs, tensor decomposition methods, particularly the resolution-of-the-identity (RI) ansatz,<sup>26–29</sup> have been broadly applied in the context of MP2 derivatives.<sup>5,6,12,30</sup> However, with increasing molecule size, even the third-order RI tensors eventually exceed the available disk space of conventional high performance computing nodes. To overcome this storage limitation, further reduction of the dimensionality of the ERIs is desirable. This can be achieved by the recently introduced tensor hypercontraction (THC) factorization of Martínez and co-workers,<sup>31–34</sup> which in general terms approximates a fourth-order integral tensor ( $\mu\nu|\delta|\lambda\sigma$ ), where  $\delta$  is a singular two electron interaction kernel, by five second-order tensors.<sup>35</sup> In the least-squares formulation of THC (LS-THC),<sup>31</sup> four of these tensors are simply obtained by evaluation of the basis functions at real-space grid nodes and the singular  $\delta$  operator is replaced by the LS-fitted  $\mathbf{Z}$  matrix. If  $\delta$  is the Coulombic  $1/r$  operator, a factorization of the regular ERIs is achieved, which has been employed in reduced scaling formulations of exact exchange,<sup>36</sup> different orders of MP $n$ ,<sup>31,33–41</sup> the random phase approximation (RPA),<sup>42</sup> complete active space perturbation theory (CASPT2),<sup>43</sup> and various flavors of CC theory,<sup>44–46</sup> as well as equation-of-motion CC (EOM-CC) theory.<sup>47</sup> Recently Matthews<sup>35</sup> thoroughly investigated amplitude factorizations within MP3, as a stepping stone toward CCSD, and noted that the LS-THC factorization of nonlocal integrals, such as the exchange integrals, incurs an additional error.

The applicability of THC for the computation of molecular gradients is largely unexplored. Song et al.<sup>39</sup> derived equations for the analytical gradient of THC-AO-MP2 with application to geometry optimizations and *ab initio* molecular dynamics (AIMD) simulations. As commonly the derivation of gradient equations of electron correlation methods is rather involved, Song et al.<sup>48</sup> also proposed an automatic differentiation scheme for the automated generation of working equations for gradients of THC-based correlation methods.

When it comes to applying tensor factorizations, be it RI or THC, two possible routes to molecular gradients can be taken: The first approach is to differentiate the RI- or THC-approximated energy equation, then the associated gradient describes the slope of an approximated potential energy surface (PES). The second approach is to take the gradient of the canonical energy and insert the approximation into the exact gradient; this way an approximate gradient is used to move along the exact PES.

The latter can lead to an unwanted buildup of errors during the course of a simulation, when used in molecular dynamics simulations. However, when used in conjunction with thermostats constant energy is traded for constant temperature, and depending on the extent of the gradient error, this approach can still be applicable but must be tested for the chosen gradient. The first approach was taken by Song et al.<sup>39</sup> for their THC-based MP2 molecular gradient, whereas in the present work, the gradient equations with respect to a perturbation, that the basis functions are independent of, are derived by inserting the THC factorization into the equations of the exact gradient. As will be discussed in section 2.4, the resulting equations are identical to the ones obtained by differentiating the THC-AO-MP2 energy equation. As a representative case of these kinds of perturbations, we apply our recently developed low-scaling LS-THC algorithm<sup>41</sup> to the AO-MP2 energy derivative with

respect to the nuclear magnetic moment, for the computation of HFCCs on the MP2 level of theory.

We present ways to efficiently treat the Coulomb- and the exchange-like part of the most expensive intermediate of the derivative within the LS-THC approximation. We demonstrate the low-scaling behavior of our THC- $\omega$ -RI-CDD-MP2 derivative method for various chemically relevant systems. We also show that THC- $\omega$ -RI-CDD-MP2 significantly outperforms our previous implementation of  $\omega$ -RI-CDD-MP2<sup>6,30</sup> for the computation of isotropic HFCCs.

## 2. THEORY

**2.1. Notation.** Throughout this publication we make use of the following indices:

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

**2.2. Review of the AO-MP2 Gradient.** The unrestricted AO-MP2 energy with a Laplace transformation<sup>49–51</sup> for the energy denominator is given as

$$E_{\text{AO-MP2}} = -\frac{1}{2} \sum_{\kappa} \omega_{\kappa} \sum_{\eta \in \{\alpha, \beta\}} \mathcal{E}^{\eta}(\kappa) \quad (1)$$

with

$$\mathcal{E}^{\eta}(\kappa) = \sum_{\mu\nu\lambda\sigma} (\underline{\mu\nu}|\underline{\lambda\sigma})_{\eta\eta} (\underline{\mu\nu}||\lambda\sigma) + \sum_{\mu\nu\lambda\sigma} (\underline{\mu\nu}|\underline{\lambda\sigma})_{\eta\eta'} (\underline{\mu\nu}|\lambda\sigma) \quad (2)$$

for which the transformed ERIs are given by

$$(\underline{\mu\nu}|\underline{\lambda\sigma})_{\eta\eta'} = \sum_{\mu'\nu'\lambda'\sigma'} \underline{P}_{\mu'\mu}^{\eta} \bar{P}_{\nu'\nu}^{\eta'} \underline{P}_{\lambda'\lambda}^{\eta'} \bar{P}_{\sigma'\sigma}^{\eta} (\mu'\nu'|\lambda'\sigma') \quad (3)$$

and  $\underline{P}$  and  $\bar{P}$  are the usual pseudodensities, given by

$$\begin{aligned} \underline{P}_{\mu'\mu}^{\eta} &= \sum_i C_{\mu'i}^{\eta} e^{\varepsilon_i^{\eta}(\xi)} C_{\mu'i}^{\eta} \\ \bar{P}_{\nu'\nu}^{\eta'} &= \sum_a C_{\nu'a}^{\eta'} e^{-\varepsilon_a^{\eta'}(\xi)} C_{\nu'a}^{\eta'} \end{aligned} \quad (4)$$

To obtain the AO-MP2 gradient, eq 1 has to be differentiated with respect to a perturbation  $\xi$ . Since the focus of this work is on HFCCs, as an example for a property for which the basis functions are independent of the perturbation—here  $\xi'$ —the following derivation is restricted to this special case. An in-depth derivation of the AO-MP2 gradient equations, as well as a comparison to the MO-MP2 gradient, is available in refs 6 and 25.

In order to obtain the gradient of the AO-MP2 energy, eq 1 is differentiated with respect to  $\xi'$

$$\frac{\partial E_{\text{AO-MP2}}}{\partial \xi'} = -\frac{1}{2} \sum_{\kappa} \omega_{\kappa} \sum_{\eta \in \{\alpha, \beta\}} \frac{\partial \mathcal{E}^{\eta}(\kappa)}{\partial \xi'} \quad (5)$$

with

$$\frac{\partial \mathcal{E}^{\eta}(\kappa)}{\partial \xi'} = 2 \sum_{\mu' \mu} \bar{\mathbf{R}}_{\mu' \mu}^{\eta}(\kappa) \frac{\partial \mathbf{P}_{\mu' \mu}^{\eta}}{\partial \xi'} + 2 \sum_{\nu' \nu} \mathbf{R}_{\nu' \nu}^{\eta}(\kappa) \frac{\partial \bar{\mathbf{P}}_{\nu' \nu}^{\eta}}{\partial \xi'} \quad (6)$$

and intermediates  $\mathbf{R}$  and  $\bar{\mathbf{R}}$  given by

$$\begin{aligned} \bar{\mathbf{R}}_{\mu' \mu}^{\eta}(\kappa) &= \sum_{\nu \lambda \sigma} (\mu' \nu | \lambda \bar{\sigma})_{\eta} (\mu \nu | \lambda \sigma) + \sum_{\nu \lambda \sigma} (\mu' \nu | \lambda \bar{\sigma})_{\eta} (\mu \nu | \lambda \sigma) \\ \mathbf{R}_{\nu' \nu}^{\eta}(\kappa) &= \sum_{\mu \lambda \sigma} (\mu \nu' | \lambda \bar{\sigma})_{\eta} (\mu \nu | \lambda \sigma) + \sum_{\mu \lambda \sigma} (\mu \nu' | \lambda \bar{\sigma})_{\eta} (\mu \nu | \lambda \sigma) \end{aligned} \quad (7)$$

The above  $\mathbf{R}$  intermediates can be thought of as the contraction of all perturbation-independent parts of the gradient. Note that eq 5 only involves the derivative of the pseudodensities and no integral derivatives, as the basis functions are taken to be independent of  $\xi'$ . In order to avoid the evaluation of the derivatives of the perturbed occupied and virtual pseudodensities, these intermediates are expanded in terms of the regular occupied and virtual densities  $\mathbf{P}_{\text{occ}}$  and  $\mathbf{P}_{\text{virt}}$  as

$$\begin{aligned} \mathbf{P}_{\mu' \mu}^{\eta} &= (e^{\tau_{\kappa} \mathbf{P}_{\text{occ}}^{\eta} \mathbf{F}^{\eta}} \mathbf{P}_{\text{occ}}^{\eta})_{\mu' \mu} \\ \bar{\mathbf{P}}_{\nu' \nu}^{\eta} &= (e^{-\tau_{\kappa} \mathbf{P}_{\text{virt}}^{\eta} \mathbf{F}^{\eta}} \mathbf{P}_{\text{virt}}^{\eta})_{\nu' \nu} \end{aligned} \quad (8)$$

Thus, the derivatives of eq 8 are given by

$$\begin{aligned} \frac{\partial \mathbf{P}_{\mu' \mu}^{\eta}}{\partial \xi'} &= \frac{\partial e^{\tau_{\kappa} \mathbf{P}_{\text{occ}}^{\eta} \mathbf{F}^{\eta}}}{\partial \xi'} \mathbf{P}_{\text{occ}}^{\eta} + e^{\tau_{\kappa} \mathbf{P}_{\text{occ}}^{\eta} \mathbf{F}^{\eta}} \frac{\partial \mathbf{P}_{\text{occ}}^{\eta}}{\partial \xi'} \\ \frac{\partial \bar{\mathbf{P}}_{\nu' \nu}^{\eta}}{\partial \xi'} &= \frac{\partial e^{-\tau_{\kappa} \mathbf{P}_{\text{virt}}^{\eta} \mathbf{F}^{\eta}}}{\partial \xi'} \mathbf{P}_{\text{virt}}^{\eta} + e^{-\tau_{\kappa} \mathbf{P}_{\text{virt}}^{\eta} \mathbf{F}^{\eta}} \frac{\partial \mathbf{P}_{\text{virt}}^{\eta}}{\partial \xi'} \end{aligned} \quad (9)$$

By further making use of the identity

$$\mathbf{P}_{\text{occ}}^{\eta} \mathbf{S} + \mathbf{P}_{\text{virt}}^{\eta} \mathbf{S} = \mathbf{I} \quad (10)$$

the perturbed virtual density can be related to the perturbed occupied density as

$$\frac{\partial \mathbf{P}_{\text{virt}}^{\eta}}{\partial \xi'} = -\frac{\partial \mathbf{P}_{\text{occ}}^{\eta}}{\partial \xi'} \quad (11)$$

Note again that eq 11 does not contain the derivative of the overlap matrix  $\mathbf{S}$ , due to  $\mathbf{S}$  being independent of  $\xi'$ . By making use of the above relations for the densities, eq 6 becomes

$$\begin{aligned} \frac{\partial \mathcal{E}^{\eta}(\kappa)}{\partial \xi'} &= 2 \text{tr} \left[ \bar{\mathbf{R}}^{\eta} e^{\tau_{\kappa} \mathbf{P}_{\text{occ}}^{\eta} \mathbf{F}^{\eta}} \frac{\partial \mathbf{P}_{\text{occ}}^{\eta}}{\partial \xi'} \right] + 2 \text{tr} \left[ \mathbf{P}_{\text{occ}}^{\eta} \bar{\mathbf{R}}^{\eta} \frac{\partial e^{\tau_{\kappa} \mathbf{P}_{\text{occ}}^{\eta} \mathbf{F}^{\eta}}}{\partial \xi'} \right] \\ &\quad - 2 \text{tr} \left[ \bar{\mathbf{R}}^{\eta} e^{-\tau_{\kappa} \mathbf{P}_{\text{virt}}^{\eta} \mathbf{F}^{\eta}} \frac{\partial \mathbf{P}_{\text{occ}}^{\eta}}{\partial \xi'} \right] \\ &\quad + 2 \text{tr} \left[ \mathbf{P}_{\text{virt}}^{\eta} \bar{\mathbf{R}}^{\eta} \frac{\partial e^{-\tau_{\kappa} \mathbf{P}_{\text{virt}}^{\eta} \mathbf{F}^{\eta}}}{\partial \xi'} \right] \end{aligned} \quad (12)$$

Cyclic permutation under the trace was applied above to obtain terms of the general form

$$\text{tr} \left[ \mathbf{B} \frac{\partial e^{\mathbf{A}}}{\partial \xi'} \right] = \text{tr} \left[ \mathbf{Y} \frac{\partial \mathbf{A}}{\partial \xi'} \right] \quad (13)$$

with

$$\mathbf{Y} = \sum_{k=1}^{\infty} \sum_{i=0}^{k-1} \frac{1}{k!} \mathbf{A}^i \mathbf{B} \mathbf{A}^{k-i-1} \quad (14)$$

which can be solved for  $\mathbf{Y}$  by recursion, as detailed in refs 25 and 12. Let  $\bar{\mathbf{Y}}^{\eta}$  be the solution of eq 13 with  $\mathbf{A} \equiv \tau_{\kappa} \mathbf{P}_{\text{occ}}^{\eta} \mathbf{F}^{\eta}$  and  $\mathbf{B} \equiv \mathbf{P}_{\text{occ}}^{\eta} \bar{\mathbf{R}}^{\eta}$ , and let  $\mathbf{Y}^{\eta}$  be the solution with  $\mathbf{A} \equiv -\tau_{\kappa} \mathbf{P}_{\text{virt}}^{\eta} \mathbf{F}^{\eta}$  and  $\mathbf{B} \equiv \mathbf{P}_{\text{virt}}^{\eta} \bar{\mathbf{R}}^{\eta}$ . Then, by making use of the other transformations outlined above, the derivative from eq 6 can be rewritten as

$$\frac{\partial \mathcal{E}^{\eta}(\kappa)}{\partial \xi'} = 2 \text{tr} \left[ \mathcal{F}^{\eta}(\kappa) \frac{\partial \mathbf{h}}{\partial \xi'} \right] + 2 \text{tr} \left[ \mathcal{P}^{\eta}(\kappa) \frac{\partial \mathbf{P}_{\text{occ}}^{\eta}}{\partial \xi'} \right] \quad (15)$$

with

$$\begin{aligned} \mathcal{F}^{\eta}(\kappa) &= \tau_{\kappa} \bar{\mathbf{Y}}^{\eta} \mathbf{P}_{\text{occ}}^{\eta} - \tau_{\kappa} \mathbf{Y}^{\eta} \mathbf{P}_{\text{virt}}^{\eta} \\ \mathcal{P}^{\eta}(\kappa) &= \tau_{\kappa} \mathbf{F}^{\eta} \bar{\mathbf{Y}}^{\eta} + \tau_{\kappa} \mathbf{F}^{\eta} \mathbf{Y}^{\eta} + \mathbf{G}[\mathcal{F}^{\eta}(\kappa)] + \bar{\mathbf{R}}^{\eta} e^{\tau_{\kappa} \mathbf{P}_{\text{occ}}^{\eta} \mathbf{F}^{\eta}} \\ &\quad - \bar{\mathbf{R}}^{\eta} e^{-\tau_{\kappa} \mathbf{P}_{\text{virt}}^{\eta} \mathbf{F}^{\eta}} \end{aligned} \quad (16)$$

where intermediate  $\mathbf{G}[\mathcal{F}^{\eta}(\kappa)]$  resembles the Fock matrix with  $\mathcal{F}^{\eta}(\kappa)$  substituting for the density matrix; that is,

$$\mathbf{G}_{\mu\nu}[\mathcal{F}^{\eta}(\kappa)] = \sum_{\lambda\sigma} \mathcal{F}_{\lambda\sigma}^{\eta}(\kappa) [2(\mu\nu|\lambda\sigma) - (\mu\sigma|\lambda\nu)] \quad (17)$$

Note that the first term in eq 15 only includes the derivative of the core Hamiltonian matrix as the integrals in the Fock matrix, from which this term originates from, are independent of  $\xi'$ .

Equation 15 permits an elegant solution for the perturbed density, avoiding the need to solve coupled-perturbed self-consistent field (CPSCF) equations for all perturbations  $\xi'$ , by means of applying a AO-based Z-vector-like method<sup>25</sup> originally proposed by Handy and Schaefer.<sup>52</sup> The implicit first derivative of the occupied density can therefore be efficiently obtained by applying the density matrix-based Laplace-transform unrestricted CPSCF (DL-UCPSCF) method by Beer and Ochsenfeld.<sup>53</sup> The intricacies of this method are detailed in refs 53, 25, and 6.

To conclude, eqs 5 and 15 yield the first derivative of the AO-MP2 energy with respect to a perturbation  $\xi'$ . More specifically, if  $\xi'$  was an external electric field then eq 5 would yield permanent dipole moments, and if  $\xi'$  was equal to the nuclear magnetic moment  $M_k$  of a given nucleus  $k$  then the isotropic contribution to the HFCC of nucleus  $k$  in the absence of spin-orbit coupling would be obtained. The latter property will be used as a sample property for the newly developed THC- $\omega$ -RI-CDD-MP2 derivative method presented in section 2.4.

**2.3. RI-CDD-MP2 HFCCs.** The computational bottleneck of obtaining the AO-MP2 gradient in eqs 5 and 15 are the integral contractions in the formation of the  $\mathbf{R}$  matrix intermediates given by eq 7. Because forming the  $\mathbf{R}$ -matrices involves the same contraction with the pseudodensities as the AO-MP2 energy, the computation of the gradient will *a priori* also have quintic scaling. As is common practice when dealing with these kinds of integral contractions, the RI approximation can be inserted into eq 7 to lower the computational cost as well as the memory requirements by avoiding the fourth-order ERI tensors. To further lower the prefactor of the integral transformations a PCD of the pseudodensities can be used, which is known as the Cholesky-decomposed pseudodensity (CDD) approach. The CDD method produces a set of so-called Cholesky pseudo-MO coefficient matrices  $\bar{\mathbf{L}}$  and  $\bar{\mathbf{L}}$  according to

$$\begin{aligned}
 P_{\mu'\mu}^{\eta} &= \sum_i^{\text{rk}(\mathbf{P}_{\text{occ}})} L_{\mu'i}^{\eta} \bar{L}_{\mu i}^{\eta} \\
 \bar{P}_{\nu\nu}^{\eta} &= \sum_a^{\text{rk}(\mathbf{P}_{\text{virt}})} \bar{L}_{\nu'a}^{\eta} \bar{L}_{\nu a}^{\eta}
 \end{aligned}
 \quad (18)$$

which are used analogously to the regular MO coefficients. The combined RI-CDD approach leads to a reformulation of the integrals incorporated in the  $\mathbf{R}$  intermediates as

$$(\underline{i}\bar{a}|\underline{j}\bar{b})_{\eta\eta'} = \sum_{\alpha} B_{\underline{i}\bar{a},\alpha}^{\eta} B_{\underline{j}\bar{b},\alpha}^{\eta'} \quad (19)$$

with

$$B_{\underline{i}\bar{a},\alpha}^{\eta} = \sum_{\beta} \sum_{\mu\nu} L_{\mu i}^{\eta} \bar{L}_{\nu a}^{\eta} (\mu\nu|\beta) (\beta|\alpha)^{-1/2} \quad (20)$$

In combination with QQR-type integral screening the RI-CDD-MP2 gradient method was shown to yield cubic scaling for the computation of molecular gradients and HFCCs.<sup>6</sup> To further reduce the scaling, Vogler et al.<sup>30</sup> employed the attenuated Coulomb metric<sup>54,55</sup> in the RI approximation as well as the scaled-opposite spin (SOS)<sup>56</sup> approximation, which removes the same spin contribution entirely. Still, the expensive formation of the  $\mathbf{R}$  intermediates has to be done for every Laplace point and thus constitutes the predominant part of the wall time for the evaluation of the MP2 gradient with respect to a perturbation  $\xi'$ , even with the  $\omega$ -RI-CDD-SOS-MP2 method.<sup>6,30</sup>

**2.4. THC-CDD-MP2 HFCCs.** ERIs are ubiquitous in electron correlation methods and their transformation and contraction usually represents the bottleneck of the calculation. This is especially the case when many different ERIs, that is, fully and partially transformed into the MO space or contracted with a perturbed density matrix, are needed, and their formation has to be carried out repeatedly, such as inside a Laplace expansion or during the iterative solution of amplitude equations. Since the scaling behavior of these operations is dependent on the dimensionality of the representation of the ERI tensor, a most compact representation is desirable. Of particular interest is thus the THC factorization, which in its AO formulation approximates an ERI as

$$(\mu\nu|\lambda\sigma) \approx \sum_{PQ} X_{\mu}^P X_{\nu}^P Z^{PQ} X_{\lambda}^Q X_{\sigma}^Q \quad (21)$$

and—in LS-THC—the  $\mathbf{X}$ -matrices are simply obtained by evaluation of basis functions at the THC grid.<sup>31</sup> The analytical expression of the  $\mathbf{Z}$ -matrix can be shown to be the solution to the normal equations associated with the least-squares equation of finding the THC factorization (see the Supporting Information). As has been shown previously,<sup>33,34,36,39,41–43</sup> the THC factorization can achieve major savings in computation time for intermediates involving ERI contractions, by reducing the representation of the ERIs to only second-order tensors. In this work, we make use of our recently reported low-scaling THC method<sup>41</sup> based on the  $\omega$ -RI approximation for the ERIs contained in  $\mathbf{Z}$  and natural blocking (NB).<sup>57,58</sup> However, in contrast to our work on THC-MP2 energies,<sup>41</sup> in the present work the AO ERIs are fitted, since the gradient for calculating the MP2 HFCCs is based on our RI-CDD approach to the computation of AO-MP2 energy gradients.<sup>6</sup> It is important to note here, that while the equations for the THC-based gradient

method are derived by inserting the THC factorization into the equations of the RI-CDD-MP2 gradient with respect to  $\xi'$ , there is no difference to directly differentiating the THC-AO-MP2 energy equations. This is the case, because neither  $\mathbf{X}$  nor  $\mathbf{Z}$  depend on the perturbation  $\xi'$  in the AO-THC formulation. This independence would not be given if either the MO-THC approach was used or for the more general derivative with respect to  $\xi$ , which would necessitate the derivatives of the THC tensors. In other words, by simply inserting the THC factorization into eq 7 the derivative of the THC- $\omega$ -RI-CDD-MP2 energy with respect to  $\xi'$  is obtained.

To reduce the computation time needed for forming the expensive  $\mathbf{R}$ -matrices of the UMP2 gradient, the AO-THC factorization is inserted into eq 7 in its RI-CDD formulation to yield

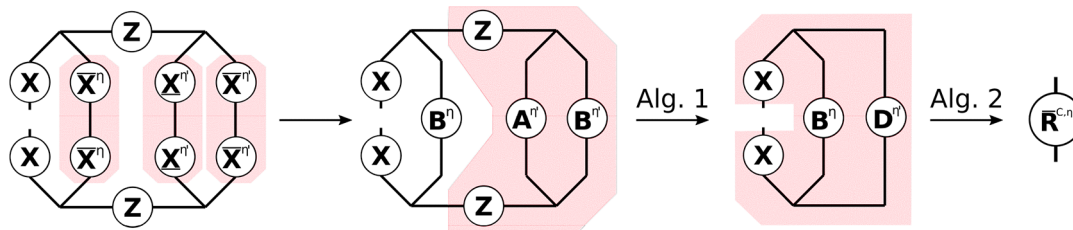
$$\begin{aligned}
 \bar{\mathbf{R}}_{\mu'\mu}^{\eta}(\kappa) &= \sum_j \sum_{ab} \sum_{PQRS} X_{\mu}^P \bar{X}_a^{P,\eta} Z^{PQ} \underline{X}_j^{Q,\eta} \bar{X}_b^{Q,\eta} \\
 &\quad [X_{\mu}^R \bar{X}_a^{R,\eta} Z^{RS} \underline{X}_j^{S,\eta} \bar{X}_b^{S,\eta} - X_{\mu}^R \bar{X}_b^{R,\eta} Z^{RS} \underline{X}_j^{S,\eta} \bar{X}_a^{S,\eta}] \\
 &\quad + \sum_j \sum_{ab} \sum_{PQRS} X_{\mu}^P \bar{X}_a^{P,\eta} Z^{PQ} \underline{X}_j^{Q,\eta'} \bar{X}_b^{Q,\eta'} \\
 &\quad X_{\mu}^R \bar{X}_a^{R,\eta} Z^{RS} \underline{X}_j^{S,\eta'} \bar{X}_b^{S,\eta'} \\
 \mathbf{R}_{\nu'\nu}^{\eta}(\kappa) &= \sum_{ij} \sum_b \sum_{PQRS} \underline{X}_i^{P,\eta} X_{\nu}^P Z^{PQ} \underline{X}_j^{Q,\eta} \bar{X}_b^{Q,\eta} \\
 &\quad [\underline{X}_i^{R,\eta} X_{\nu}^R Z^{RS} \underline{X}_j^{S,\eta} \bar{X}_b^{S,\eta} - \underline{X}_i^{R,\eta} \bar{X}_b^{R,\eta} Z^{RS} \underline{X}_j^{S,\eta} X_{\nu}^S] \\
 &\quad + \sum_{ij} \sum_b \sum_{PQRS} \underline{X}_i^{P,\eta} X_{\nu}^P Z^{PQ} \underline{X}_j^{Q,\eta'} \bar{X}_b^{Q,\eta'} \\
 &\quad \underline{X}_i^{R,\eta} X_{\nu}^R Z^{RS} \underline{X}_j^{S,\eta'} \bar{X}_b^{S,\eta'}
 \end{aligned}
 \quad (22)$$

where  $\underline{\mathbf{X}}$  and  $\bar{\mathbf{X}}$  are the collocation matrices  $\mathbf{X}$  transformed into the occupied and virtual Cholesky pseudo-MO basis, respectively. In MP2 it is often advisable to treat Coulomb- and exchange-like contributions separately,<sup>58</sup> thus the  $\mathbf{R}$ -matrices are partitioned into  $\mathbf{R}^C$ , the Coulomb-like contribution, and  $\mathbf{R}^X$ , the exchange-like contribution.

**2.4.1. THC R-Matrices: Coulomb-like Contribution.** The Coulomb-like parts  $\underline{\mathbf{R}}^C$  and  $\bar{\mathbf{R}}^C$  are given by

$$\begin{aligned}
 \bar{\mathbf{R}}_{\mu'\mu}^{\eta,C}(\kappa) &= \sum_j \sum_{ab} \sum_{PQRS} X_{\mu}^P \bar{X}_a^{P,\eta} Z^{PQ} \underline{X}_j^{Q,\eta} \bar{X}_b^{Q,\eta} \\
 &\quad X_{\mu}^R \bar{X}_a^{R,\eta} Z^{RS} \underline{X}_j^{S,\eta} \bar{X}_b^{S,\eta} \\
 &\quad + \sum_j \sum_{ab} \sum_{PQRS} X_{\mu}^P \bar{X}_a^{P,\eta} Z^{PQ} \underline{X}_j^{Q,\eta'} \bar{X}_b^{Q,\eta'} \\
 &\quad X_{\mu}^R \bar{X}_a^{R,\eta} Z^{RS} \underline{X}_j^{S,\eta'} \bar{X}_b^{S,\eta'} \\
 \mathbf{R}_{\nu'\nu}^{\eta,C}(\kappa) &= \sum_{ij} \sum_b \sum_{PQRS} \underline{X}_i^{P,\eta} X_{\nu}^P Z^{PQ} \underline{X}_j^{Q,\eta} \bar{X}_b^{Q,\eta} \\
 &\quad \underline{X}_i^{R,\eta} X_{\nu}^R Z^{RS} \underline{X}_j^{S,\eta} \bar{X}_b^{S,\eta} \\
 &\quad + \sum_{ij} \sum_b \sum_{PQRS} \underline{X}_i^{P,\eta} X_{\nu}^P Z^{PQ} \underline{X}_j^{Q,\eta'} \bar{X}_b^{Q,\eta'} \\
 &\quad \underline{X}_i^{R,\eta} X_{\nu}^R Z^{RS} \underline{X}_j^{S,\eta'} \bar{X}_b^{S,\eta'}
 \end{aligned}
 \quad (23)$$

and can—in analogy to the THC-MP2 energy—be efficiently computed using sparse linear algebra.<sup>41</sup> In doing so,  $\underline{\mathbf{R}}^C$  and  $\bar{\mathbf{R}}^C$  are especially efficient to compute, as their formation only involves BLAS level 3 operations. For an efficient implementation it is important to realize, that for large enough molecules and appropriate ordering of the THC grid points, the collocation



**Figure 1.** Tensor network representation of the contractions performed by Algorithms 1 and 2 for the formation of a Coulomb-like contribution  $\bar{\mathbf{R}}^{C,\eta}$ . By symmetry,  $\bar{\mathbf{R}}^{C,\eta}$  can be formed analogously, but with intermediates **A** and **B** interchanged. Tensors contracted in ensuing steps are highlighted.

matrices **X** become sparse, while the **Z**-matrix, being the representation of the long-ranged  $1/r$  operator, will always remain dense. By closer inspection of eq 23, it can be noticed, that the **X**-matrices corresponding to the MOs of the ket of the decomposed ERI and the **Z**-matrix are identical for all terms. Therefore, these matrices can be collected in an intermediate  $\mathbf{D}^\eta$  given by Algorithm 1, where intermediates **A** and **B** are given by

$$\begin{aligned} A^{QS,\eta} &= \sum_j \underline{X}_j^{Q,\eta} \underline{X}_j^{S,\eta}, \\ B^{QS,\eta} &= \sum_b \bar{X}_b^{Q,\eta} \bar{X}_b^{S,\eta}, \end{aligned} \quad (24)$$

and represent the occupied and virtual pseudodensity in the grid basis, respectively. Additionally, the  $\Lambda$ -factorization of the **Z**-matrix, i.e.,  $\mathbf{Z} \equiv \Lambda \Lambda^T$ , is used to lower the prefactor of this step.<sup>34,41</sup>

### Algorithm 1 Precontraction

- 1: **procedure** BUILD\_D(**A** $^\eta$ , **B** $^\eta$ ,  $\Lambda$ )
- 2:  $C^{QS,\eta} = A^{QS,\eta} B^{QS,\eta}$
- 3:  $D'_{\alpha\alpha'}{}^\eta = \sum_{QS} \Lambda_\alpha^Q C^{QS,\eta} \Lambda_{\alpha'}^S$
- 4:  $D^{PR,\eta} = \sum_{\alpha\alpha'} \Lambda_\alpha^P D'_{\alpha\alpha'}{}^\eta \Lambda_{\alpha'}^R$
- 5: **return** **D** $^\eta$
- 6: **end procedure**

As forming intermediate **D** requires a series of dense matrix–matrix-multiplications, this step will require the majority of the computation time for  $\bar{\mathbf{R}}^C$  and  $\bar{\mathbf{R}}^C$ . However, Algorithm 1 has to be performed only once per Laplace point and electron spin. The final contribution to the **R**-matrices can then simply be obtained by a Schur product and two matrix–matrix-multiplications given by Algorithm 2.

### Algorithm 2 Formation of a Coulomb-like contribution $\bar{\mathbf{R}}^{C,\eta}$ .

- 1: **procedure** BUILD\_R<sup>C</sup>(**X**, **C** $^\eta$ , **D** $^{\eta'}$ )
- 2:  $E^{PR,\eta} = C^{PR,\eta} D^{PR,\eta'}$
- 3:  $R_{\lambda'\lambda}^{C,\eta} = \sum_{PR} X_{\lambda'}^P E^{PR,\eta} X_\lambda^R$
- 4: **return** **R**<sup>C, $\eta$</sup>
- 5: **end procedure**

The effects of Algorithms 1 and 2 can be best understood by visualizing the underlying tensor contractions in a tensor network diagram, as given in Figure 1. For an introduction of tensor network diagrams, also in the context of THC, refer to the work by Schutski et al.<sup>46</sup> Algorithms 1 and 2 are then pieced

together with the algorithm for the exchange-like part, detailed in the next section, for the final Algorithm 4 in section 2.4.3.

**2.4.2. THC R-Matrices: Exchange-like Contribution.** Usually when higher than second-order tensors arise, the associated tensor contractions are either carried out with the tensors reshaped into matrices or batched over the dimensions exceeding matrix dimensionality in so-called tensor slices. For an efficient contraction when iterating over tensor slices it is advisable to make use of an underlying structure to reduce the dimensions of the slices, either by matrix decomposition of the slice or by neglecting noncontributing elements. If whole rows and columns are excluded based on some significance criterion, one arrives at the natural blocking (NB) formalism<sup>57,58</sup> for tensor contractions. NB relies on significance lists, which in general terms describe which pairs, of a general index pair  $i$  and  $j$ , contribute to a tensor contraction involving these indices. Mathematically speaking these lists are sets and thus can, in set-builder notation, be represented as

$$j_i \equiv \{j\}_i = \{j | |A_{ij}| > \epsilon_{\text{NB}}\} \quad (25)$$

where **A** is a screening matrix involving indices  $i$  and  $j$ , and  $\epsilon_{\text{NB}}$  is the NB screening threshold. To avoid confusion, we use  $j_i$  as a shorthand notation for the set of all significant  $j$  for a particular index  $i$  and  $\{j_i\}$  for the set of all  $j_i$ . If two elements in a set  $\{j_i\}$  are identical, it technically becomes a multiset, as elements in sets are only allowed to have a multiplicity of 1. In set terminology, the set  $\{i\}_j$  is the transpose of  $\{j\}_i$  and can analogously determined from  $\mathbf{A}^T$  as

$$i_j \equiv \{i\}_j = \{i | |A_{ji}| > \epsilon_{\text{NB}}\} \quad (26)$$

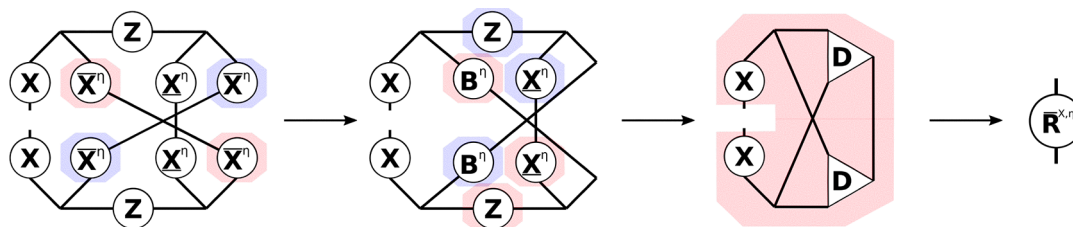
or directly from  $\{j_i\}$ . Another important quantity is the number of significant pairs  $N_{ij}$ , which is defined as

$$N_{ij} = \sum_{i_j \in \{i\}_j} |i_j| \quad (27)$$

NB and THC work particularly well together for exchange-like contractions of ERIs, as within the THC formalism the necessary screening matrices can easily be constructed as outlined in the following. The following prototypical exchange-like ERI contraction from the THC-CDD-MP2 energy expression will serve as an example:

$$\begin{aligned} & (\underline{i}\bar{a}|\underline{j}\bar{b})(\underline{i}\bar{b}|\underline{j}\bar{a}) \\ & \approx \sum_{PQRS} \underline{X}_i^P \bar{X}_a^P \underline{Z}^{PQ} \underline{X}_j^Q \bar{X}_b^Q \cdot \underline{X}_i^R \bar{X}_b^R \underline{Z}^{RS} \underline{X}_j^S \bar{X}_a^S \end{aligned} \quad (28)$$

Two types of indices are present in the above equation, the orbital indices  $i$  and  $j$  (occupied space) as well as  $a$  and  $b$  (virtual space) and the THC auxiliary indices  $P$ ,  $Q$ ,  $R$ , and  $S$ . In the following discussion we use the LS formulation of THC, but the statements also hold for other THC variants. Thus, there are



**Figure 2.** Tensor network representation of the contractions performed by Algorithm 3 for the formation of an exchange-like contribution  $\bar{\mathbf{R}}^{X,\eta}$ . By symmetry,  $\mathbf{R}^{X,\eta}$  can be formed analogously, however, in Algorithm 3 a different approach is used to reduce the prefactor. Tensors contracted in ensuing steps are highlighted.

three general types of index pairs: orbital–orbital, orbital–grid, and grid–grid index pairs. The set of significance lists for orbital–grid type pairs is especially easy to construct, since it can be directly derived from the collocation matrices  $\mathbf{X}$ . For example, the set of all significant grid points  $P$  for a given occupied orbital  $i$  can be constructed as

$$P_i \equiv \{P\}_i = \{P \mid |\underline{X}_i^P| > \epsilon_{\text{NB}}\} \quad (29)$$

Significant orbital–orbital pairs are also easily obtained from the collocation matrices; for example, the set of significant virtual orbitals  $a$  for a given occupied orbital  $i$  can be built as

$$a_i \equiv \{a\}_i = \{a \mid \sum_P |\underline{X}_i^P| |\bar{X}_a^P| > \epsilon_{\text{NB}}\} \quad (30)$$

The screening criterion from eq 30 can also be interpreted to yield only  $ia$  pairs, for which the orbitals have significant overlap and which produce non-negligible charge densities. For the development of low-scaling exchange-type contractions it is important to make use of the exponential coupling between all orbital indices. Orbital  $i$  couples to orbital  $a$  in an exponentially decaying fashion in the bra of the first ERI in eq 28. Likewise, orbital  $j$  also couples to orbital  $a$  in an exponentially decaying fashion in the ket of the second ERI. Thus, there is indirect exponential coupling between orbitals  $i$  and  $j$  via orbital  $a$  and the set of significant orbitals  $i$  for a given orbital  $j$  can be derived from the sets  $\{a_j\}$ , which is identical to  $\{a_i\}$ , and  $\{a_i\}$ :

$$i_j \equiv \{i\}_j = \{i \mid a_i \cap a_j \neq \emptyset\} \quad (31)$$

With the significance lists given by eqs 29, 30, and 31, an asymptotically linear scaling algorithm for the exchange-like energy contribution to MP2 can be devised. The algorithm is detailed in the Supporting Information and close to linear scaling is demonstrated. Here, however, the focus lies on the exchange-like parts of the  $\mathbf{R}$ -matrices, which are conceptually similar, but different in that two AO indices remain uncontracted. The tensor contractions necessary for forming the  $\mathbf{R}^X$  parts can again best be understood from the tensor network diagram in Figure 2.

While Figure 2 only shows the  $\bar{\mathbf{R}}^X$  part, the  $\mathbf{R}^X$  contribution can be constructed analogously. First, intermediates  $\mathbf{A}$  or  $\mathbf{B}$  are formed from the collocation matrices  $\underline{\mathbf{X}}$  or  $\bar{\mathbf{X}}$ , respectively. Second, the remaining collocation matrices, the  $\mathbf{Z}$ -matrices and intermediates  $\mathbf{A}$  or  $\mathbf{B}$  are contracted to the third-order tensor intermediate  $\mathbf{D}$ , where the symmetry of the tensors is used to reduce the operation count. The idea for an efficient implementation is then as follows: to avoid storing third-order tensors, the contraction is batched over the occupied orbital index common to both  $\mathbf{R}$  and  $\bar{\mathbf{R}}$ , and to reduce the cost of the `dgemm` operations within the loop NB is applied. In this way separate algorithms for  $\mathbf{R}^X$  and  $\bar{\mathbf{R}}^X$  can be formulated, which are

presented in the Supporting Information. By closer inspection, however, it can be seen, that both  $\mathbf{R}^X$ -matrices share the most expensive to compute intermediate, which incorporates the  $\mathbf{Z}$ -matrix. Therefore, a joint computation as given by Algorithm 3 is preferred.

**Algorithm 3** Formation of the exchange-like contributions  $\underline{\mathbf{R}}^{X,\eta}$  and  $\bar{\mathbf{R}}^{X,\eta}$ .

```

1: procedure BUILD_RX( $\mathbf{X}$ ,  $\underline{\mathbf{X}}^\eta$ ,  $\bar{\mathbf{X}}^\eta$ ,  $\mathbf{A}^\eta$ ,  $\mathbf{B}^\eta$ ,  $\mathbf{Z}$ )
2:   build  $\{S_j\} (\equiv \{Q_j\})$  from  $|\underline{X}_j^{S,\eta}|$  (eq. 29)
3:   build  $\{b_R\}$  from  $|\bar{X}_b^{R,\eta}|$  (eq. 29)
4:   build  $\{b_j\}$  from  $\sum_Q |\underline{X}_j^{Q,\eta}| |\bar{X}_b^{Q,\eta}|$  (eq.
5:   30)
6:   build  $\{R_j\} (\equiv \{P_j\})$  from  $\{b_R\}$  and  $\{b_j\}$ 
7:   (eq. 31)
8:    $\mathbf{E}, \mathbf{F} = \mathbf{0}_{N_{\text{grid}}, N_{\text{grid}}}$ 
9:   for  $j = 1$  to  $j \leq \text{rk}(\mathbf{P}^\eta)$  do
10:    for all  $R \in \{R\}_j$  and  $S \in \{S\}_j$  do
11:       $C^{RS,\eta} = Z^{RS} \underline{X}_j^{S,\eta}$ 
12:    end for
13:    for all  $P \in \{P\}_j$  and  $R \in \{R\}_j$  do
14:       $D^{PR,\eta} = \sum_{Q \in \{Q\}_j} C^{PQ,\eta} B^{QR,\eta}$ 
15:    end for
16:    for all  $P \in \{P\}_j$  and  $R \in \{R\}_j$  do
17:       $E^{PR,\eta} += D^{PR,\eta} D^{RP,\eta}$ 
18:    end for
19:    for all  $P \in \{P\}_j$  and  $R \in \{R\}_j$  do
20:       $D^{PR,\eta} \times = A^{PR,\eta}$ 
21:    end for
22:    for all  $P \in \{P\}_j$  and  $S \in \{S\}_j$  do
23:       $F^{PS,\eta} += \sum_{R \in \{R\}_j} D^{PR,\eta} C^{RS,\eta}$ 
24:    end for
25:  end for
26:   $\bar{\mathbf{R}}^{X,\eta} = \sum_{P,R} X_{\mu'\mu}^P E^{PR,\eta} X_\mu^R$ 
27:   $\underline{\mathbf{R}}^{X,\eta} = \sum_{P,S} X_{\nu'\nu}^P F^{PS,\eta} X_\nu^S$ 
28:  return  $\bar{\mathbf{R}}^{X,\eta}$ ,  $\underline{\mathbf{R}}^{X,\eta}$ 
29: end procedure

```

First, according to eqs 29, 30, and 31 all necessary significance lists are computed and then the precursor intermediates to  $\mathbf{R}^X$  and  $\bar{\mathbf{R}}^X$ , that is, the matrices  $\mathbf{E}$  and  $\mathbf{F}$ , are accumulated in a loop over the common occupied orbital index  $j$ . Lines 12 and 21

represent the bottleneck of the algorithm as their evaluation formally scales as  $O(N_{\text{grid}}^3 N_{\text{occ}})$ . However, since all involved grid indices  $P$ ,  $Q$ , and  $R$  are connected to  $j$  through the significance lists  $\{S_j\}$  (identical to  $\{Q_j\}$ ) and  $\{R_j\}$  (identical to  $\{P_j\}$ ), the size of the involved matrices in NB format is asymptotically constant. Therefore, as  $N_{\text{occ}}$  grows linearly with the molecule size, the formation of the  $\mathbf{R}^X$ -parts can be done in linear scaling time. For this, appropriate thresholds have to be chosen for the formation of the significance lists. Instead of using the same threshold for all pairs, we identify three different types of lists: (1) grid-occupied orbital index pairs, e.g.,  $\{S_j\}$ , (2) grid-virtual orbital index pairs, e.g.,  $\{b_R\}$ , where, in both cases, the orbital index is coupled to the grid index directly by a collocation matrix, and (3) virtual-occupied orbital index pairs, e.g.,  $\{b_j\}$ , which are coupled in real-space over grid points. While the final algorithm only requires the  $\{S_j\}$  and  $\{R_j\}/\{P_j\}$  lists, the selection of index pairs included in  $\{b_R\}$  and  $\{b_j\}$  is still important, as the  $\{R_j\}/\{P_j\}$  lists are built from these lists analogously to eq 31.

**2.4.3. THC R-Matrices: Final Algorithm.** Piecing together Algorithms 1 and 2 for the Coulomb-like part and Algorithm 3 for the exchange-like part, the final algorithm for the formation of the  $\mathbf{R}$ -matrices is given by Algorithm 4.

---

#### Algorithm 4 Final algorithm for $\overline{\mathbf{R}}$ and $\underline{\mathbf{R}}$

---

```

1: procedure BUILD_R( $\mathbf{X}, \underline{\mathbf{X}}^\eta, \overline{\mathbf{X}}^\eta, \mathbf{Z}, \Lambda$ )
2:    $\mathbf{A}^\alpha, \mathbf{A}^\beta, \mathbf{B}^\alpha, \mathbf{B}^\beta, \mathbf{D}^\alpha, \mathbf{D}^\beta = \mathbf{0}_{N_{\text{grid}}, N_{\text{grid}}}$ 
3:   for  $\eta \in \{\alpha, \beta\}$  do
4:      $\mathbf{A}^\eta = \underline{\mathbf{X}}^\eta (\underline{\mathbf{X}}^\eta)^\text{T}$ 
5:      $\mathbf{B}^\eta = \overline{\mathbf{X}}^\eta (\overline{\mathbf{X}}^\eta)^\text{T}$ 
6:      $\mathbf{D}^\eta = \text{BUILD\_D}(\mathbf{A}^\eta, \mathbf{B}^\eta, \Lambda)$ 
7:   end for
8:    $\overline{\mathbf{R}}^\alpha, \overline{\mathbf{R}}^\beta, \underline{\mathbf{R}}^\alpha, \underline{\mathbf{R}}^\beta = \mathbf{0}_{N_{\text{bf}}, N_{\text{bf}}}$ 
9:   for  $\eta \in \{\alpha, \beta\}$  do
10:     $\overline{\mathbf{R}}^\eta += \text{BUILD\_R}^C(\mathbf{X}, \mathbf{B}^\eta, \mathbf{D}^\eta)$ 
11:     $\overline{\mathbf{R}}^\eta += \text{BUILD\_R}^C(\mathbf{X}, \mathbf{B}^\eta, \mathbf{D}^\eta)$ 
12:     $\underline{\mathbf{R}}^\eta += \text{BUILD\_R}^C(\mathbf{X}, \mathbf{A}^\eta, \mathbf{D}^\eta)$ 
13:     $\underline{\mathbf{R}}^\eta += \text{BUILD\_R}^C(\mathbf{X}, \mathbf{A}^\eta, \mathbf{D}^\eta)$ 
14:     $\overline{\mathbf{R}}^\eta, \underline{\mathbf{R}}^\eta += \text{BUILD\_R}^X(\mathbf{X}, \underline{\mathbf{X}}^\eta, \overline{\mathbf{X}}^\eta,$ 
       $\mathbf{A}^\eta, \mathbf{B}^\eta, \mathbf{Z})$ 
15:   end for
16:   return  $\overline{\mathbf{R}}^\alpha, \overline{\mathbf{R}}^\beta, \underline{\mathbf{R}}^\alpha, \underline{\mathbf{R}}^\beta$ 
17: end procedure

```

---

Overall, Algorithm 4 has to be executed once per Laplace point and is separated into a precontraction phase and the phase for the actual formation of the contributions to  $\underline{\mathbf{R}}$  and  $\overline{\mathbf{R}}$ . In the precontraction phase intermediates  $\mathbf{D}^\eta$  are formed, which represent the most expensive part for the Coulomb-like terms, as the subsequent calls to Algorithm 2 only contribute a Schur product and two dgemm operations. These dgemm calls contribute a negligible overhead compared with line 3 of Algorithm 1, as the dimensions of the matrices involved are reduced. In total, however, Algorithm 3 will dominate the runtime for forming the  $\mathbf{R}$ -matrices due to its formal  $O(N^4)$  scaling.

### 3. COMPUTATIONAL DETAILS

The above-described THC- $\omega$ -RI-CDD-MP2 HFCC code is implemented within our quantum chemistry package FERMIONS++.<sup>59–61</sup> For the THC-based HFCC calculations the hand-optimized grids by Martínez and co-workers<sup>38</sup> were used together with the Dunning cc-pVXZ ( $X \in \{D, T\}$ ) basis sets<sup>62</sup> and the corresponding auxiliary basis sets. For the phosphorus atoms in the DNA backbone, the fluorine grids were used without loss of accuracy as reported in our work on THC-MP2 energies.<sup>41</sup> All calculations were carried out without the frozen-core approximation. All preceding SCF calculations were converged to an energy difference of  $10^{-8}$  H and a FPS–SPF commutator difference of  $10^{-7}$  using DIIS acceleration.<sup>63</sup> For the gradient calculations seven Laplace points were used in the expansion and the DL-UCPSCF algorithm was converged to an error of  $10^{-4}$  for all molecules of the benchmark set in section 4. For all subsequent calculations on larger molecules, a threshold of  $10^{-3}$  was used. These settings were shown to yield errors below 1 MHz.<sup>6,30</sup> For the assessment of the accuracy of the THC- $\omega$ -RI-CDD-MP2 HFCCs against other methods, the standard orientation was used. For the THC factorization of the ERIs an attenuation strength of 0.1 in the attenuated Coulomb metric was used<sup>54,55</sup> and the same general protocol for screening based on integral partition bounds (IPBs)<sup>64</sup> and NB, as in our work on THC-MP2 energies, was followed, although adjusted to fit ERIs in the AO basis.<sup>41</sup> All timings are done on an AMD EPYC 7302 (3.30 GHz) CPU node with 256 GB RAM and 1.7 TB of SSD disk space.

### 4. RESULTS AND DISCUSSION

First, the accuracy of the newly developed THC- $\omega$ -RI-CDD-MP2 method for the calculation of isotropic HFCCs is assessed against our reference  $\omega$ -RI-CDD-MP2 implementation.<sup>6,30</sup> Next, the thresholds necessary for the screening in the expensive exchange-like contribution  $\mathbf{R}^X$  are optimized on a set of medium-sized organic radicals. Finally, the scaling of the THC- $\omega$ -RI-CDD-MP2 method is analyzed and timings are compared to the  $\omega$ -RI-CDD-MP2 reference for a set of representative radicals.

**4.1. Accuracy of THC- $\omega$ -RI-CDD-MP2 HFCCs.** Throughout this publication, our  $\omega$ -RI-CDD-MP2 implementation for the computation of HFCCs by Vogler et al.,<sup>6,30</sup> which was verified against the RI-MP2 implementation in the ORCA program package,<sup>65</sup> will serve as reference. The original implementation, however, made use of the SOS approximation and excluded the exchange-like terms. To enable a fair comparison, the exchange-like terms were added analogously to the earlier implemented RI-CDD-based AO-MP2 gradient,<sup>6</sup> albeit with the attenuated Coulomb metric for the RI integrals. For the comparison, first the accuracy of the presented THC- $\omega$ -RI-CDD-MP2 method for HFCCs is assessed. The THC- $\omega$ -RI-CDD-MP2 method is benchmarked using a set of 12 organic radicals from a recent study<sup>7</sup> on the effects of electron correlation, molecular dynamic contributions, and solvation effects on HFCCs. Mean absolute deviations (MAD), root-mean-square deviations (RMSD), and absolute maximum deviations (MAX) are given in Table 1. We note that we used all 12 radicals for the comparison, even though, as Vogler et al.<sup>7</sup> pointed out, some molecules are spin contaminated. While the latter certainly has an effect on the reliability of the results, when comparing to experiment, it should not influence the comparison of different methods.

**Table 1.** Errors of the HFCCs Obtained with the THC- $\omega$ -RI-CDD-MP2 Method Compared to the  $\omega$ -RI-CDD-MP2 Reference Implementation for the HFCC Benchmark Set from Vogler et al.<sup>7</sup> and the cc-pVXZ/cc-pVXZ-RI ( $X \in \{D, T\}$ ) Basis Sets

basis set	MAD <sup>a</sup>	RMSD <sup>a</sup>	MAX <sup>a</sup>
cc-pVDZ	0.304	0.478	1.822
cc-pVTZ	0.092	0.167	0.675

<sup>a</sup>Deviations in MHz.

Table 1 shows that the mean errors for the THC- $\omega$ -RI-CDD-MP2 method are below 1 MHz for both basis sets, while the MAX error corresponds to atoms with high spin density, that is, <sup>19</sup>F in the CF<sub>3</sub> radical for the double- $\zeta$  basis set and <sup>11</sup>B in the BH<sub>3</sub> radical for the triple- $\zeta$  basis set. While for these nuclei the absolute error is larger, the relative error is still below 1%, as due to their high spin density the HFCCs are large in magnitude. The origin of these errors is mainly based on the following two shortcomings: (1) Using the hand-optimized THC grids by Martínez and co-workers<sup>38</sup> for AO-THC incurs additional errors over MO-THC as these grids were optimized for fitting MO-based ERIs, for which the orbital space is much more compact compared to ERIs in the AO basis. This phenomenon was observed similarly in our recent work on THC- $\omega$ -RI-CDD-MP2 energies.<sup>41</sup> (2) Since grid-based THC uses DFT-like integration grids, THC-based properties are likewise prone to not being rotationally invariant. In DFT the problem is alleviated through larger grids, which THC cannot make use of without forfeiting the reduction in computational cost compared to the respective canonical method. However, we found that these rotational errors are on the order of 0.01–0.05 MHz, depending on the magnitude of the spin density on the respective nucleus. Overall, we consider a mean deviation of less than 1 MHz to be less than the method error of RI-MP2 and certainly accurate enough, when compared against experimental results, where, as shown by Vogler et al.,<sup>7</sup> other effects like dynamic contributions or solvation effects contribute significantly.

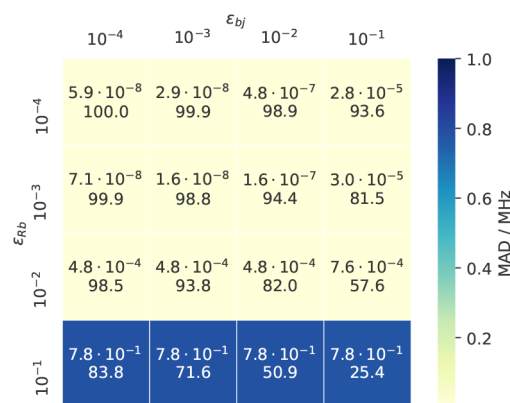
**4.2. Threshold Optimization.** After having established that the presented THC- $\omega$ -RI-CDD-MP2 method provides reliable HFCCs, the focus is now on optimizing the time complexity of the underlying algorithm while preserving the accuracy. An obvious point for optimization is the formation of the exchange-like parts  $R^X$ , for which Algorithm 3 has quartic scaling if no screening is applied. However, as discussed in section 2.4.2, by applying NB with carefully chosen thresholds, the formation of  $R^X$  should be possible with linear time complexity. As outlined in section 2.4.2, different thresholds will be used in the screening process for the different types of significance lists. We associate the thresholds  $\epsilon_{S_j}$ ,  $\epsilon_{bR}$ , and  $\epsilon_{bj}$  with the significance lists  $\{S_j\}$ ,  $\{b_R\}$ , and  $\{b_j\}$ , respectively. Since  $\epsilon_{S_j}$  directly determines the pairs included in  $\{S_j\}$ , and  $\epsilon_{bR}/\epsilon_{bj}$  only indirectly determine the pairs in  $\{R_j\}$ , the optimization of these thresholds is simplified by separately optimizing  $\epsilon_{S_j}$  and  $\epsilon_{bR}/\epsilon_{bj}$ . The thresholds are first optimized for the smaller double- $\zeta$  basis set and later transferred to the triple- $\zeta$  basis. For this, we chose a set of six medium-sized radical molecules and supramolecular assemblies, which are large enough for the screening to have effect. For further information on this benchmark set see the Supporting Information. Table 2 summarizes the errors and the resulting average number of significant pairs  $\bar{N}_{S_j}$  in  $\{S_j\}$  for the optimization of  $\epsilon_{S_j}$ .

**Table 2.** Threshold Optimization of  $\epsilon_{S_j}$ : Errors of the HFCCs and Average Numbers of Significant Pairs ( $\bar{N}_{S_j}$ ) from the Screening Benchmark Set, Obtained with the Chosen Threshold for  $\epsilon_{S_j}$  ( $\epsilon_{bR} = 0$ ,  $\epsilon_{bj} = 0$ , cc-pVDZ)

$\epsilon_{S_j}$	$\bar{N}_{S_j}$ <sup>a</sup>	MAD <sup>b</sup>	RMSD <sup>b</sup>	MAX <sup>b</sup>
$10^{-6}$	77.6	$2.8 \times 10^{-7}$	$8.7 \times 10^{-7}$	$1.1 \times 10^{-5}$
$10^{-5}$	57.9	$8.2 \times 10^{-6}$	$2.0 \times 10^{-5}$	$1.8 \times 10^{-4}$
$10^{-4}$	36.7	$1.5 \times 10^{-4}$	$3.5 \times 10^{-4}$	$3.4 \times 10^{-3}$
$10^{-3}$	12.8	$3.5 \times 10^{-3}$	$1.1 \times 10^{-2}$	$1.9 \times 10^{-1}$
$10^{-2}$	5.8	$9.2 \times 10^{-2}$	$3.6 \times 10^{-1}$	$6.2 \times 10^0$

<sup>a</sup>Ratio in %. <sup>b</sup>Deviations in MHz.

While for threshold values of  $10^{-6}$  through  $10^{-4}$  the error remains negligible, the screening shows an effect in that  $\bar{N}_{S_j}$  indicates that only roughly a third of the pairs in  $\{S_j\}$  are necessary for this accuracy. The mean errors grow roughly linearly with loosening thresholds and remain sufficiently small for a range of threshold values. For  $\epsilon_{S_j} \geq 10^{-3}$  especially the MAX error deteriorates above 1 MHz, while the MAD remains below 0.1 MHz. For the optimization of  $\epsilon_{bR}$  and  $\epsilon_{bj}$ , combinations of thresholds have to be considered, since they determine, through eq 31, the significant pairs in  $\{R_j\}$ . The results of this optimization are shown in Figure 3 as a heatmap.



**Figure 3.** Threshold optimization of  $\epsilon_{bR}$  and  $\epsilon_{bj}$ : In each cell the MADs (in MHz) of the HFCCs (top value) and the average numbers of significant pairs ( $\bar{N}_{R_j}$  in %, bottom value) are given as an average from the screening benchmark set ( $\epsilon_{S_j} = 0$ , cc-pVDZ).

From Figure 3 it can be seen, that the MAD is stable through a wide range of threshold values, only significantly worsening when choosing  $\epsilon_{bR} > 10^{-2}$ , irrespective of the value chosen for  $\epsilon_{bj}$ . The observation that even for looser thresholds, for example,  $\epsilon_{bR} = \epsilon_{bj} = 10^{-2}$ , still roughly 80% of the pairs in  $\{R_j\}$  are significant stems from the fact that indices  $R$  and  $j$  are coupled indirectly over a virtual orbital  $b$ . According to eq 31 a pair of indices  $R$  and  $j$  is only considered insignificant, if they do not share significant overlap with any virtual orbital  $b$ . Therefore,  $\bar{N}_{R_j}$  will always be greater than  $\bar{N}_{S_j}$  for any sensibly chosen combination of thresholds.

With the separate optimization of  $\epsilon_{S_j}$  and  $\epsilon_{bR}/\epsilon_{bj}$  as a starting point, different combinations of these three thresholds were tested. The best trade-off between accuracy and the number of significant pairs appeared to be for the thresholds  $\epsilon_{S_j} = 10^{-3}$ ,  $\epsilon_{bR} = 10^{-2}$ , and  $\epsilon_{bj} = 10^{-2}$ , for which the errors are summarized in Table 3.

The chosen combination of thresholds provides good accuracy for both basis sets, with the errors for the triple- $\zeta$



**Table 3. Errors of the HFCCs from the Screening Benchmark Set, Obtained with the Chosen Thresholds ( $\epsilon_{S_j} = 10^{-3}$ ,  $\epsilon_{b_R} = 10^{-2}$ ,  $\epsilon_{b_j} = 10^{-2}$ ) for the THC- $\omega$ -RI-CDD-MP2 method referenced against the same method with disabled screening**

basis set	MAD <sup>a</sup>	RMSD <sup>a</sup>	MAX <sup>a</sup>
cc-pVDZ	0.004	0.011	0.189
cc-pVTZ	0.023	0.077	0.941

<sup>a</sup>Deviations in MHz.

basis being somewhat larger. The latter could be improved through a separate optimization of the thresholds for the triple- $\zeta$  basis. However, in view of the fact that the mean errors are still below 0.1 MHz, the thresholds optimized for the double- $\zeta$  basis set seem to be suitable for the larger basis set as well.

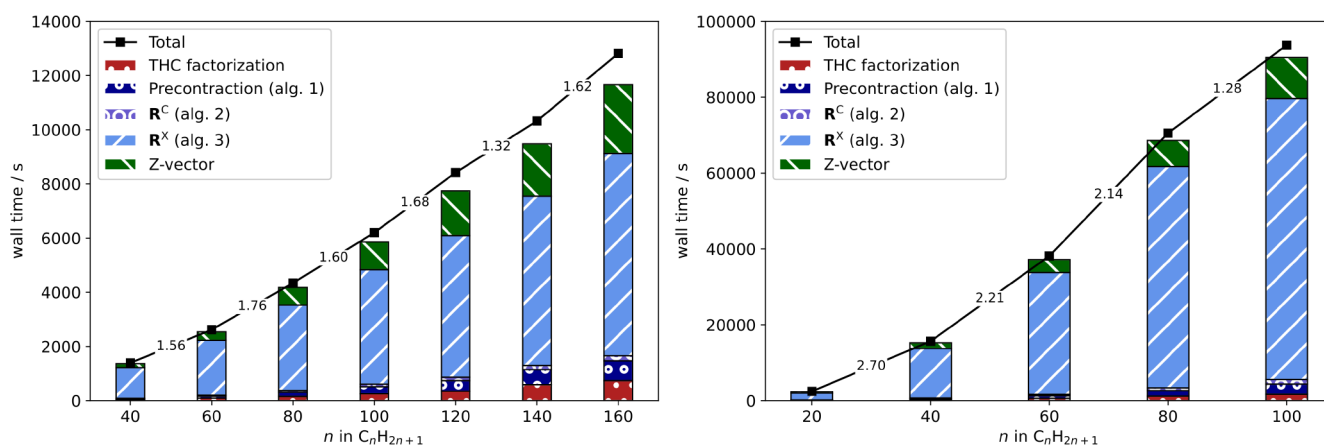
**4.3. Timings and Scaling.** With the optimized screening thresholds at hand, the scaling behavior of the THC- $\omega$ -RI-CDD-MP2 method for the computation of HFCCs is analyzed and compared to the previous  $\omega$ -RI-CDD-MP2 implementation.<sup>6,30</sup> For the assessment of the asymptotic scaling behavior, HFCCs for a series of linear alkyl radicals  $C_nH_{2n+1}$  are computed. In Figure 4 the timings are shown together with the underlying contributions from the most significant steps for both basis sets.

As is evident from Figure 4, the overall scaling behavior is governed by the contribution from the exchange-like parts of  $R$  (Algorithm 3), while the Coulomb-like terms (Algorithms 1 and 2) and the overhead from obtaining the THC factorization only contribute marginally. For both basis sets, the THC- $\omega$ -RI-CDD-MP2 method reaches subquadratic scaling, while the scaling is also partly influenced by the Fock matrix builds in the  $Z$ -vector step. The increased cost of the  $Z$ -vector step for larger fragment sizes in the case of the double- $\zeta$  basis set is also the reason for the overall scaling exponent slightly deteriorating beyond  $C_{140}H_{281}$  to 1.62. To prevent this unfavorable scaling for the larger triple- $\zeta$  basis set, the recommendations by Laqua et al.,<sup>66</sup> which are default settings in FERMIONS++, are followed, and the recently presented seminumerical exchange method (sn-Link) is used for the exchange part of the Fock matrices. The latter makes the overall scaling for the cc-pVTZ basis set almost entirely be governed by Algorithm 3 (blue bars). Therefore, the scaling reduces to close to linear for the largest fragment size considered. The same is true for the cc-pVDZ basis set, for which Algorithm 3 reaches an apparent asymptotic scaling of 1.3.

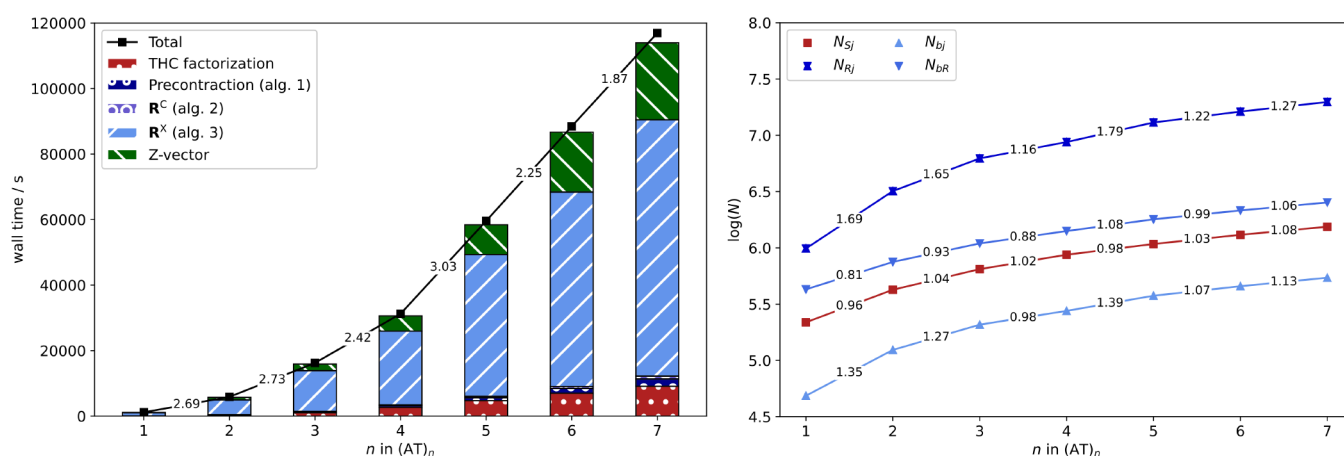
To go toward more chemically relevant systems and beyond what was possible with our previous  $\omega$ -RI-CDD-MP2 implementation, the scaling behavior for spin-labeled adenine–thymine base pair stacks  $(AT)_n$  is assessed.

Figure 5 (left) shows the scaling behavior for spin-labeled DNA fragments up to seven repetition units or 5101 basis functions. As expected, the onset of subquadratic scaling is for greater fragment sizes compared to the alkyl radicals. Nonetheless, the scaling exponent decreases to 1.87 for  $(AT)_6 \rightarrow (AT)_7$ . Further reduction with increasing fragment size can be expected based on the growth rate of the number of significant pairs in  $\{S_j\}$  and  $\{R_j\}$ . In Figure 5 (right) the logarithm of the total number of significant pairs is shown for all index pairs relevant for Algorithm 3. The runtime of Algorithm 3 is mainly governed by  $N_{S_j}$  and  $N_{R_j}$  and relies on only a constant number of grid points being significant for a given occupied orbital, see section 2.4.2. If only a constant number of grid points is significant for a given occupied orbital, then  $N_{S_j}$  and  $N_{R_j}$  will grow linearly with increasing molecule size. The latter is demonstrated for  $N_{S_j}$ , that is, the grid point-occupied orbital pair directly coupled by a collocation matrix.  $N_{R_j}$  is inherently greater than  $N_{S_j}$ , since  $\{R_j\}$  is formed from eq 31 with coupling of  $R$  and  $j$  over a virtual orbital  $b$ . The latter is also the reason for the scaling exponent not quite reducing to 1.0 and there being an onset for the close to linear scaling. This also explains why Algorithm 3 reaches subquadratic scaling for the DNA system, but not quite linear scaling. Nonetheless, the THC- $\omega$ -RI-CDD-MP2 method also reaches subquadratic scaling for the spin-labeled DNA fragments under consideration and allows for the computation of HFCCs for almost 500 atoms and more than 5000 basis functions.

While the asymptotic scaling exponent of a quantum chemical method is certainly important for the treatment of large (bio)chemical systems, the prefactor oftentimes determines the applicability of a method for a certain problem. In other words, a method can be linear scaling, but have a prefactor so large, that calculations still remain unfeasible. Another aspect which determines feasibility are memory/storage requirements, oftentimes governed by the necessity to store ERIs or amplitude tensors present in electron correlation methods. These aspects are considered in the following for a comparison of the  $\omega$ -RI-CDD-MP2 method in its all-nuclei variant<sup>6,30</sup> and the presented THC- $\omega$ -RI-CDD-MP2 method for a collection of organic radicals. An additional comparison against the selected-nuclei



**Figure 4.** Detailed timings for the computation of HFCCs with the THC- $\omega$ -RI-CDD-MP2 method (black) for linear alkyl radicals  $C_nH_{2n+1}$  as well as significant contributions from underlying steps (colored bars) for the cc-pVDZ (left) and cc-pVTZ (right) basis sets. The numbers between fragments correspond to the scaling with respect to the preceding fragment.



**Figure 5.** Detailed timings for the computation of HFCCs with the THC- $\omega$ -RI-CDD-MP2 method (black) for spin-labeled (AT) $_n$  radicals as well as significant contributions from underlying steps (colored bars) for the cc-pVDZ basis set. Scaling behavior of the number of significant index pairs  $\{S_j\}$  (red) and  $\{R_j\}$  (dark blue) for increasing fragment sizes (right). The numbers between fragments correspond to the scaling with respect to the preceding fragment.

variant of  $\omega$ -RI-CDD-MP2,<sup>30</sup> which was proposed to alleviate some of the shortcomings of the  $\omega$ -RI-CDD-MP2 method, is given in the Supporting Information. Here, the focus is on typical applications on medium-sized organic radicals. Table 4

**Table 4.** Comparison of the Memory Requirements  $M_{\text{est}}$ , Timings  $t$ , and Relative Speedups  $S$  of the THC- $\omega$ -RI-CDD-MP2 Method with the Previously Implemented  $\omega$ -RI-CDD-MP2 Method

system	$N_{\text{bf}}$	$\omega$ -RI-CDD-MP2		THC- $\omega$ -RI-CDD-MP2		
		$M_{\text{est}}/\text{GB}$	$t/\text{h}$	$M_{\text{est}}/\text{GB}$	$t/\text{h}$	$S$
cc-pVDZ						
C <sub>60</sub> H <sub>121</sub>	1445	201.0	101.8	11.7	0.7	145
TEMPO <sub>H<sub>2</sub>O</sub>	1444	208.5	345.9	11.7	1.5	230
(glu) <sub>4</sub>	1074	90.0	380.2	5.9	0.7	540
PTMA <sub>3</sub>	1102	93.8	620.9	6.5	1.2	517
(AT) <sub>2</sub>	1566	286.0	1111.0	23.1	1.9	585
cc-pVTZ						
C <sub>20</sub> H <sub>41</sub>	1174	66.1	363.3	8.7	0.6	606
TEMPO	582	8.2	160.1	2.1	0.2	891
Tyr	530	6.6	99.4	1.4	0.1	780
Thy	1260	86.6	1346.2 <sup>a</sup>	8.7	1.1	1224
(AT) <sub>1</sub>	1982	340.4	2727.8 <sup>a</sup>	20.7	3.1	880

<sup>a</sup>Timings are estimated conservatively based on the time taken for the first Laplace point.

summarizes the estimated storage requirements  $M_{\text{est}}$  and the wall times  $t$  for the radicals, as well as the speedups  $S$  relative to the  $\omega$ -RI-CDD-MP2 implementation. More information on how the storage requirements are estimated is given in the Supporting Information.

THC has a natural advantage over RI-based methods when it comes to storage requirements. In THC-based methods the largest tensors necessary to keep in memory or store on disk are second-order tensors of dimension  $N_{\text{grid}} \times N_{\text{grid}}$ . For RI-based methods the dimensionality of the factorized representation of the fourth-order ERI tensor increases to three with dimensions  $N_{\text{occ}} \times N_{\text{virt}} \times N_{\text{aux}}$  in the MO basis. Therefore, THC- $\omega$ -RI-CDD-MP2 is superior with an order of magnitude less storage requirements, as can be seen from Table 4. Furthermore, the THC-based method considerably outperforms  $\omega$ -RI-CDD-

MP2 in terms of computation time with speedups up to roughly 600 for the double- $\zeta$  basis set and 800–1200 for the triple- $\zeta$  basis set. The relative speedups are somewhat reduced when comparing the methods for predominantly linear and very sparse systems like linear alkyl radicals, signifying that the  $\omega$ -RI-CDD-MP2 uses the sparsity well in these model cases. For larger and more globular structures, the computational savings with the THC- $\omega$ -RI-CDD-MP2 method are significantly greater, indicating that Algorithm 4 utilizes the sparsity well, even in nonlinear systems. The latter, and the reduced memory requirements make the THC- $\omega$ -RI-CDD-MP2 method attractive for the computation of HFCCs of large, globular systems, as they are commonly encountered in proteins and enzymes. Furthermore, due to the greatly reduced computational cost, the method is applicable in double-hybrid functionals, where usually the MP2 part is the computational bottleneck, while also allowing for sampling of multiple points on the PES.

## 5. CONCLUSIONS AND OUTLOOK

In this work, we presented the THC- $\omega$ -RI-CDD-MP2 method for the efficient and accurate computation of isotropic HFCCs for large organic radicals. As usual for MP2 methods, the exchange-like terms, here  $\mathbf{R}^X$ , govern the scaling and runtime of the method. This issue was addressed through screening based on the THC collocation matrices  $\mathbf{X}$  in combination with natural blocking for the tensor contractions. An asymptotically linear scaling recipe for the contraction of exchange-like terms in THC format is provided. This recipe is applied to the  $\mathbf{R}^X$  terms, reducing the formal quartic scaling to effectively subquadratic, as shown for linear alkyl radicals and spin-labeled DNA strands. The THC- $\omega$ -RI-CDD-MP2 method furthermore highlights the attractiveness of THC-based methods for derivatives of electron correlation methods. Derivative calculations usually involve more types of ERIs, for example, half-transformed integrals, integrals with mixed spin in open-shell calculations, or integrals contracted with perturbed densities. This generally increases storage requirements compared to single point calculations. Furthermore, the computational cost of the method is strongly increased due to additional ERI contractions. Both challenges are overcome by using THC-factorized ERIs, for which only second-order tensors have to be stored and integral transformations and contractions can be easily reduced to simple

dgemm operations. The advantages of THC-based gradient methods are demonstrated for HFCC calculations on a range of medium-sized organic radicals and spin-labeled DNA strands with more than 5000 basis functions.

For future applications, the availability of THC grids has to be improved, as ideally EPR-specific basis sets,<sup>67</sup> and corresponding THC grids, should be used for the calculations presented. The grids used throughout this publications were hand optimized<sup>38</sup> and are only viable for the cc-pVXZ ( $X \in \{D, T\}$ ) basis sets. Furthermore, these grids were optimized based on MO-THC-MP2 and not for AO-THC, as used throughout this work. This incurs additional errors due to the larger fitting space. Different techniques for the on-the-fly generation of THC grids have been proposed, either based on PCD of the THC metric in a larger parent grid basis,<sup>40</sup> or based on centroidal Voronoi tessellation (CVT),<sup>36</sup> and could be applied for the calculation of THC-MP2 HFCCs in future work.

Finally, the developed THC- $\omega$ -RI-CDD-MP2 method can easily be used for the MP2 part of double-hybrid functionals and, once appropriate grids are available, enable accurate HFCC predictions for large molecules. Furthermore, due to the reduced computational complexity and memory requirements, THC- $\omega$ -RI-CDD-MP2, in conjunction with an appropriate double-hybrid functional, is an attractive candidate for QM/MM HFCC calculations.

## ■ ASSOCIATED CONTENT

### SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jctc.2c00118>.

Detailed algorithm for the asymptotically linear scaling formation of THC-MP2 exchange-like energy contribution as well as the individual formation of  $\mathbf{R}^X$  and  $\mathbf{R}^X$ , detailed results for the benchmark calculations, details on the memory requirement estimation and comparison of selected-nuclei  $\omega$ -RI-CDD-MP2 against THC- $\omega$ -RI-CDD-MP2 (PDF) Structures (ZIP)

## ■ 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](https://orcid.org/0000-0002-4189-6558); Email: [christian.ochsenfeld@cup.uni-muenchen.de](mailto: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/doi/10.1021/acs.jctc.2c00118>

### Funding

Open access funded by Max Planck Society.

### 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. The authors thank J. Kussmann (LMU Munich) for providing a development version of the FERMIONS++ program package.

## ■ 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. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1*, 509–530.
- (3) Bent, G. D. Many-body perturbation theory electronic structure calculations for the methoxy radical. II. Hyperfine coupling coefficients. *J. Chem. Phys.* **1994**, *100*, 8219.
- (4) Kossmann, S.; Neese, F. Correlated ab initio spin densities for larger molecules: Orbital-optimized spin-component-scaled MP2 method. *J. Phys. Chem. A* **2010**, *114*, 11768–11781.
- (5) Sandhoefer, B.; Kossmann, S.; Neese, F. Derivation and assessment of relativistic hyperfine-coupling tensors on the basis of orbital-optimized second-order Møller-Plesset perturbation theory and the second-order Douglas-Kroll-Hess transformation. *J. Chem. Phys.* **2013**, *138*, 104102.
- (6) Vogler, S.; Ludwig, M.; Maurer, M.; Ochsenfeld, C. Low-scaling first-order properties within second-order Møller-Plesset perturbation theory using Cholesky decomposed density matrices. *J. Chem. Phys.* **2017**, *147*, 024101.
- (7) Vogler, S.; Dietschreit, J. C. B.; Peters, L. D. M.; Ochsenfeld, C. Important components for accurate hyperfine coupling constants: electron correlation, dynamic contributions, and solvation effects. *Mol. Phys.* **2020**, *118*, e1772515.
- (8) 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.
- (9) Gauss, J. Effects of electron correlation in the calculation of nuclear magnetic resonance chemical shifts. *J. Chem. Phys.* **1993**, *99*, 3629–3643.
- (10) 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.
- (11) 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.
- (12) Glasbrenner, M.; Vogler, S.; Ochsenfeld, C. 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.* **2021**, *155*, 224107.
- (13) Kossmann, S.; Kirchner, B.; Neese, F. Performance of modern density functional theory for the prediction of hyperfine structure: Meta-GGA and double hybrid functionals. *Mol. Phys.* **2007**, *105*, 2049–2071.
- (14) Sæbø, S.; Pulay, P. Local configuration interaction: An efficient approach for larger molecules. *Chem. Phys. Lett.* **1985**, *113*, 13–18.
- (15) Sæbø, S.; Pulay, P. Local treatment of electron correlation. *Annu. Rev. Phys. Chem.* **1993**, *44*, 213–236.
- (16) El Azhary, A.; Rauhut, G.; Pulay, P.; Werner, H.-J. Analytical energy gradients for local second-order Møller-Plesset perturbation theory. *J. Chem. Phys.* **1998**, *108*, 5185.

- (17) Schütz, M.; Werner, H. J.; Lindh, R.; Manby, F. R. Analytical energy gradients for local second-order Møller-Plesset perturbation theory using density fitting approximations. *J. Chem. Phys.* **2004**, *121*, 737.
- (18) Pinski, P.; Neese, F. Communication: Exact analytical derivatives for the domain-based local pair natural orbital MP2 method (DLPNO-MP2). *J. Chem. Phys.* **2018**, *148*, 031101.
- (19) Pinski, P.; Neese, F. Analytical gradient for the domain-based local pair natural orbital second order Møller-Plesset perturbation theory method (DLPNO-MP2). *J. Chem. Phys.* **2019**, *150*, 164102.
- (20) 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.
- (21) 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.
- (22) 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*, 054102.
- (23) Kristensen, K.; Jørgensen, P.; Jansik, B.; Kjærgaard, T.; Reine, S. Molecular gradient for second-order Møller-Plesset perturbation theory using the divide-expand-consolidate (DEC) scheme. *J. Chem. Phys.* **2012**, *137*, 114102.
- (24) Bykov, D.; Kristensen, K.; Kjærgaard, T. The molecular gradient using the divide-expand-consolidate resolution of the identity second-order Møller-Plesset perturbation theory: The DEC-RI-MP2 gradient. *J. Chem. Phys.* **2016**, *145*, 024106.
- (25) Schweizer, S.; Doser, B.; Ochsenfeld, C. An atomic orbital-based reformulation of energy gradients in second-order Møller-Plesset perturbation theory. *J. Chem. Phys.* **2008**, *128*, 154101.
- (26) Whitten, J. L. Coulombic potential energy integrals and approximations. *J. Chem. Phys.* **1973**, *58*, 4496.
- (27) Dunlap, B. I.; Connolly, J. W.; Sabin, J. R. On some approximations in applications of  $X\alpha$  theory. *J. Chem. Phys.* **1979**, *71*, 3396.
- (28) Vahtras, O.; Almlöf, J.; Feyereisen, M. W. Integral approximations for LCAO-SCF calculations. *Chem. Phys. Lett.* **1993**, *213*, 514–518.
- (29) 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.
- (30) Vogler, S.; Savasi, G.; Ludwig, M.; Ochsenfeld, C. Selected-Nuclei Method for the Computation of Hyperfine Coupling Constants within Second-Order Møller-Plesset Perturbation Theory. *J. Chem. Theory Comput.* **2018**, *14*, 3014–3024.
- (31) Parrish, R. M.; Hohenstein, E. G.; Martínez, T. J.; Sherrill, C. D. Tensor hypercontraction. II. Least-squares renormalization. *J. Chem. Phys.* **2012**, *137*, 224106.
- (32) Parrish, R. M.; Hohenstein, E. G.; Schunck, N. F.; Sherrill, C. D.; Martínez, T. J. Exact tensor hypercontraction: A universal technique for the resolution of matrix elements of local finite-range  $N$ -body potentials in many-body quantum problems. *Phys. Rev. Lett.* **2013**, *111*, 1–5.
- (33) 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.
- (34) Song, C.; Martínez, T. J. Atomic orbital-based SOS-MP2 with tensor hypercontraction. II. Local tensor hypercontraction. *J. Chem. Phys.* **2017**, *146*, 034104.
- (35) Matthews, D. A. A critical analysis of least-squares tensor hypercontraction applied to MP3. *J. Chem. Phys.* **2021**, *154*, 134102.
- (36) 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.* **2020**, *16*, 243–263.
- (37) 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*, 044103.
- (38) 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.
- (39) Song, C.; Martínez, T. J. Analytical gradients for tensor hypercontracted MP2 and SOS-MP2 on graphical processing units. *J. Chem. Phys.* **2017**, *147*, 161723.
- (40) Matthews, D. A. Improved Grid Optimization and Fitting in Least Squares Tensor Hypercontraction. *J. Chem. Theory Comput.* **2020**, *16*, 1382–1385.
- (41) Bangerter, F. H.; Glasbrenner, M.; Ochsenfeld, C. Low-Scaling Tensor Hypercontraction in the Cholesky Molecular Orbital Basis Applied to Second-Order Møller-Plesset Perturbation Theory. *J. Chem. Theory Comput.* **2021**, *17*, 211–221.
- (42) Shenvi, N.; Van Aggelen, H.; Yang, Y.; Yang, W. Tensor hypercontracted ppRPA: Reducing the cost of the particle-particle random phase approximation from  $O(r6)$  to  $O(r4)$ . *J. Chem. Phys.* **2014**, *141*, 024119.
- (43) Song, C.; Martínez, T. J. Reduced scaling CASPT2 using supporting subspaces and tensor hyper-contraction. *J. Chem. Phys.* **2018**, *149*, 044108.
- (44) Hohenstein, E. G.; Kokkila, S. I.; 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.
- (45) Parrish, R. M.; Sherrill, C. D.; Hohenstein, E. G.; Kokkila, S. I.; Martínez, T. J. Communication: Acceleration of coupled cluster singles and doubles via orbital-weighted least-squares tensor hypercontraction. *J. Chem. Phys.* **2014**, *140*, 181102.
- (46) Schutski, R.; Zhao, J.; Henderson, T. M.; Scuseria, G. E. Tensor-structured coupled cluster theory. *J. Chem. Phys.* **2017**, *147*, 184113.
- (47) Hohenstein, E. G.; Kokkila, S. I.; Parrish, R. M.; Martínez, T. J. Tensor hypercontraction equation-of-motion second-order approximate coupled cluster: Electronic excitation energies in  $O(N^4)$  time. *J. Phys. Chem. B* **2013**, *117*, 12972–12978.
- (48) Song, C.; Martínez, T. J.; Neaton, J. B. A diagrammatic approach for automatically deriving analytical gradients of tensor hypercontracted electronic structure methods. *J. Chem. Phys.* **2021**, *155*, 024108.
- (49) Almlöf, J. Elimination of energy denominators in Møller-Plesset perturbation theory by a Laplace transform approach. *Chem. Phys. Lett.* **1991**, *181*, 319.
- (50) Häser, M.; Almlöf, J. Laplace transform techniques in Møller-Plesset perturbation theory. *J. Chem. Phys.* **1992**, *96*, 489.
- (51) Häser, M. Møller-Plesset (MP2) perturbation theory for large molecules. *Theor. Chim. Acta* **1993**, *87*, 147–173.
- (52) Handy, N. C.; Schaefer, H. F. On the evaluation of analytic energy derivatives for correlated wave functions. *J. Chem. Phys.* **1984**, *81*, 5031–5033.
- (53) Beer, M.; Ochsenfeld, C. Efficient linear-scaling calculation of response properties: Density matrix-based Laplace-transformed coupled-perturbed self-consistent field theory. *J. Chem. Phys.* **2008**, *128*, 221102.
- (54) 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.
- (55) 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.
- (56) 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.
- (57) 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.

(58) 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.

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

(60) Kussmann, J.; Ochsenfeld, C. Preselective screening for linear-scaling exact exchange-gradient calculations for graphics processing units and general strong-scaling massively parallel calculations. *J. Chem. Theory Comput.* **2015**, *11*, 918–922.

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

(62) 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.

(63) Pulay, P. Convergence acceleration of iterative sequences. the case of scf iteration. *Chem. Phys. Lett.* **1980**, *73*, 393–398.

(64) 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.

(65) Neese, F. The ORCA program system. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 73–78.

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

(67) Jakobsen, P.; Jensen, F. Probing basis set requirements for calculating hyperfine coupling constants. *J. Chem. Phys.* **2019**, *151*, 174107.