The Journal of Chemical Physics

Implementation of the iterative triples model CC3 for excitation energies using pair natural orbitals and Laplace transformation techniques

Cite as: J. Chem. Phys. **153**, 034109 (2020); https://doi.org/10.1063/5.0012597
Submitted: 02 May 2020 . Accepted: 22 June 2020 . Published Online: 16 July 2020

Marius S. Frank, Gunnar Schmitz, and Christof Hättig

View Online Export Citation CrossMark

ARTICLES YOU MAY BE INTERESTED IN

The ORCA quantum chemistry program package
The Journal of Chemical Physics **152**, 224108 (2020); https://doi.org/10.1063/5.0004608

Recent developments in the PySCF program package
The Journal of Chemical Physics **153**, 024109 (2020); https://doi.org/10.1063/5.0006074

NECI: N-Electron Configuration Interaction with an emphasis on state-of-the-art stochastic methods
The Journal of Chemical Physics **153**, 034107 (2020); https://doi.org/10.1063/5.0005754

Lock-in Amplifiers up to 600 MHz
Watch



© 2020 Author(s).

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

Implementation of the iterative triples model CC3 for excitation energies using pair natural orbitals and Laplace transformation techniques

Cite as: J. Chem. Phys. 153, 034109 (2020); doi: 10.1063/5.0012597
Submitted: 2 May 2020 • Accepted: 22 June 2020 •
Published Online: 16 July 2020

View Online Export Citation CrossMark

Marius S. Frank,[1,a] Gunnar Schmitz,[2,b] and Christof Hättig[1,a]

AFFILIATIONS

[1] Lehrstuhl für Theoretische Chemie, Ruhr-Universität Bochum, D-44801 Bochum, Germany
[2] Deparment of Chemistry, Aarhus University, DK-8000 Aarhus C, Denmark

[a] **Authors to whom correspondence should be addressed:** Marius.Frank@rub.de and christof.haettig@rub.de
[b] Gunnar.Schmitz@rub.de

ABSTRACT

We present a pair natural orbital (PNO)-based implementation of CC3 excitation energies, which extends our previously published state-specific PNO ansatz for the solution of the excited state eigenvalue problem to methods including connected triple excitations. A thorough analysis of the equations for the excited state triples amplitudes is presented from which we derive a suitable state-specific triple natural orbital basis for the excited state triples amplitudes, which performs equally well for local and non-local excitations. The accuracy of the implementation is evaluated using a large and diverse test set. We find that for states with small contributions from double excitations, a T0 approximation to PNO-CC3 yields accurate results with a mean absolute error (MAE) for $T_{\mathrm{PNO}} = 10^{-7}$ in the range of 0.02 eV. However, for states with larger double excitation contributions, the T0 approximation is found to yield significantly less accurate results, while the Laplace-transformed variant of PNO-CC3 shows a uniform accuracy for singly and doubly excited states (MAE and maximum error of 0.01 eV and 0.07 eV for $T_{\mathrm{PNO}} = 10^{-7}$, respectively). Finally, we apply PNO-CC3 to the calculation of the first excited state of berenil at a S_1 minimum geometry, which is shown to be close to a conical intersection. This calculation in the aug-cc-pVTZ basis set (more than 1300 basis functions) is the largest calculation ever performed with CC3 on excitation energies.

Published under license by AIP Publishing. https://doi.org/10.1063/5.0012597

I. INTRODUCTION

State-of-the art coupled cluster (CC) models that include connected triple excitations are able to accurately reproduce experimental results with high accuracy, e.g., for reaction energies often within the chemical accuracy of ≈1 kcal/mol and excitation energies for one-electron transitions within 0.1 eV. Unfortunately, the applicability of such accurate (triple excitations including) coupled cluster methods is restricted to rather small molecular systems with canonical implementations due to high computational costs in terms of floating-point operations (FLOPs), memory, and disk space usage. For example, the costs for the widely used and highly accurate approximate triples models as coupled cluster singles and doubles with perturbative triples, CCSD(T), coupled cluster singles and doubles with perturbative triples for excitation energies, CCSDR(3), and coupled cluster singles and doubles with approximate iterative triples scale as $\mathcal{O}(\mathcal{N}^7)$ in terms of FLOPs and as $\mathcal{O}(\mathcal{N}^4)$ in terms of disk space demands with the system size $\mathcal{N}$. This steep scaling makes it impossible to apply these methods routinely to molecules with more than about 20 non-hydrogen atoms with basis sets that allow to exploit the accuracy of these methods. CC3, which has been designed for response calculations on excitation energies and frequency-dependent properties, is, due to its iterative nature and the larger number of equations that need to be solved for these properties, restricted to even smaller molecular systems.

Fortunately, during the last few years, a lot of progress has been made on approaches that avoid the inherently steep scaling of the canonical coupled cluster methods without sacrificing their accuracy. Most of these approaches take advantage of the local nature of dynamic electron correlation. This locality

Published under license by AIP Publishing

is difficult to exploit if canonical molecular orbitals (MOs) are used. Therefore, these methods usually employ localized molecular orbitals (LMOs), which allow to restrict the excitation space to local correlation domains, and are, hence, called local correlation methods.

Over the past few decades, many different flavors of local correlation methods have been proposed and shown to be efficient and accurate.[1–27] One of the most popular and promising local correlation approaches nowadays is based on the representation of the virtual orbital space by pair natural orbitals (PNOs). It has its roots in the 1970s when it was introduced by Meyer[28] and refined by him,[29–31] Ahlrichs,[32,33] Staemmler,[32,34] and Kutzelnigg,[32] and co-workers. About a decade ago, the PNO approach has been revived and since then refined[35–40] by Neese and co-workers who recognized that the bottleneck of the early PNO-based methods—the integral transformations to the PNO basis—can be circumvented using the resolution-of-the-identity (RI) approximation.[41,42] Within the PNO approximation, local correlation domains are obtained by diagonalization of approximate model densities. The eigenvalues (occupation numbers) are used to select the most important orbitals for the description of the electron correlation of a LMO pair *ij*. Restriction of the correlation domains to PNOs with eigenvalues above a given threshold T_{PNO} leads to a size of the virtual spaces, which increases eventually only linearly with the system size.

Highly efficient, linear scaling, and accurate PNO-based implementations of single reference methods, such as MP2, CCSD, and CCSD(T) and their explicitly correlated[35–40,43–49] counterparts, as well as of the multireference methods NEVPT2[50] and CASPT2[51] for ground state (GS) energy calculations have already been presented by different groups. Most of these implementations make use of projected atomic orbitals (PAOs)[52–57] or orbital specific virtuals (OSVs)[44,47–49] as an intermediate basis during the PNO construction in order to achieve a linear cost-scaling. First-order molecular properties and gradients (geometry optimizations) using PNOs have recently also become available.[58–60]

Excitation energies and response properties in a truncated PNO space are an important next step. However, the description of excited states (ESs) or other non-local optical properties is a special challenge for local correlation methods since the excitation process or the response to external fields is often non-local. The development of PAO-based response methods for models beyond second order,[61,62] for which an explicit treatment and storage of doubles amplitudes cannot longer be avoided, has been given up. This is due to the fact that huge correlation domains that grow with the number of excited states were needed, which made these approaches rather inefficient.[63,64] In addition, methods based on natural orbitals (NOs)[65] and natural transition orbitals (NTOs),[66] which have recently become an active field of research, lack yet either general applicability or efficiency.

A novel route to tackle the excited state problem with local correlation methods has been recently proposed by Dutta *et al.*[67] They proposed to solve the ground state CC equations in the PNO basis and to transform the ground-state amplitudes and other intermediates afterward back to the canonical MO basis to solve the excited state eigenvalue problem. Combining this approach with the similarity-transformed equation of motion CCSD (STEOM-CCSD) leads to formal $\mathcal{O}(\mathcal{N}^5)$-scaling costs. More recently, Dutta *et al.* presented (ground state) PNO-based implementations of CCSD ionization potentials[68] (IPs) and electron affinities[69] (EAs), which are necessary ingredients for lower-scaling variants of STEOM-CCSD. However, besides the advantage of having a common PNO basis and, hence, being able to use well-established multistate Block-Davidson algorithms for the solution of the excited state eigenvalue problem, a major disadvantage of this approach is that it and also its local variants inherit the difficulties of STEOM-CCSD,[70] e.g., the selection of a basis of IP and EA states and difficulties to generalize the approach to methods with connected triples such as CCSDR(T), CCSDR(3),[71] or CC3.[72]

An alternative route, which also bears the advantage of being able to use standard multistate solvers that, compared to single state (i.e., one-state-at-a-time) solvers, are more robust, has been developed by Peng *et al.*[73] They use state-averaged PNOs obtained from densities that are derived from approximate configuration interaction singles with perturbative doubles [CIS(D)] calculations and averaged over a certain number of states. The wavefunctions of all excited states are then restricted to the same local correlation domains obtained from this density. Even though giving accurate excitation energies, the computational costs increase for this approach, in particular, for extended systems with several chromophores not only with the size of individual chromophores but also with the number of states included.

Hättig and Helmich, in contrast, proposed to use state-specific PNOs for the solution of the excited state eigenvalue problem. Here, a separate PNO basis is constructed for each excited state from approximate CIS(D) densities. The advantage of this approach is that the PNO basis adapts itself to the nature of the excited state and is, hence, as compact as possible for each state. Furthermore, the demands in computational resources such as FLOPs, memory, and disk space scale linearly with the size of the chromophore. The state-specific PNO approach is in our view the most promising local correlation approach for excited state calculations on extended systems, in particular, if they contain multiple chromophores. It has already been proven to be efficient and accurate for CC2,[74] ADC(2)-x,[75] and CCSD.[76]

The aim of this paper is to extend the PNO-based CC response approach to models including connected triples. We use for this approach, in the current work, the coupled cluster singles and doubles with approximate triples model CC3[72,77] together with state-specific PNOs.[78] This is motivated by its simplicity and its thorough grounding in perturbation theory. We did not consider here perturbative triples corrections even though their accuracy is similar to CC3 since they involve additional computational steps and, with that, pitfalls when implemented with PNOs. The most important feature that a model for excitation energies with connected triples should have is a systematic improvement upon CCSD excitation energies for singly as well as for doubly excited states. In the following, the concept of triples natural orbitals (TNOs) and the selection scheme for LMO triples will be extended to excited states. After discussing the formal theory, the accuracy and the efficiency of the implementation will be evaluated. Finally, we show for a real-life example how PNO-CC3 extends the possibilities of coupled cluster response calculations.

Published under license by AIP Publishing

II. THEORY AND IMPLEMENTATION

A. The CC3 model

In single-reference coupled cluster theory, the wavefunction is parameterized as

$$|\mathrm{CC}\rangle = e^{T}|\mathrm{HF}\rangle, \tag{1}$$

where $|\mathrm{HF}\rangle$ denotes the reference determinant (usually from a Hartree–Fock calculation) and $T = T_1 + T_2 + \cdots$ is the cluster operator, where $T_i = \sum_{\mu_i} t_{\mu_i} \tau_{\mu_i}$ are the contributions from the ith excitation class with the excitation operators τ_{μ_i} and amplitudes t_{μ_i}.

The CC3 model[72] is an iterative approximation to coupled cluster singles, doubles, and triples (CCSDT), where only the leading nonvanishing order contributions from the connected triple excitations T_3 are kept. However, in contrast to approximate triples models for ground-state energies, the singles amplitudes in CC3 are treated as zeroth-order parameters due to their special role in response theory.[72] The residuals of the CC3 cluster equations, which should vanish at the converged solution, are[72]

$$\Omega_{\mu_1} = \langle \mu_1 | e^{-T_2} \hat{H} e^{T_2} | \mathrm{HF}\rangle + \langle \mu_1 | [\hat{H}, T_3] | \mathrm{HF}\rangle = 0, \tag{2}$$

$$\Omega_{\mu_2} = \langle \mu_2 | e^{-T_2} \hat{H} e^{T_2} | \mathrm{HF}\rangle + \langle \mu_2 | [\hat{H}, T_3] | \mathrm{HF}\rangle = 0, \tag{3}$$

$$\Omega_{\mu_3} = \langle \mu_3 | [F, T_3] | \mathrm{HF}\rangle + \langle \mu_3 | [[\hat{H}, T_2] | \mathrm{HF}\rangle = 0. \tag{4}$$

The energy is obtained as $E = \langle \mathrm{HF}|H|\mathrm{CC}\rangle$. Above, we have introduced $\hat{H}$ as shorthand notation for the T_1-similarity transformed Hamiltonian[79] $\hat{H} = \exp(-T_1) H \exp(T_1)$.

In the following, we denote with Δ_{μ_3} the orbital energy differences $\Delta^{ijk}_{abc} = \epsilon_i + \epsilon_j + \epsilon_k - \epsilon_a - \epsilon_b - \epsilon_c$, where indices $i, j, k \ldots$ denote the occupied molecular orbitals (MOs) and $a, b, c \ldots$ denote the virtual MOs. For canonical MOs, Eq. (4) can formally be solved as

$$t_{\mu_3} = \Delta_{\mu_3}^{-1} \langle \mu_3 | [\hat{H}, T_2] | \mathrm{HF}\rangle. \tag{5}$$

This allows us to compute and process, whenever needed, a batch of triples amplitudes t_{μ_3} on-the-fly and solve, e.g., the CC3 amplitude equations without storing a full rank-6 triples tensor, which would limit the applicability of the method by huge storage and input/output (I/O) demands. Equation (5) is similar to the equation of the perturbative triples used in CCSD(T) but employs the t_1-dressed Hamiltonian $\hat{H}$ and doubles amplitudes (in T_2) that are determined iteratively at the CC3 level.

In coupled cluster response theory, excitation energies are obtained as eigenvalues of the Jacobian $\mathbf{A}$ of the ground-state residual or vector function Ω_{μ_i},

$$A_{\mu_i \nu_j} = \frac{\partial \Omega_{\mu_i}}{\partial t_{\nu_j}}. \tag{6}$$

For CC3, the Jacobian obtained from Eqs. (2) to (4) is

$$\mathbf{A} = \begin{pmatrix} \langle \mu_1 | [\hat{H} + [\hat{H}, T_2], \tau_{\nu_1}] | \mathrm{HF}\rangle & \langle \mu_1 | [\hat{H}, \tau_{\nu_2}] | \mathrm{HF}\rangle & \langle \mu_1 | [\hat{H}, \tau_{\nu_3}] | \mathrm{HF}\rangle \\ \langle \mu_2 | [\hat{H} + [\hat{H}, T_2 + T_3], \tau_{\nu_1}] | \mathrm{HF}\rangle & \langle \mu_2 | [\hat{H} + [\hat{H}, T_2], \tau_{\nu_2}] | \mathrm{HF}\rangle & \langle \mu_2 | [\hat{H}, \tau_{\nu_3}] | \mathrm{HF}\rangle \\ \langle \mu_3 | [[\hat{H}, T_2], \tau_{\nu_1}] | \mathrm{HF}\rangle & \langle \mu_3 | [\hat{H}, \tau_{\nu_2}] | \mathrm{HF}\rangle & \langle \mu_3 | [F, \tau_{\nu_3}] | \mathrm{HF}\rangle \end{pmatrix}. \tag{7}$$

The eigenvalues are usually determined by iterative procedures that require the implementation of linear transformations of a trial vector $\mathbf{R}$ with the Jacobian,

$$\rho_{\mu_i} = \sum_{\nu_j} A_{\mu_i \nu_j} R_{\nu_j}. \tag{8}$$

With the auxiliary operators $R_i = \sum_{\mu_i} R_{\mu_i} \tau_{\mu_i}$ and the one-index transformed Hamiltonian[80] $\bar{H} = [\hat{H}, R_1]$, the linear transformations with the CC3 Jacobian can be written as

$$\rho_{\mu_1} = \langle \mu_1 | \bar{H} + [\bar{H}, T_2] + [\hat{H}, R_2] + [\hat{H}, R_3] | \mathrm{HF}\rangle, \tag{9}$$

$$\rho_{\mu_2} = \langle \mu_2 | \bar{H} + [\bar{H}, T_2 + T_3] + [\hat{H} + [\hat{H}, T_2], R_2] + [\hat{H}, R_3] | \mathrm{HF}\rangle, \tag{10}$$

$$\rho_{\mu_3} = \langle \mu_3 | [\bar{H}, T_2] + [\hat{H}, R_2] + [F, R_3] | \mathrm{HF}\rangle. \tag{11}$$

The triples part of the CC3 eigenvectors, E_{μ_i}, has to fulfill the relation[77]

$$E_{\mu_3} = (\Delta_{\mu_3} + \omega)^{-1} \langle \mu_3 | [\bar{H}, T_2] + [\hat{H}, E_2] | \mathrm{HF}\rangle, \tag{12}$$

where ω denotes the eigenvalue (excitation energy). This can be used to reformulate the eigenvalue problem for the CC3 excitation energies from a linear eigenvalue problem in the singles, doubles, and triples space into a non-linear eigenvalue problem in the singles and doubles space only. The effective Jacobi matrix for the non-linear problem is

$$\sum_{j=1}^{2} \sum_{\nu_j} A^{\mathrm{eff}}_{\mu_i \nu_j}(\omega) E_{\nu_j} = \sum_{j=1}^{2} \sum_{\nu_j} \left(A_{\mu_i \nu_j} + \sum_{\kappa_3} \frac{A_{\mu_i \kappa_3} A_{\kappa_3 \nu_j}}{\Delta_{\kappa_3} + \omega} \right) E_{\nu_j} = \omega E_{\mu_i}. \tag{13}$$

This reformulation allows us to solve the eigenvalue problem without storage of full triples vectors.

The time determining steps of the CC3 Jacobian right transformations scale for canonical implementations as $\mathcal{O}(\mathcal{N}^7)$, which restricts the applicability of them to small molecular systems. The aim of this work is to reduce this steep scaling using PNOs and TNOs, which will be described in Secs. II B–II F.

B. Ground and excited state PNOs

PNOs are linear combinations of virtual molecular orbitals

Published under license by AIP Publishing

$$|\tilde{a}_{ij}\rangle = \sum_{a} d^{ij}_{a\tilde{a}_{ij}} |a\rangle, \tag{14}$$

with expansion coefficients $d^{ij}_{a\tilde{a}_{ij}}$ that are initially obtained from the eigenvectors of the correlation contribution from a MO pair ij to the virtual–virtual block of the one-electron density,

$$\sum_{ab} d^{ij}_{a\tilde{a}_{ij}} D^{ij}_{ab} d^{ij}_{b\tilde{b}_{ij}} = n^{ij}_{\tilde{a}_{ij}} \delta_{\tilde{a}_{ij}\tilde{b}_{ij}}. \tag{15}$$

Here and in the following, PNOs will be denoted by a bar accent and the pair index will be omitted if it is clear from the context which pair a PNO belongs to. The eigenvalues $n^{ij}_{\bar{a}_{ij}}$ of the density $\mathbf{D}^{ij}$ have the meaning of occupation numbers and correlate with the importance of a given PNO for the correlation contribution to the wavefunction. If the occupied orbital space is represented by LMOs, the eigenvalues decay rapidly with the index $\bar{a}_{ij}$ and with the distance between the LMOs. Hence, it becomes practical to define a truncation threshold T_{PNO} and discard PNOs (i.e., truncate the virtual space) with $n^{ij}_{\bar{a}_{ij}} < T_{\mathrm{PNO}}$. In this way, the accuracy of a calculation can, in principle, be controlled by a simple truncation threshold.

For the ground state, the PNOs are usually constructed from approximate MP2 densities,[28]

$$D^{ij,\mathrm{gs}}_{ab} = \sum_{c} \left(\tilde{t}^{ij,(1)}_{ac} t^{ij,(1)}_{bc} + \tilde{t}^{ij,(1)}_{ca} t^{ij,(1)}_{cb} \right), \tag{16}$$

with $\tilde{t}^{ij,(1)}_{ab} = 2t^{ij,(1)}_{ab} - t^{ij,(1)}_{ba}$ and the approximate first-order amplitudes

$$t^{ij,(1)}_{ab} = \frac{(ai|bj)}{F_{ii} + F_{jj} - \epsilon_a - \epsilon_b}, \tag{17}$$

where $(ai|bj)$ denote the electron repulsion integrals (ERIs), F_{ii} denote the diagonal Fock matrix elements in the LMO basis, and ϵ_a denote orbital energies for canonical virtual MOs.

For PNOs of excited states, we use an approximate correlation contribution to the density,[78]

$$D^{ij,\mathrm{es}}_{ab} = \sum_{c} \left(\tilde{R}^{ij,(1)}_{ac} R^{ij,(1)}_{bc} + \tilde{R}^{ij,(1)}_{ca} R^{ij,(1)}_{cb} \right), \tag{18}$$

which is calculated from CIS(D)-like excited state first-order amplitudes,[78]

$$R^{ij,(1)}_{ab} = \frac{(ai|\tilde{bj})}{F_{ii} + F_{jj} - \epsilon_a - \epsilon_b + \omega} \tag{19}$$

and $\tilde{R}^{ij,(1)}_{ab} = 2R^{ij,(1)}_{ab} - R^{ij,(1)}_{ba}$. The excited-state singles-dressed integrals are computed as

$$(ai|\tilde{bj}) = \sum_{\mu\nu\kappa\lambda} P^{ij}_{ab} \left(\bar{\Lambda}^{p}_{\mu a} C_{\nu i} + C_{\mu a} \bar{\Lambda}^{h}_{\nu i} \right) C_{\kappa b} C_{\lambda j} (\mu\nu|\kappa\lambda), \tag{20}$$

where $\mathbf{C}$ denotes the LMO coefficients and $\bar{\Lambda}^{p}_{\mu a} = -\sum_i R^i_a C_{\mu i}$ and $\bar{\Lambda}^{h}_{\mu i} = \sum_a R^i_a C_{\mu a}$ are the modified MO coefficients, dressed with the singles part of the current trial vector R^i_a. The symmetrization operator P^{ij}_{ab} is defined as $P^{ij}_{ab} f^{ij}_{ab} = f^{ij}_{ab} + f^{ji}_{ba}$.

If the PNOs would be obtained from Eqs. (16) and (18) without any further approximations, the computational complexity of the PNO construction would increase with the system size as $\mathcal{O}(\mathcal{N}^5)$. To reduce this scaling, the densities are evaluated in a so-called pre-PNO basis[44] $\{\tilde{a}_{ij}\}$, which is obtained by merging the orbital specific virtual[81] (OSV) spaces $\{\tilde{a}_i\}$ and $\{\tilde{a}_j\}$ of the LMOs i and j. In this basis, the approximate first-order amplitudes are computed with a lower cost scaling as[48,74]

$$t^{ij,(1)}_{\tilde{a}_{ij}\tilde{b}_{ij}} = \frac{(\tilde{a}_{ij} i|\tilde{b}_{ij} j)}{F_{ii} + F_{jj} - \epsilon_{\tilde{a}_{ij}} - \epsilon_{\tilde{b}_{ij}}}, \tag{21}$$

$$R^{ij,(1)}_{\tilde{a}_{ij}\tilde{b}_{ij}} = \frac{(\tilde{a}_{ij} i|\tilde{\tilde{b}}_{ij} j)}{F_{ii} + F_{jj} - \epsilon_{\tilde{a}_{ij}} - \epsilon_{\tilde{b}_{ij}} + \omega}, \tag{22}$$

where $\epsilon_{\tilde{a}_{ij}}$ are the eigenvalues of the Fock matrix that is diagonalized in the pre-PNO basis.

The ground-state OSVs are obtained by diagonalizing the MP2-like densities [Eq. (16)] for the diagonal pairs. The excited-state OSVs are obtained from densities [Eq. (18)] that are calculated from modified pseudo-amplitudes[74]

$$R^{\prime ii,(1)}_{ab} = \frac{(ai|bi) + (\overline{ai|bj})}{2F_{ii} - \epsilon_a - \epsilon_b + \omega} \tag{23}$$

with integrals

$$(\overline{ai|bj}) = \sum_{\mu\nu\kappa\lambda} \Big[\bar{\Lambda}^{h}_{\nu i} \bar{\Lambda}^{h}_{\lambda j} C_{\mu a} C_{\kappa b} + \bar{\Lambda}^{p}_{\mu a} \bar{\Lambda}^{p}_{\kappa b} C_{\nu i} C_{\lambda j} + P^{ij}_{ab} \bar{\Lambda}^{p}_{\mu a} \bar{\Lambda}^{h}_{\lambda j} C_{\nu i} C_{\kappa b} \Big] (\mu\nu|\kappa\lambda) \tag{24}$$

in order to obtain OSVs that are suitable for expanding the one-index transformed integrals $(ai|\tilde{bj})$ of LMO pairs.

Furthermore, we pre-screen the LMO pairs based on approximate pair contributions to the correlation energy. The estimates for the pair correlation energies $e^{(2)}_{ij}$ for the ground state and $\omega^{(2)}_{ij}$ for the excited state are calculated in the OSV basis $\{\tilde{a}_i\}$ and $\{\tilde{b}_j\}$ as

$$e^{(2)}_{ij} = \sum_{\tilde{a}_{ij}\tilde{b}_{ij}} \frac{(\tilde{a}_i i|\tilde{b}_j j)}{F_{ii} + F_{jj} - \epsilon_{\tilde{a}_i} - \epsilon_{\tilde{b}_j}}, \tag{25}$$

$$\omega^{(2)}_{ij} = \sum_{\tilde{a}_{ij}\tilde{b}_{ij}} \frac{(\tilde{a}_i i|\tilde{\tilde{b}}_j j)}{F_{ii} + F_{jj} - \epsilon_{\tilde{a}_i} - \epsilon_{\tilde{b}_j} + \omega}. \tag{26}$$

Ground state pairs with $e^{(2)}_{ij} < \frac{1}{3} T_{pair}$ and excited state pairs with $\omega^{(2)}_{ij} < \frac{1}{3} T_{pair}$ are discarded. For a more detailed description of our PNO construction, we refer the reader to Refs. 48, 78, and 74.

C. Triples selection and TNO construction for the ground state

The construction of a triples-specific basis from approximate triples amplitudes using similar strategies as for the PNOs is—at least for approximate triples models such as CCSD(T) and CC3—not practical since the computational costs for the required ERIs and their contraction with the doubles amplitudes to approximate triples amplitudes are high. As an alternative, Riplinger *et al.*[82] proposed to restrict the ground state triples amplitudes to LMO triples ijk, where ij, ik, and jk are significant pairs. Only triples for which

Published under license by AIP Publishing

each of those pairs is part of the ground state pair list are considered to be important for the triples contribution of the CCSD(T) model. Furthermore, the triples list is further compressed by only including triples where at most one pair is a so-called weak pair. Such a weak pair is not screened out by the MP2 pair correlation estimate but also not included in the CCSD treatment for which these authors applied a tighter pair selection threshold. This scheme for the selection of non-negligible LMO triples is very similar to the one initially proposed by Schütz[18] in the context of PAO-based perturbative triples corrections.

The triples selection, as described above, can be rationalized by analyzing the explicit expression for the ground state CCSD(T) triples equations[83]

$$t_{abc}^{ijk} = -\frac{1}{\Delta_{abc}^{ijk}} P_{abc}^{ijk} \left\{ \sum_d t_{ad}^{ij}(ck|bd) - \sum_l t_{ab}^{il}(ck|lj) \right\}, \tag{27}$$

where the operator P_{abc}^{ijk} acts as $P_{abc}^{ijk} f_{abc}^{ijk} = f_{abc}^{ijk} + f_{acb}^{ikj} + f_{bac}^{jik} + f_{cab}^{kij} + f_{bca}^{jki} + f_{cba}^{kji}$. Here and in the following, we restrict the discussion to the spin-free closed-shell case. The multipole expansion of two-electron integrals

$$|(pq|rs)| \leq \frac{M_{pq}^{(0)} M_{rs}^{(0)}}{R} + \frac{M_{pq}^{(1)} M_{rs}^{(0)} + M_{pq}^{(0)} M_{rs}^{(1)}}{R^2} + \frac{M_{pq}^{(2)} M_{rs}^{(0)} + 2M_{pq}^{(1)} M_{rs}^{(1)} + M_{pq}^{(0)} M_{rs}^{(2)}}{R^3} + \mathcal{O}(R^{-4}), \tag{28}$$

where indices p, q, r, s denote general MOs, holds for (approximately) non-overlapping charge-distributions, which in local correlation treatments is a reasonable assumption for weak and distant pairs.[84] In Eq. (28), upper indices 0, 1, and 2 denote the monopoles, dipoles, and quadrupoles of the charge clouds pq and $R = |\vec{R}_{pq} - \vec{R}_{rs}|$ is the distance between the expansion centers $\vec{R}_{pq}$ and $\vec{R}_{rs}$ of these charge distributions. The $M^{(n)}$ denote the absolute values of the corresponding multipole integrals,

$$M_{pq}^{(l)} = \sqrt{\sum_m \left| \int \phi_\mu(\mathbf{r}) \phi_\nu(\mathbf{r}) r^l \sqrt{\frac{4\pi}{2n+1}} Y_{lm}(\mathbf{r}) d\mathbf{r} \right|^2}, \tag{29}$$

where Y_{lm} are the spherical harmonics of order l with $-l \leq m \leq l$. The monopole is directly related to the overlap, $M_{pq}^{(0)} = -S_{pq}$, and reduces for orthonormal orbitals to $M_{pq}^{(0)} = -\delta_{pq}$.

Applying the multipole expansion to the two terms in Eq. (27) separately gives

$$\left| \sum_d t_{ad}^{ij}(ck|bd) \right| \leq \sum_d |t_{ad}^{ij}| \left(\frac{M_{ck}^{(1)} M_{bd}^{(0)}}{R^2} + \frac{M_{ck}^{(2)} M_{bd}^{(0)} + 2M_{ck}^{(1)} M_{bd}^{(1)}}{R^3} \right) + \mathcal{O}(\mathcal{R}^{-7}), \tag{30}$$

$$\left| \sum_l t_{ab}^{il}(ck|lj) \right| \leq \sum_l |t_{ab}^{il}| \left(\frac{M_{ck}^{(1)} M_{lj}^{(0)}}{R^2} + \frac{M_{ck}^{(2)} M_{lj}^{(0)} + 2M_{ck}^{(1)} M_{lj}^{(1)}}{R^3} \right) + \mathcal{O}(\mathcal{R}^{-7}), \tag{31}$$

where, on the right-hand side, it was used that, for weak or distant pairs, the amplitudes t_{ab}^{ij} decay as $1/R_{ij}^3$ with the distance between the LMOs i and j, as it is apparent from the integrals $(ai|bj)$, which enter the MP1 amplitude equations and only contain dipole–dipole interactions in the lowest non-vanishing order. In Eqs. (30) and (31), R refers to the distance between $\vec{R}_{ck}$ and, respectively, $\vec{R}_{bd}$ and $\vec{R}_{lj}$ and $\mathcal{R} = \min(R_{ij}, R)$.

These equations suggest that the triples amplitudes would have rather long-ranging non-negligible contributions since the leading terms of the multipole expansion of the integrals decay only as $\mathcal{O}(R^{-2})$. This would imply that only two LMOs of a triple ijk have to be close to each other, those for which the indices appear in the amplitude t_{ab}^{ij} or in a charge distribution with a common index with the amplitude $t_{ab}^{il} M_{lj}^{(n)}$. However, the third LMO could be rather far away from both of them as the contribution to the nominator of the triples amplitude would decay only as $1/R^2$ with the distance between the charge centers in the ERI.

If both terms in Eq. (27) are combined, the long-range $\mathcal{O}(R^{-2})$ contributions cancel out each other and the dipole–dipole contribution of the integrals, which decays as $\mathcal{O}(R^{-3})$, becomes the leading non-vanishing term in the multipole expansion of the triples amplitudes,

$$\left| \sum_d t_{ad}^{ij}(ck|bd) - \sum_l t_{ab}^{il}(ck|lj) \right| \leq 2 \sum_d |t_{ad}^{ij}| \frac{M_{ck}^{(1)} M_{bd}^{(1)}}{R^3} + 2 \sum_l |t_{ab}^{il}| \frac{M_{ck}^{(1)} M_{lj}^{(1)}}{R^3} + \mathcal{O}(\mathcal{R}^{-7}). \tag{32}$$

From Eq. (32), it can, thus, be concluded that the three LMOs of a triple ijk are in the long-range regime coupled by dipole interactions. Hence, for a non-vanishing contribution from a strong pair ij for which the contractions $\sum_d t_{ab}^{ij} \phi_b(\mathbf{r}) \phi_d(\mathbf{r})$ and $\sum_l t_{ab}^{il} \phi_l(\mathbf{r}) \phi_j(\mathbf{r})$ are large, the LMO k must be close to either i or j and, hence, in addition to the pair ij, one of the pairs ik or jk has to be a strong pair.

This forms the basis for the pre-screening of LMO triples used in local correlation methods.[18,82] The PNO-CCSD(T) implementation,[46] which we used as a starting point for the current developments, employed as an estimate for the magnitude of the correlation energy contribution from a LMO triple ijk,

$$E_{\text{guess}}^{ijk} = 3.5 \left[\frac{(e_1^{ijk} - e_2^{ijk})(e_2^{ijk} - e_3^{ijk})(e_3^{ijk} - e_1^{ijk})}{2\left\{ (e_2^{ijk} - e_3^{ijk}) \log e_1^{ijk} + (e_3^{ijk} - e_1^{ijk}) \log e_2^{ijk} + (e_1^{ijk} - e_2^{ijk}) \log e_3^{ijk} \right\}} \right]^{0.4} \tag{33}$$

Published under license by AIP Publishing

with $e_x^{ijk} = e^{ij}e^{ik}\delta_{x,1} + e^{ij}e^{jk}\delta_{x,2} + e^{ik}e^{jk}\delta_{x,3}$ calculated from the CCSD pair energies if all three of the e_x^{ijk} differ by at least by 10^{-12}. If only two of these products of pair energies differ, the importance of a triple is estimated from these two products as

$$E_{\text{guess}}^{ijk} = 3.5\left[\frac{e_1^{ijk} - e_2^{ijk}}{\log e_1^{ijk} - \log e_2^{ijk}}\right]^{0.9}. \tag{34}$$

Triples with E_{guess}^{ijk} below a threshold T_{trip} are discarded. In the meantime, we have replaced the CCSD pair energies by the estimate for the MP2 correlation energies $e_{ij}^{(2)}$ from Eq. (25) to avoid artificial differences between spin-orbital and spin-free calculations. Since E_{guess}^{ijk} is a rather empirical estimate for the size of the triples energy, the selection threshold T_{trip} is chosen rather tight, relative to the TNO selection threshold T_{TNO} (*vide infra*) to not diminish the accuracy.

For the selected triples *ijk*, an initial pre-TNO basis $\{\bar{a}_{ijk}\}$ is constructed by merging the ground state OSV sets of the LMOs *i*, *j*, and *k*. Then, a density for the triple *ijk* is calculated by averaging the approximate MP2 pair densities [cf. Eqs. (16) and (17)] for the three pairs,

$$D_{\bar{a}\bar{b}}^{ijk,\text{gs}} = \frac{1}{3}\left(S_{\bar{a}\bar{a}}^{ijk,\,ij} D_{\bar{a}\bar{b}}^{ij,\text{gs}} S_{\bar{b}\bar{b}}^{ij,ijk} + S_{\bar{a}\bar{a}}^{ijk,\,ik} D_{\bar{a}\bar{b}}^{ik,\text{gs}} S_{\bar{b}\bar{b}}^{ik,ijk} + S_{\bar{a}\bar{a}}^{ijk,\,jk} D_{\bar{a}\bar{b}}^{jk,\text{gs}} S_{\bar{b}\bar{b}}^{jk,ijk}\right). \tag{35}$$

In the above equation, the pre-TNO/PNO overlap matrix

$$S_{\bar{a}\bar{a}}^{ijk,\,ij} = \sum_c d_{c\bar{a}}^{ijk} d_{c\bar{a}}^{ij} \tag{36}$$

with the PNO and pre-TNO coefficients $d_{c\bar{a}}^{ij}$ and $d_{c\bar{a}}^{ijk}$ was used to transform the densities from the PNO to the pre-TNO basis. Similar as in the PNO construction, the initial TNO coefficients are obtained from the eigenvectors of this density and only the TNOs for eigenvalues $\eta_{\bar{a}_{ijk}}^{ijk}$ above a threshold T_{TNO} are kept. Thereafter, the selected TNOs for a triple *ijk* are transformed to a semi-canonical basis in which the Fock matrix is diagonal.

D. Triples selection and TNOs for excited states

To arrive at a reasonable selection of the triples for excited states, we follow a similar multipole analysis-based strategy as used in Sec. II C for the ground-state triples. For simplicity, we assume that the reference wavefunction fulfills the Brillouin condition so that the ground state singles t_{μ_1} vanish in the first-order and can be ignored for qualitative considerations. Before discussing the amplitude equations, we analyze the multipole expansion of the integrals dressed with the singles part of the eigenvector E_a^i that appear in the nominators of the triples equations,

$$(pq|\bar{r}s) = P_{pr}^{qs}\sum_t\left[E_t^q(pt|rs) - E_p^t(tq|rs)\right], \tag{37}$$

with the understanding that E_q^t is nonzero only if *t* is an index of an occupied and *p* is an index of a virtual MO. A separate multipole expansion of the two terms in Eq. (37) gives

$$\left|\sum_t E_t^q(pt|rs)\right| \le \sum_t |E_t^q|\left(\frac{M_{pt}^{(0)}M_{rs}^{(0)}}{R} + \frac{M_{pt}^{(1)}M_{rs}^{(0)} + M_{pt}^{(0)}M_{rs}^{(1)}}{R^2} + \frac{M_{pt}^{(2)}M_{rs}^{(0)} + 2M_{pt}^{(1)}M_{rs}^{(1)} + M_{pt}^{(0)}M_{rs}^{(2)}}{R^3}\right) + \mathcal{O}(R^{-4}), \tag{38}$$

$$\left|\sum_t E_p^t(tq|rs)\right| \le \sum_t |E_p^t|\left(\frac{M_{tq}^{(0)}M_{rs}^{(0)}}{R'} + \frac{M_{tq}^{(1)}M_{rs}^{(0)} + M_{tq}^{(0)}M_{rs}^{(1)}}{R'^2} + \frac{M_{tq}^{(2)}M_{rs}^{(0)} + 2M_{tq}^{(1)}M_{rs}^{(1)} + M_{tq}^{(0)}M_{rs}^{(2)}}{R'^3}\right) + \mathcal{O}(R'^{-4}), \tag{39}$$

where R and R' denote the distances between the center of the orbital product *rs* and the centers of *pt* and *tq*, respectively. The first contribution is non-zero only if *q* is an occupied MO contributing to E_{μ_1}, and the second contribution is non-zero only if *p* is a virtual MO contributing to E_{μ_1}. Long-ranging charge–charge terms arise for non-local transitions if *p* is a virtual MO located at or close to the electron-accepting chromophore and, at the same time, *q* is an occupied MO at or close to the electron-donating chromophore. If the two terms are considered together, the long-ranging $\mathcal{O}(R^{-1})$ and $\mathcal{O}(R'^{-1})$ contributions eventually cancel out each other if both distances are large compared to the distance between the centers of *pt* and *tq* and the dipole–monopole interactions become the leading non-vanishing terms in the multipole expansion of the dressed integrals,

$$\left|(pq|\bar{r}s)\right| \le P_{pr}^{qs}\sum_t |E_t^q|\left(\frac{M_{pt}^{(1)}M_{rs}^{(0)}}{R^2} + \frac{M_{pt}^{(2)}M_{rs}^{(0)} + 2M_{pt}^{(1)}M_{rs}^{(1)}}{R^3}\right) + P_{pr}^{qs}\sum_t |E_p^t|\left(\frac{M_{tq}^{(1)}M_{rs}^{(0)}}{R^2} + \frac{M_{tq}^{(2)}M_{rs}^{(0)} + 2M_{tq}^{(1)}M_{rs}^{(1)}}{R^3}\right) + \mathcal{O}(R^{-4}). \tag{40}$$

For a long-range charge-transfer (CT) state where the occupied and virtual natural transition orbitals (NTOs) are located on separate chromophores, this will, however, only happen for rather large distances.

For the dressed integrals $(ai|\bar{b}j)$ appearing in the nominator of excited state doubles amplitudes of transitions dominated by one-electron excitations, the undressed orbital pair is always charge free, and hence, $\mathcal{O}(\mathcal{R}^{-3})$-decaying dipole–dipole interactions become the lowest non-vanishing order contributions to the multipole expansion for distances fulfilling the above condition,

$$\left|(ai|\bar{b}j)\right| \le 2P_{ab}^{ij}\left[\sum_c |E_c^i|\frac{M_{ac}^{(1)}M_{bj}^{(1)}}{R^3} + \sum_k |E_a^k|\frac{M_{ki}^{(1)}M_{bj}^{(1)}}{R^3}\right] + \mathcal{O}(R^{-4}) \tag{41}$$

$$\approx 2P_{ab}^{ij}\left[|\beta_{i\sigma^o}|\frac{M_{a\sigma^v}^{(1)}M_{bj}^{(1)}}{R^3} + |\beta_{a\sigma^v}|\frac{M_{\sigma^o i}^{(1)}M_{bj}^{(1)}}{R^3}\right] + O(R^{-4}). \tag{42}$$

In Eq. (42), we assumed, for simplicity, that E_{μ_1} is dominated by a single pair of virtual $|\sigma^v\rangle$ and occupied NTOs $|\sigma^o\rangle$ with the corresponding coefficients $\beta_{a\sigma^v}$ and $\beta_{i\sigma^o}$, i.e., $E_a^i \approx \beta_{a\sigma^v}\beta_{i\sigma^o}$.

Published under license by AIP Publishing

For local excitations where occupied and virtual NTOs are close to each other, only pairs ij contribute that are part of or close to the same chromophore, and the corresponding PNOs are also close to this chromophore. However, for charge-transfer excitations where the occupied and virtual NTOs are spatially separated, two different types of pairs ij are significant: those where both LMOs i and j are close to the occupied NTO. For them, the doubles amplitude matrices E^{ij}_{ab} have a rank of 1 (in the following called rank-1 pairs[74]), and one of their PNOs is close to the virtual NTO, while the other is close to the occupied NTO. For the second class of pairs, one LMO is close to the occupied NTO, while the other is close to the virtual NTO. Their doubles amplitude matrices have (for small molecular systems) a rank that can be as large as the number of virtual orbitals (in the following called rank-v pairs[74]), and all of their PNOs are located close to the virtual NTO.

We can now apply this to arrive at a scheme for selecting non-negligible LMO triples for the excited state eigenvectors. For this, we start from the (closed-shell) expression for the excited-state triples amplitudes in the CCSDR(T) model,

$$R^{ijk}_{abc} = \frac{1}{\Delta^{ijk}_{abc} + \omega} P^{ijk}_{abc} \Big\{ \sum_{d\neq b} R^{ij}_{ad}(ck|bd) - \sum_{l\neq j} R^{il}_{ab}(ck|lj) + R^{ij}_{ab}[(ck|bb) - (ck|jj)] + \sum_{d\neq b} t^{ij}_{ad}(ck|\bar{b}d) - \sum_{l\neq j} t^{il}_{ab}(ck|\bar{l}j) + t^{ij}_{ab}[(ck|\bar{b}b) - (ck|\bar{j}j)] \Big\}. \tag{43}$$

TABLE I. Multipole expansion of the individual contributions to the excited state triples amplitudes. R_1 refers to the distance between the centers *lk* and *bd*, R_1' refers to the distance between *ce* and *bd*, R_1'' refers to the distance between *ck* and *ld*, R_2 refers to the distance between *mk* and *lj*, R_2' refers to the distance between *cd* and *lj*, R_2'' refers to the distance between *ck* and *le*, R refers to the distance between *ck* and *bd*, and R' refers to the distance between *ck* and *lj*.

Excited state doubles contributions

$$\left|\sum_d R^{ij}_{ad}(ck|bd)\right| \le \sum_d |R^{ij}_{ad}| \left(\frac{M^{(1)}_{ck} M^{(0)}_{bd}}{R^2} + \frac{M^{(2)}_{ck} M^{(0)}_{bd} + 2M^{(1)}_{ck} M^{(1)}_{bd}}{R^3} \right) + \mathcal{O}(\mathcal{R}^{-7}),$$

$$\left|\sum_l R^{il}_{ab}(ck|lj)\right| \le \sum_l |R^{il}_{ab}| \left(\frac{M^{(1)}_{ck} M^{(0)}_{lj}}{R'^2} + \frac{M^{(2)}_{ck} M^{(0)}_{lj} + 2M^{(1)}_{ck} M^{(1)}_{lj}}{R'^3} \right) + \mathcal{O}(\mathcal{R}^{-7})$$

Contributions to $t^{ij}_{ad}(ck|\bar{b}d)$

$$\left|\sum_d t^{ij}_{ad} \sum_l R^l_c(lk|bd)\right| \le \sum_d |t^{ij}_{ad}| \sum_l |R^l_c| \left(\frac{M^{(0)}_{lk} M^{(0)}_{bd}}{R_1} + \frac{M^{(1)}_{lk} M^{(0)}_{bd} + M^{(0)}_{lk} M^{(1)}_{bd}}{R_1^2} + \frac{M^{(2)}_{lk} M^{(0)}_{bd} + 2M^{(1)}_{lk} M^{(1)}_{bd} + M^{(0)}_{lk} M^{(2)}_{bd}}{R_1^3} \right) + \mathcal{O}(\mathcal{R}^{-7})$$

$$\left|\sum_d t^{ij}_{ad} \sum_e R^k_e(ce|bd)\right| \le \sum_d |t^{ij}_{ad}| \sum_e |R^k_e| \left(\frac{M^{(0)}_{ce} M^{(0)}_{bd}}{R_1'} + \frac{M^{(1)}_{ce} M^{(0)}_{bd} + M^{(0)}_{ce} M^{(1)}_{bd}}{R_1'^2} + \frac{M^{(2)}_{ce} M^{(0)}_{bd} + 2M^{(1)}_{ce} M^{(1)}_{bd} + M^{(0)}_{ce} M^{(2)}_{bd}}{R_1'^3} \right) + \mathcal{O}(\mathcal{R}^{-7})$$

$$\left|\sum_d t^{ij}_{ad} \sum_l R^l_b(ck|ld)\right| \le 2 \sum_d |t^{ij}_{ad}| \sum_l |R^l_b| \frac{M^{(1)}_{ck} M^{(1)}_{ld}}{R_1''^3} + \mathcal{O}(\mathcal{R}^{-7})$$

Contributions to $t^{il}_{ab}(ck|\bar{l}j)$

$$\left|\sum_l t^{il}_{ab} \sum_m R^m_c(mk|lj)\right| \le \sum_l |t^{il}_{ab}| \sum_m |R^m_c| \left(\frac{M^{(0)}_{mk} M^{(0)}_{lj}}{R_2} + \frac{M^{(1)}_{mk} M^{(0)}_{lj} + M^{(0)}_{mk} M^{(1)}_{lj}}{R_2^2} + \frac{M^{(2)}_{mk} M^{(0)}_{lj} + 2M^{(1)}_{mk} M^{(1)}_{lj} + M^{(0)}_{mk} M^{(2)}_{lj}}{R_2^3} \right) + \mathcal{O}(\mathcal{R}^{-7})$$

$$\left|\sum_l t^{il}_{ab} \sum_d R^k_d(cd|lj)\right| \le \sum_l |t^{il}_{ab}| \sum_d |R^k_d| \left(\frac{M^{(0)}_{cd} M^{(0)}_{lj}}{R_2'} + \frac{M^{(1)}_{cd} M^{(0)}_{lj} + M^{(0)}_{cd} M^{(1)}_{lj}}{R_2'^2} + \frac{M^{(2)}_{cd} M^{(0)}_{lj} + 2M^{(1)}_{cd} M^{(1)}_{lj} + M^{(0)}_{cd} M^{(2)}_{lj}}{R_2'^3} \right) + \mathcal{O}(\mathcal{R}^{-7})$$

$$\left|\sum_l t^{il}_{ab} \sum_e R^j_e(ck|le)\right| \le 2 \sum_l |t^{ij}_{ad}| \sum_e |R^j_e| \frac{M^{(1)}_{ck} M^{(1)}_{le}}{R_2''^3} + \mathcal{O}(\mathcal{R}^{-7})$$

Published under license by AIP Publishing

The contributions obtained by applying the multipole expansion to the individual contributions to the excited state triples amplitudes are summarized in Table I. According to the previous discussion for the integrals $(pq|rs)$, it follows that the potentially very long-ranging contributions (monopole–monopole interactions) to $t^{ij}_{ab}[(ck|bb) - (ck|jj)]$ cancel each other out if the distances R_1 and R_2 as well as R'_1 and R'_2 are large compared to the distance between the orbitals b and j. This happens on the same length scale on which the $\mathcal{O}(\mathcal{R}^{-2})$ contributions become negligible in the nominators for the ground-state triples amplitudes.

On the other hand, the dipole–monopole contributions from $t^{ij}_{ad}[(ck|bd)$ and $t^{il}_{ab}(ck|lj)]$ do not cancel, in contrast to the ground state. However, these rather long-ranging contributions from the integrals do, for all of these terms, decay with the inverse of the distance between the occupied or virtual NTO and a LMO, which is part of a triple ijk. Since the leading term in the multipole expansion of the integrals $(ai|bj)$ that are used for the construction of the excited state PNOs is of this dipole–monopole type, we conclude that the LMO indices of the ground-state doubles amplitudes and the LMO k appearing in the integrals have to be connected by an excited state pair relation, i.e., one of the pair ik and jk has to be a weak and the other one has to be a strong excited state pair, if ijk should be a significant excited state triple.

The analysis of the long-ranging contributions from the excited state doubles amplitudes $R^{ij}_{ab}[(ck|bb) - (ck|jj)]$, which decay with $\mathcal{O}(\mathcal{R}^{-2})$, is more complicated since one has to distinguish for ij rank-1 or rank-v pairs, and the interchange of the indices ai and bj (corresponding to the action of P^{ij}_{ab} in the equation for the excited state PNOs) also leads to different contributions. For example, if ij refers to a rank-1 pair where i contributes to or is part of the occupied NTO, the term $R^{ij}_{ab}(ck|bb)$ contains a dipole–monopole interaction between the orbital products ck and bb (see Fig. 1). Hence, this dipole–monopole interaction directly refers to the distance between k and j due to the exponential distance dependence between the orbitals b and j. On the other hand, j is also only connected to the occupied NTO via a dipole–monopole interaction (as stated previously when discussing the excited state pairs). This would imply that the pair kj is neither part of the ground-state nor part of the excited-state LMO pair list. The ground- and excited-state pair lists are the truncated estimates that use integrals, which contain dipole–dipole (ground state pairs) and dipole–monopole interactions between two LMOs or a LMO and the occupied or virtual NTO (excited state pairs), respectively, as most long ranging contributions. However, if the distance between ck and bb as well as jj is large compared to the distance between the orbitals b and j, this long-ranging contribution cancels out against the dipole–monopole contribution arising from $R^{ij}_{ab}(ck|jj)$. This occurs on the same length scale on which in the nominators for the ground state triples amplitudes, the $\mathcal{O}(\mathcal{R}^{-2})$ contributions become negligible.

The analysis made so far suggests that similar to the ground state LMO triples, the excited LMO triples ijk also have to be connected to each other by three pair relations, i.e., all three pairs ij, ik, and jk have to be significant excited state or ground-state pairs. If the contribution from $R^{ji}_{ba}(ck|bb)$ is considered (interchanged indices ai and bj corresponding to the action of P^{ij}_{ab} in the equation for the excited state PNOs), the situation changes significantly (as depicted in Fig. 2). In this case, the orbital b is close to or contributes to the virtual NTO and is coupled to the orbital product ck via a $\mathcal{O}(\mathcal{R}^{-2})$ decaying dipole–monopole interaction. This implies that jk is a rank-v excited-state pair, but there is no pair relation between the LMOs i and k since i is coupled to the occupied NTO via a dipole–monopole interaction, while k is coupled to the virtual NTO via a dipole–monopole interaction. Hence, k can be rather far from the virtual NTO, and i can be rather far from the occupied NTO. The important finding here is that there is no term in the equations for the excited state triples amplitudes, which would cancel this long-ranging contribution. This implies that for charge-transfer states, the excited state LMO triples ijk only need to contain two pair relations. There can be two LMOs that are not connected to each other by an excited- or ground-state pair relation. A similar analysis holds for the cases where ij refers to a rank-v pair.

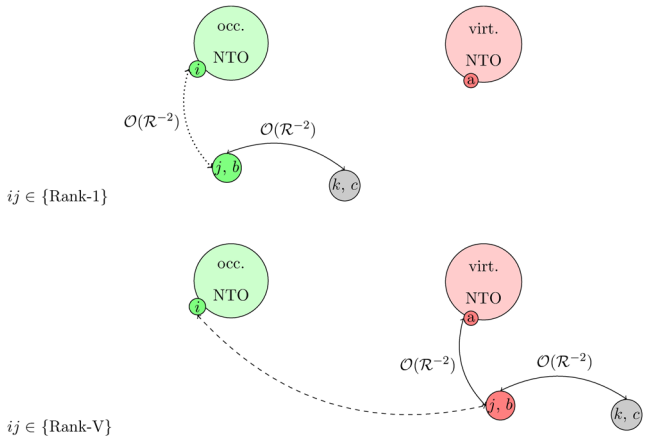


FIG. 1. Pictorial representation of the long-ranging contributions $R^{ij}_{ab}(ck|bb)$ and $R^{ij}_{ab}(ck|jj)$ to the excited state triples amplitudes, which only require one pair-relation but do cancel out each other. The dashed line indicates a rank-v excited state pair relation, while the dotted line refers to a rank-1 excited state pair relation. Green circles denote orbitals that are coupled to the occupied NTO (occ. NTO), and red circles refer to orbitals that are coupled to the virtual NTO (virt. NTO). Gray circles denote orbitals that are not directly coupled to the chromophore.

The selection of the triples for the excited-state amplitudes has to account for these differences relative to the ground state. In contrast to the ground-state LMO triples where all three LMO indices ijk are coupled to each other by pair relations, the excited-state LMO triples only have to fulfill two pair relations to give rise to significant contributions. Only for the special case of local excitations, where the occupied and virtual NTOs are localized on a common small set of atoms, the three LMOs of the non-negligible excited-state triples will also be connected by three pair relations. Hence, the long-ranging contributions of the term $\sum_d R^{ij}_{ad}(ck|bd)$, which in the charge-transfer case do only involve two pair relations cancel out against the long-ranging

Published under license by AIP Publishing

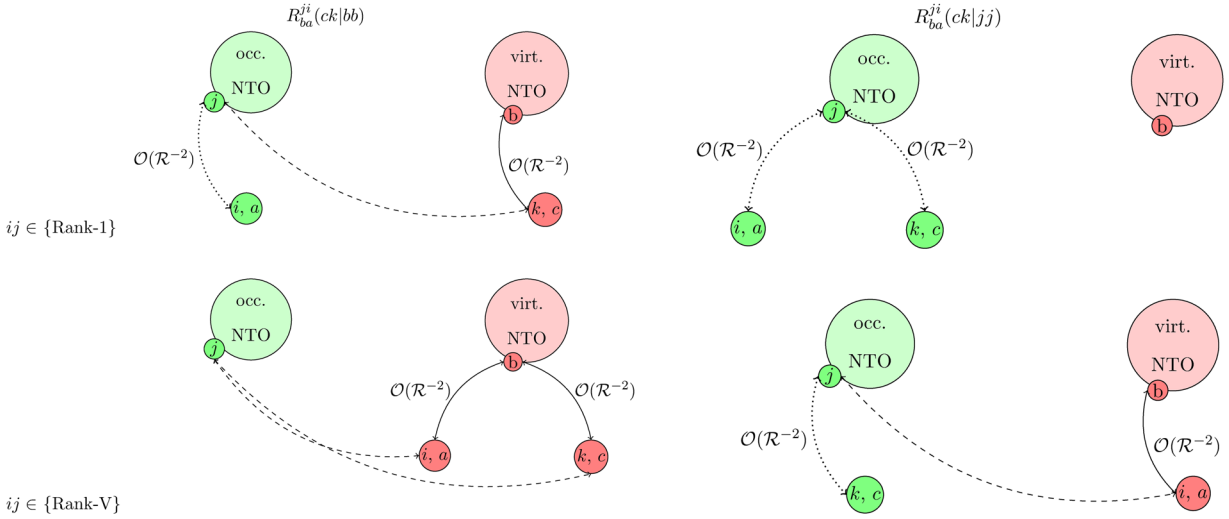


FIG. 2. Pictorial representation of the long-ranging contributions $R^{ji}_{ba}(ck|bb)$ and $R^{ji}_{ba}(ck|jj)$ to the excited state triples amplitudes, which only require two pair-relations and do not cancel out each other. The dashed lines indicate a rank-v excited state pair relation, while the dotted lines refer to a rank-1 excited state pair relation. Green circles denote orbitals that are coupled to the occupied NTO (occ. NTO), and red circles refer to orbitals that are coupled to the virtual NTO (virt. NTO).

contributions from $\sum_l R^{il}_{ab}(ck|lj)$. An example that demonstrates the LMO triples selection for charge-transfer and local excitations is given in Sec. IV D.

From Eq. (43), it follows that the excited state TNO basis has to span the virtual spaces for the ground state doubles amplitudes t_{μ_2}, the excited state doubles amplitudes R_{μ_2}, and the undressed and R_1-dressed ERIs with sufficient accuracy. To achieve this, we construct the approximate triples density,

$$\begin{aligned} D^{ijk}_{\tilde{a}\tilde{b}} = \frac{1}{6}\Big(& S^{ijk,\,ij}_{\tilde{a}\tilde{a}^0} D^{ij,\mathrm{gs}}_{\tilde{a}^0\tilde{b}^0} S^{ij,ijk}_{\tilde{b}^0\tilde{b}} + S^{ijk,\,ik}_{\tilde{a}\tilde{a}^0} D^{ik,\mathrm{gs}}_{\tilde{a}^0\tilde{b}^0} S^{ik,ijk}_{\tilde{b}^0\tilde{b}} \\ & + S^{ijk,\,jk}_{\tilde{a}\tilde{a}^0} D^{jk,\mathrm{gs}}_{\tilde{a}^0\tilde{b}^0} S^{jk,ijk}_{\tilde{b}^0\tilde{b}} + S^{ijk,\,ij}_{\tilde{a}\tilde{a}} D^{ij,\mathrm{es}}_{\tilde{a}\tilde{b}} S^{ij,ijk}_{\tilde{b}\tilde{b}} \\ & + S^{ijk,\,ik}_{\tilde{a}\tilde{a}} D^{ik,\mathrm{es}}_{\tilde{a}\tilde{b}} S^{ik,ijk}_{\tilde{b}\tilde{b}} + S^{ijk,\,jk}_{\tilde{a}\tilde{a}} D^{jk,\mathrm{es}}_{\tilde{a}\tilde{b}} S^{jk,ijk}_{\tilde{b}\tilde{b}} \Big), \end{aligned} \tag{44}$$

where $\mathbf{D}^{\mathrm{es}}$ denotes approximate CIS(D) densities [cf. Eqs. (18) and (19)], evaluated in the excited state PNO basis $\{\bar{a}\}$, and $\mathbf{D}^{\mathrm{gs}}$ are approximate MP2 densities [cf. Eqs. (16) and (17)], evaluated in the ground state PNO basis $\{\bar{a}^0\}$. The selection of the excited state TNOs for a triple *ijk* starts from a pre-TNO basis, obtained by merging the excited state OSVs of the LMOs *i*, *j*, and *k*. Since the excited state OSVs are constructed such that they form a suitable basis for integrals $(ai|bj)$ and $(ai|\bar{b}j)$, the pre-TNO basis obtained this way is also adequate for the ground- and excited-state amplitudes. The excited state TNOs are then obtained by diagonalizing the triples density $\mathbf{D}^{ijk}$. Afterwards, the Fock matrix is diagonalized for each triple in its selected excited state TNO basis.

E. Evaluation of the excited state triples amplitudes: T0 approximation and Laplace transformation

The evaluation of the triples amplitudes is more difficult in a localized MO basis than in the canonical basis because the Fock matrix is no longer diagonal and the amplitudes have now to be determined as (approximate) solutions of a set of coupled linear equations. This problem appears already in local variants of CCSD(T). Several strategies have been proposed to tackle this. In the so-called T0 approximation, the off-diagonal elements in the occupied/occupied block of the Fock matrix are neglected. It is a simple and computationally inexpensive approximation that performs well in many cases, but in other cases, it causes relatively large errors.[18,46,85] This makes it desirable to have a possibility to include the triples correction in a more accurate way to verify the results obtained with the T0 approximation. This led Schütz[18] to the development of the T1 approximation, which includes the couplings through F_{ij} to the first order and corresponds to a modified first iteration of an iterative algorithm, which can also be used to solve for the local CCSD(T) triples amplitudes. Even though the T1 approximation, in contrast to the iterative solution, still avoids the I/O of triples amplitudes, it is computationally far more expensive than the T0 approximation. Furthermore, Guo *et al.* demonstrated an improved linear scaling triples correction in their domain-based local PNO (DLPNO) framework, where a small set of triples is selected for which the amplitudes are iteratively refined.[86]

In the context of a PNO-CCSD(F12*)(T) implementation,[46] we proposed recently to use a numerical Laplace transformation to approximate the energy denominator. This is systematically improvable (through the number of quadrature grid points) and also avoids the I/O of triples amplitudes, which even in a local implementation could be a severe bottleneck. Since the performance of these approaches have not yet been studied for excitation energies, we implemented, for the present work, both the T0 approximation [CC3(T0)] and the Laplace transformation (CC3) technique to investigate their performance for excitation energies.

Within the T0 approximation, the excited state triples amplitudes are in the LMO and TNO basis approximated as

Published under license by AIP Publishing

$$R^{ijk}_{\tilde{a}\tilde{b}\tilde{c}} = \frac{1}{\Delta^{ijk}_{\tilde{a}\tilde{b}\tilde{c}} + \omega} P^{ijk}_{\tilde{a}\tilde{b}\tilde{c}} \Bigg\{ \sum_{\tilde{d}} (\bar{c}k|\hat{\bar{b}}\bar{d}) \sum_{\tilde{a}'\tilde{d}'} S^{ijk,\,ij}_{\tilde{a}\tilde{a}'} R^{ij}_{\tilde{a}'\tilde{d}'} S^{ij,ijk}_{\tilde{d}'\tilde{d}} - \sum_{l} (\bar{c}k|\hat{lj}) \sum_{\tilde{a}\tilde{b}} S^{ijk,\,il}_{\tilde{a}\tilde{a}} R^{il}_{\tilde{a}\tilde{b}} S^{il,ijk}_{\tilde{b}\tilde{b}} + \sum_{\tilde{d}} (\bar{c}k|\bar{\tilde{b}}\bar{d}) \sum_{\tilde{a}^0\tilde{d}^0} S^{ijk,\,ij}_{\tilde{a}\tilde{a}^0} t^{ij}_{\tilde{a}^0\tilde{d}^0} S^{ij,ijk}_{\tilde{d}^0\tilde{d}} - \sum_{l} (\bar{c}k|\bar{lj}) \sum_{\tilde{a}^0_{il}\tilde{b}^0} S^{ijk,\,il}_{\tilde{a}\tilde{a}^0} t^{il}_{\tilde{a}^0\tilde{b}^0} S^{il,ijk}_{\tilde{b}^0\tilde{b}} \Bigg\} \tag{45}$$

with $\Delta^{ijk}_{\tilde{a}\tilde{b}\tilde{c}} = F_{ii} + F_{jj} + F_{kk} - \epsilon_{\tilde{a}} - \epsilon_{\tilde{b}} - \epsilon_{\tilde{c}}$, where $\epsilon_{\tilde{a}}$ and F_{ii} denote the Fock matrix eigenvalues in the excited state TNO basis and the diagonal Fock matrix elements in the LMO basis, respectively. The PNO/TNO overlap matrices that are used to transform the virtual indices of the doubles amplitudes from the PNO to the TNO basis are defined as

$$S^{ijk,\,ij}_{\tilde{a}\tilde{b}} = \sum_{a} d^{ijk}_{a\tilde{a}} d^{ij}_{a\tilde{b}}. \tag{46}$$

The latter equations for the excited state triples amplitudes have been implemented analogously to the (T0) correction for ground state energies reported in Ref. 46.

The Laplace transform of the energy denominator[87] can be expressed as

$$\frac{1}{\epsilon_i + \epsilon_j + \epsilon_k - \epsilon_a - \epsilon_b - \epsilon_c} = \int_0^\infty e^{-(\epsilon_i+\epsilon_j+\epsilon_k-\epsilon_a-\epsilon_b-\epsilon_c)x} dx \tag{47}$$

$$\approx \sum_{z}^{n_z} w_z e^{-(\epsilon_i+\epsilon_j+\epsilon_k-\epsilon_a-\epsilon_b-\epsilon_c)t_z}, \tag{48}$$

where w_z denotes the quadrature weights, t_z denotes the position, and n_z denotes the number of the grid points.[88,89] The advantage of the Laplace transformation is that the off-diagonal Fock matrix elements can also be included without the need of an iterative solution scheme because the right-hand side of Eq. (48) is factorizable. After applying the Laplace transformation, the excited state triples amplitudes can be evaluated as follows in the excited-state TNO basis:

$$R^{ijk}_{\tilde{a}\tilde{b}\tilde{c}} = P^{ijk}_{\tilde{a}\tilde{b}\tilde{c}} \sum_{z}^{n_L} w_z \Bigg\{ \sum_{\tilde{d}_{ijk}} (\bar{c}^z k^z|\bar{\tilde{b}}^z \bar{d}) \sum_{\tilde{a}'\tilde{d}'} S^{ijk,\,ij,\,z}_{\tilde{a}\tilde{a}'} R^{i^z j^z}_{\tilde{a}'\tilde{d}'} S^{ij,ijk}_{\tilde{d}'\tilde{d}} - \sum_{l} (\bar{c}^z k^z|\hat{lj}^z) \sum_{\tilde{a}'\tilde{b}'} S^{ijk,\,il,\,z}_{\tilde{a}\tilde{a}'} R^{i^z l}_{\tilde{a}'\tilde{b}'} S^{il,ijk,z}_{\tilde{b}'\tilde{b}} + \sum_{\tilde{d}} (\bar{c}^z k^z|\bar{\tilde{b}}^z \bar{d}) \sum_{\tilde{a}^0\tilde{d}^0} S^{ijk,\,ij,\,z}_{\tilde{a}\tilde{a}^0} t^{i^z j^z}_{\tilde{a}^0\tilde{d}^0} S^{ij,ijk}_{\tilde{d}^0\tilde{d}} - \sum_{l} (\bar{c}^z k^z|\bar{lj}^z) \sum_{\tilde{a}^0\tilde{b}^0} S^{ijk,\,il,\,z}_{\tilde{a}\tilde{a}^0} t^{i^z l}_{\tilde{a}^0\tilde{b}^0} S^{il,ijk,z}_{\tilde{b}^0\tilde{b}} \Bigg\}. \tag{49}$$

The grid point dependent LMOs $|i^z\rangle$ and TNOs $|\tilde{a}^z\rangle$ are generated using grid point dependent transformation matrices

$$Y^z_{\mu\tilde{a}_{ijk}} = \sum_{a} Y_{\mu a} e^{-\left(\epsilon_a + \frac{\omega}{6}\right)t_z} d^{ijk}_{a\tilde{a}}, \tag{50}$$

$$X^z_{\mu i} = \sum_{j} X_{\mu j} e^{\left(\epsilon_j - \frac{\omega}{6}\right)t_z} T_{ij}, \tag{51}$$

where the matrices $\mathbf{Y}$ and $\mathbf{X}$ can contain the (unmodified) MO coefficients $C_{\mu p}$ or the coefficients for t_{μ_1} or R_{μ_1}-dressed particle and hole transformation matrices.[80] The matrix T_{ij} transforms from the canonical MO to the LMO basis.

The grid point dependent PNO/TNO overlap matrices are defined as

$$S^{ijk,\,ij,\,z}_{\tilde{a}\tilde{b}} = \sum_{a} d^{ijk}_{a\tilde{a}} e^{-\left(\epsilon_a + \frac{\omega}{6}\right)t_z} d^{ij}_{a\tilde{b}} \tag{52}$$

and transform from the unmodified PNO to the grid point dependent TNO basis. Additionally, the ground- and excited-state doubles amplitudes with one and two grid point dependent LMO indices are needed. Amplitudes with one grid point dependent LMO index are, similar as in the PNO-CCSD(T) implementation,[46] calculated as

$$R^{ij^z}_{\tilde{a}\tilde{b}} = R^{ij}_{\tilde{a}\tilde{b}} T^z_{jj} + \sum_{k\neq j} T^z_{kj} \sum_{\tilde{a}'\tilde{b}'} S^{ij,ik}_{\tilde{a}\tilde{a}'} R^{ik}_{\tilde{a}'\tilde{b}} S^{ik,ij}_{\tilde{a}'\tilde{a}}, \tag{53}$$

and amplitudes with two grid point dependent LMO indices are obtained by applying the above equation to the remaining untransformed LMO index of $R^{ij^z}_{\tilde{a}\tilde{b}}$. The transformation matrix T^z_{ij} that transforms in Eq. (53) from the unmodified LMO basis to the grid point dependent LMO basis is given by

$$T^z_{ij} = \sum_{k} T_{ik} e^{(\epsilon_k - \omega)t_z} T_{jk}. \tag{54}$$

F. Solution of the PNO-CC3 ground and excited state equations

Since the triples contributions to the residual of the ground-state cluster equations and the contraction of the Jacobian with a trial vector from the right are similar, we discuss them together with a focus on the latter case. We only describe the contributions beyond CCSD and refer the reader for the implementation of the CCSD vector function and Jacobian transformations in a PNO basis to Refs. 45 and 76.

Table II summarizes the CC3 contributions to the singles and doubles parts of the transformed vector. It can be seen that the time-determining $\mathcal{O}(\mathcal{N}^7)$-scaling steps of the canonical implementation become, in the local implementation, linear scaling in the limit that the number of significant LMO pairs and triples grows only linear with the system size and the number of PNOs and TNOs per pair and triple becomes size-independent, respectively. For the terms $\rho^{ij,G_{DT}}_{\tilde{a}\tilde{b}}$ and $\rho^{ij,I}_{\tilde{a}\tilde{b}}$, this is a consequence from the use of TNOs and the restriction of the summations over occupied orbitals that follows from the sparse triples list $\{ijk\}$. For $\rho^{ij,H}_{\tilde{a}\tilde{b}}$, the contribution from the ground-state wave function can be restricted to triples jkl, where at least one LMO forms an excited state pair with i, and the excited-state contribution can be restricted to triples jkl where at least one LMO forms an excited-state or a ground-state pair with i. Similar considerations allow us to restrict the summation over LMOs l in the equations for the excited state triples amplitudes.

In the implementation of the triples contribution for PNO-CC3, we follow a similar strategy as used for PNO-CCSD(T) and PNO-CCSD where only compact intermediates and integrals are stored on disk.[46,76] However, for PNO-CC3, a larger number of

Published under license by AIP Publishing

TABLE II. Explicit expressions for the PNO-CC3 contributions to the singles and doubles part of the Jacobian transformed vector ρ_{μ_1} and ρ_{μ_2}. For ease of notation, the triples and pair indices of PNOs and TNOs are not given explicitly.

Contribution	Formal scaling CC3	Formal scaling PNO-CC3
$\rho^{I}_{ai} = 2\sum_{jk}\sum_{\bar{a}} d^{ijk}_{a\bar{a}} \sum_{\bar{b}\bar{c}} \left(R^{ijk}_{\bar{a}\bar{b}\bar{c}} - R^{ijk}_{\bar{c}\bar{b}\bar{a}}\right)\left(2(j\bar{b}\vert k\bar{c}) - (j\bar{c}\vert k\bar{b})\right)$	$\mathcal{O}(\mathcal{N}^6)$	$\mathcal{O}(\mathcal{N}^2)$
$\rho^{ij,G}_{\bar{a}'\bar{b}'} = 2\sum_{\bar{a}\bar{b}}\sum_{k} S^{ij,ijk}_{\bar{a}'\bar{a}} S^{ij,ijk}_{\bar{b}'\bar{b}} P^{ij}_{\bar{a}\bar{b}} \sum_{\bar{c}\bar{d}} \left(2R^{ijk}_{\bar{c}\bar{b}\bar{d}} - R^{ijk}_{\bar{c}\bar{d}\bar{b}} - R^{ijk}_{\bar{d}\bar{c}\bar{b}}\right)(\bar{a}\hat{\bar{c}}\vert k\bar{d})$ $+ 2\sum_{\bar{a}^0\bar{b}^0}\sum_{k} S^{ij,ijk}_{\bar{a}'\bar{a}^0} S^{ij,ijk}_{\bar{b}'\bar{b}^0} P^{ij}_{\bar{a}^0\bar{b}^0} \sum_{\bar{c}^0\bar{d}^0} \left(2t^{ijk}_{\bar{c}^0\bar{b}^0\bar{d}^0} - t^{ijk}_{\bar{c}^0\bar{d}^0\bar{b}^0} - t^{ijk}_{\bar{d}^0\bar{c}^0\bar{b}^0}\right)(\bar{a}^0\hat{\bar{c}^0}\vert k\bar{d}^0)$	$\mathcal{O}(\mathcal{N}^7)$	$\mathcal{O}(\mathcal{N})$
$\rho^{ij,H}_{\bar{a}'\bar{b}'} = -2\sum_{\bar{a}\bar{b}} P^{ij}_{\bar{a}\bar{b}} \sum_{kl} S^{ij,jkl}_{\bar{a}'\bar{a}} S^{ij,jkl}_{\bar{b}'\bar{b}} \sum_{\bar{c}} \left(2R^{jkl}_{\bar{b}\bar{a}\bar{c}} - R^{jkl}_{\bar{c}\bar{a}\bar{b}} - R^{jkl}_{\bar{b}\bar{c}\bar{a}}\right)(ki\hat{\vert}l\bar{c})$ $- 2\sum_{\bar{a}^0\bar{b}^0} P^{ij}_{\bar{a}^0\bar{b}^0} \sum_{kl} S^{ij,jkl}_{\bar{a}'\bar{a}^0} S^{ij,jkl}_{\bar{b}'\bar{b}^0} \sum_{\bar{c}} \left(2t^{jkl}_{\bar{b}^0\bar{a}^0\bar{c}^0} - t^{jkl}_{\bar{c}^0\bar{a}^0\bar{b}^0} - t^{jkl}_{\bar{b}^0\bar{c}^0\bar{a}^0}\right)(ki\vert l\bar{c}^0)$	$\mathcal{O}(\mathcal{N}^7)$	$\mathcal{O}(\mathcal{N})$
$\rho^{ij,I}_{\bar{a}'\bar{b}'} = 2\sum_{\bar{a}\bar{b}}\sum_{k} S^{ij,ijk}_{\bar{a}'\bar{a}} S^{ij,ijk}_{\bar{b}'\bar{b}} P^{ij}_{\bar{a}\bar{b}} \sum_{\bar{c}} \left(R^{ijk}_{\bar{a}\bar{b}\bar{c}} - R^{ijk}_{\bar{c}\bar{b}\bar{a}}\right)\hat{F}_{k\bar{c}}$ $2\sum_{\bar{a}\bar{b}}\sum_{k} S^{ij,ijk}_{\bar{a}'\bar{a}^0} S^{ij,ijk}_{\bar{b}'\bar{b}^0} P^{ij}_{\bar{a}^0\bar{b}^0} \sum_{\bar{c}^0} \left(t^{ijk}_{\bar{a}^0\bar{b}^0\bar{c}^0} - t^{ijk}_{\bar{c}^0\bar{b}^0\bar{a}^0}\right)\hat{F}_{k\bar{c}^0}$	$\mathcal{O}(\mathcal{N}^6)$	$\mathcal{O}(\mathcal{N})$

different two-electron integrals is needed than for PNO-CCSD(T), and several of them are dressed with singles amplitudes that change during the iterations. For instance, the integrals $(ck\hat{\vert}bd)$ and $(ck\hat{\vert}lj)$ are cubic in t_{μ_1}, and $(ck\bar{\vert}bd)$ and $(ck\bar{\vert}lj)$ contain three R_{μ_1}-dressed indices. Therefore, we compute all integrals needed for the CC3 triples contributions in each iteration in an AO-integral-direct way.

For the solution of the eigenvalue problem with state-specific PNOs, we use, at the CCSD level, a single-state Davidson algorithm with Wielandt deflation, as described in detail in Refs. 74 and 76. In the partitioned formulation, which avoids the storage of triples amplitudes, CC3 requires the solution of a non-linear eigenvalue problem of the form $\mathbf{A}^{\mathrm{eff}}(\omega)\mathbf{E}(\omega) = \omega\mathbf{E}(\omega)$ where the effective Jacobi matrix and the trial vectors defined in Eqs. (13) and (12) depend on the excitation energy. For such a non-linear eigenvalue problem, the standard Wielandt deflation is not applicable. The solution of a nonlinear eigenvalue problem could be avoided by solving the eigenvalue problem for the full Jacobian in the singles, doubles, and triples space. This would, however, require that the triples basis vectors and their linear transformations with $\mathbf{A}$ are stored on disk for several iterations and states. For the singles and doubles amplitudes, this in a PNO-based implementation is not a problem. However, even in a local implementation, the storage requirements and the I/O for the triples amplitudes can become demanding, if not prohibitive. Hence, in our implementation, we follow a different route for solving the excited state eigenvalue problem, which is similar to strategies used in canonical CC3 implementations: For this, we first solve the PNO-CCSD ground-state equations and eigenvalue problem to obtain accurate start guesses for the CC3 amplitudes and eigenvectors. After obtaining the solution of the PNO-CC3 ground-state amplitude equations, we construct for each excited-state a new PNO basis from the singles parts of the converged PNO-CCSD eigenvectors and excitation energies. The doubles parts of the PNO-CCSD eigenvectors are projected onto the new PNO basis, and the non-linear PNO-CC3 eigenvalue problem is solved in this basis using a direct inversion of the iterative subspace (DIIS) algorithm as described originally for a canonical RI-CC2 implementation.[90]

III. COMPUTATIONAL DETAILS

Our implementation is part of the pnoccsd program[45,48] in a local development version of the TURBOMOLE quantum chemistry program package.[91] All calculations employ the frozen core approximation with the 1*s* orbitals of second-row elements excluded from the post-Hartree–Fock calculations. The Hartree–Fock calculations have been done with the dscf[92] and the canonical CC3 reference calculations with the ccsdf12 program of TURBOMOLE. For the benchmark calculations on the Schreiber test set[93] and for the calculations on alkene chains, we used the TZVP orbital basis set.[94] For the calculations on berenil, we used the aug-cc-pVDZ and the aug-cc-pVTZ basis sets.[95,96] If not otherwise noted, the corresponding optimized auxiliary basis sets have been used[97,98] for the resolution-of-the-identity (RI) approximation for two-electron integrals in the post-HF calculations. The geometries of Schreiber's test set have been taken from Ref. 93.

For the PNO-CC3 and PNO-CC3(T0) calculations, we used intrinsic bond orbitals[99] (IBOs) as the localization scheme for the occupied orbital space and carried out calculations with PNO thresholds of $\log(T_{\mathrm{PNO}}) = -6, -7, -8$. The OSV, TNO, and pair thresholds are directly coupled to the PNO threshold as $T_{\mathrm{OSV}} = 0.1 \cdot T_{\mathrm{PNO}}$, $T_{\mathrm{TNO}} = T_{\mathrm{PNO}}$, and $T_{\mathrm{pair}} = (0.1 \cdot T_{\mathrm{PNO}})^{\frac{2}{3}}$, respectively. The list of LMO triples for the triple excitation part of the eigenvectors are, if not otherwise noted, selected based on three pair relations, as described in Sec. II D.

Published under license by AIP Publishing

IV. RESULTS AND DISCUSSION

A. Choice of the Laplace grid

To investigate the dependence of the accuracy on the number of Laplace grid points, the first three excited states of a subset of molecules from Schreiber's test set (ethene, butadiene, cyclopropene, formaldehyde, acetone, and formamide) have been computed with different numbers of Laplace grid points (n_L = 1, ..., 5) and, for comparison, with the T0 approximation and canonical CC3. For these calculations, the TZVP orbital basis was used in combination with the def2-QZVPP auxiliary basis[100] to minimize the errors arising from the different use of the RI approximation in canonical and PNO-based CC3. The deviations have been computed relative to the canonical CC3 result.

Figure 3 shows the results of the statistical analysis of these states and molecules based on the mean absolute errors (MAEs) and maximum errors (MAX). It can be seen that for PNO-CC3, the errors are already converged with n_L = 3 for all investigated PNO thresholds. This corresponds to a threshold for the convergence of the Laplace transformation, which is in the range of 10^{-2}–$3.2 \cdot 10^{-2}$. For the loose threshold $T_{\mathrm{PNO}} = 10^{-6}$, the benefit from the Laplace transformation is small, irrespective of the number of grid points. The maximum errors for PNO-CC3 and PNO-CC3(T0) are basically identical since the errors are dominated for this PNO threshold by the PNO and TNO approximations. For the remaining PNO thresholds, we find with $n_L = 3$ a decrease in the errors by a factor of up to 4. For $n_L < 3$, the errors are larger than for the T0 approximation since in this case the largest error arises from the non-converged Laplace transformation.

The fact that the MAE for PNO-CC3 is for the loose threshold $T_{\mathrm{PNO}} = 10^{-6}$ lowest with $n_L = 1$ may be explained by a fortunate error compensation. Figure 3 suggests that for $n_L < 3$, the errors for $T_{\mathrm{PNO}} = 10^{-7}$ are lower than for the tighter PNO thresholds $T_{\mathrm{PNO}} = 10^{-8}$ and $T_{\mathrm{PNO}} = 10^{-9}$. This is due to the fact that with very few Laplace grid points, the deviations from the canonical result are dominated by the error from the numerical Laplace integrals and not by the error from the truncation of the virtual spaces. With a larger number of grid points for the Laplace transformation, we observe the expected systematic convergence of the results with the PNO threshold until at $T_{\mathrm{PNO}} = 10^{-9}$ the deviations are dominated by differences in the use of the RI approximation in the PNO and the canonical calculations.

B. Choice of the TNO threshold

To investigate the dependence of the errors in the excitation energies on the TNO threshold, we performed PNO-CC3 (with n_L = 4) calculations with different combinations of PNO and TNO thresholds (the same thresholds were used for the ground and the excited states) and the same set of molecules and states as in Sec. IV A. Again, the calculations were done in the TZVP orbital and def2-QZVPP auxiliary basis.

The results are summarized in Fig. 4. The dependence of the errors on the TNO threshold becomes rather weak once $T_{\mathrm{TNO}} \lessapprox T_{\mathrm{PNO}}$. The reason for this is that the TNO space is already restricted by the PNO truncation but not vice versa. Figure 4 suggests that the PNO-CC3 results do not significantly improve if the TNO threshold is chosen tighter than T_{PNO}. This was also observed by Schmitz and Hättig[46] for PNO-CCSD(T) ground-state energies. Hence, we suggest to set T_{TNO} for the ground and excited states to T_{PNO}.

C. Accuracy of PNO-CC3 excitation energies

For benchmarking the accuracy of PNO-CC3 and PNO-CC3(T0), we chose Schreiber's test set[93] that consists of about 150 excited states in 28 molecules. The states of the nucleobases adenine, thymine, uracil, and cytosine were taken from Ref. 101 since Ref. 93 did not report CC3 values for these molecules. Furthermore, we

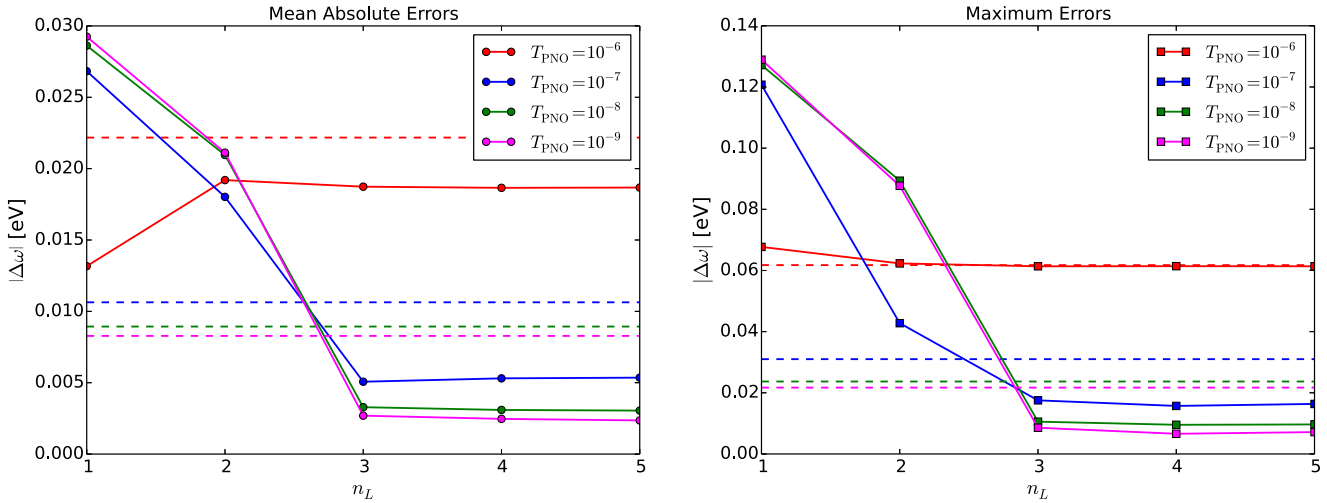


FIG. 3. Mean absolute and maximum errors for PNO-CC3 as a function of the number of Laplace grid points. The dotted lines indicate the PNO-CC3(T0) errors for the respective PNO-threshold.

Published under license by AIP Publishing

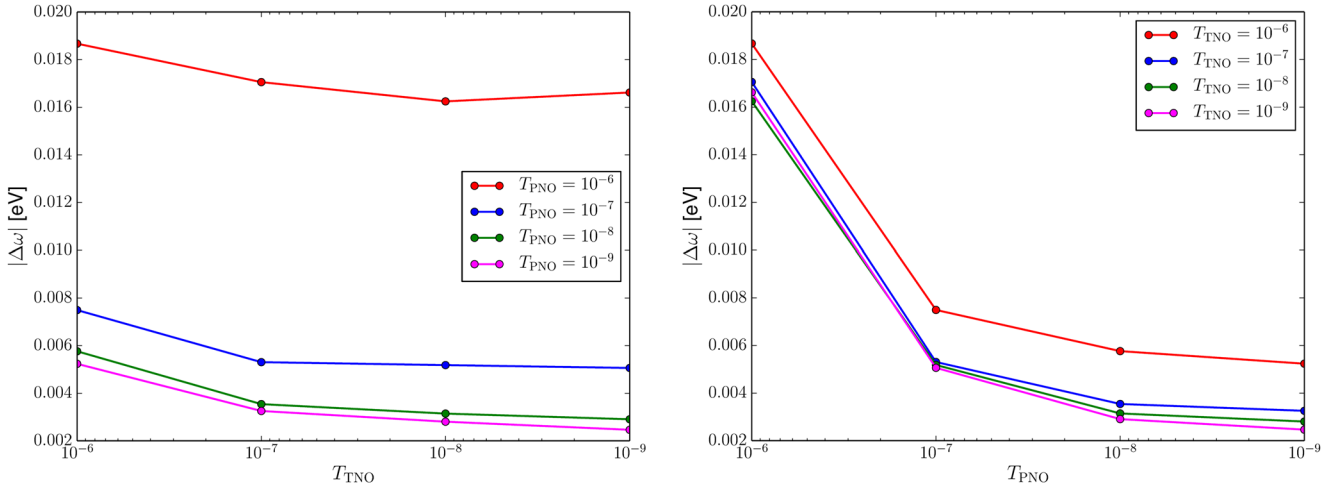


FIG. 4. Mean absolute errors of a subset of Schreiber's test set as a function of the TNO threshold (left) and as a function of the PNO threshold (right).

excluded the highest excited state of thymine and acetamide, since a definite assignment of these two states was no possible.

Figure 5 and Table III show a statistical analysis of the results for all states in the test set. For comparison, we included the results for PNO-CCSD for this test set obtained in Ref. 76. For PNO-CC3, the MAE and maximum errors over all states show a rapid convergence toward the canonical result. The curves are approximately parallel to those for PNO-CCSD, only shifted to slightly larger errors. The maximum errors are for PNO-CC3 approximately twice as large as for PNO-CCSD, while for the MAE, the errors are approximately by a factor of 1.6 larger than for PNO-CCSD. For PNO-CC3(T0), the deviations from the canonical CC3 results include in addition to the PNO and TNO truncation errors also the errors from the T0 approximation. The latter limit the convergence to the conical limit, and thus, PNO-CC3(T0) has significantly larger deviations, which below $T_{\mathrm{PNO}} = 10^{-7}$ approach constant values. This is especially pronounced for states with a significant doubles contribution ($\%T_1 \leq 80\%$) where the maximum error for PNO-CC3(T0) does not drop below 0.15 eV. Apparently, for doubly excited states, the errors arising from the T0 approximation have about the same magnitude as the PNO and TNO truncation errors at $T_{\mathrm{PNO}} = 10^{-7}$. In contrast to this, the deviations of PNO-CC3(T0) decrease for singly excited states ($\%T_1 > 80\%$) with $T_{\mathrm{PNO}} = 10^{-8}$ to a MAE of 0.01 eV and a maximum error, which is approximately three times larger than for its Laplace-transformed counterpart. Hence, it performs for singly excited states significantly better than for doubly excited states. For the loose threshold $T_{\mathrm{PNO}} = 10^{-6}$, the errors of PNO-CC3 and PNO-CC3(T0) are very similar, and there is no benefit from going beyond the T0 approximation. PNO-CC3(T0) with $T_{\mathrm{PNO}} = 10^{-6}$ can, thus, be used as an economic approach to check for the importance of the triples contribution. If a higher accuracy is needed, the Laplace transformation provides a systematic convergence to the canonical CC3 with decreasing threshold T_{PNO}.

Figure 6 visualizes the distribution of the errors. This shows that the PNO results tend to convergence from above to the canonical values. The PNO and TNO truncation leads on average to a blue shift in the excitation energies.

D. Charge-transfer excitations

Charge-Transfer (CT) excitations pose a challenge to local correlation methods due to their non-local character. Already in previous publications[76,78] on PNO-CC2 and PNO-CCSD excitation energies, it has been found that charge-transfer excitations need to be treated with special care to preserve the accuracy of the PNO approximation for non-local excitations.

We used the model system C_2H_4–C_2F_4 in the cc-pVDZ basis set to study the performance of PNO-CC3 for CT excitations. Figure 7 shows the errors for the lowest local and CT $\pi\pi^*$ excitations. Two types of results are shown. In one set of calculations, we included, for the excited states, only LMO triples *ijk* where the three LMOs are, similar as for ground state triples, connected by three pair relations, and a second set of calculations where for the excited state also the LMO triples *ijk* with only two pair relations have been included.

Figure 7 clearly shows the importance of the relaxed constraints for the triples selection for CT states. While for excitations localized on either C_2H_4 or C_2F_4 both triples selections yield almost identical results, there are significant differences for both of the CT excitations. If only those triples are included for which all three LMOs *ijk* are related by a pair relation (black lines in Fig. 7), the errors for PNO-CC3 for the loose and moderate PNO truncation thresholds 10^{-6} and 10^{-7} are at least 0.15 eV—almost one order of magnitude larger than the errors observed for local excitations in this model system and for the test set in Sec. IV C. If for the excited states also the triples with only two pair relations are included, the errors for the CT states decrease to values of similar magnitude as for the local excitations and a similar convergence pattern. The fact that both selection schemes for the excited states converge to the same results at $T_{\mathrm{PNO}} = 10^{-8}$ is explained by the fact that for this model system and

Published under license by AIP Publishing

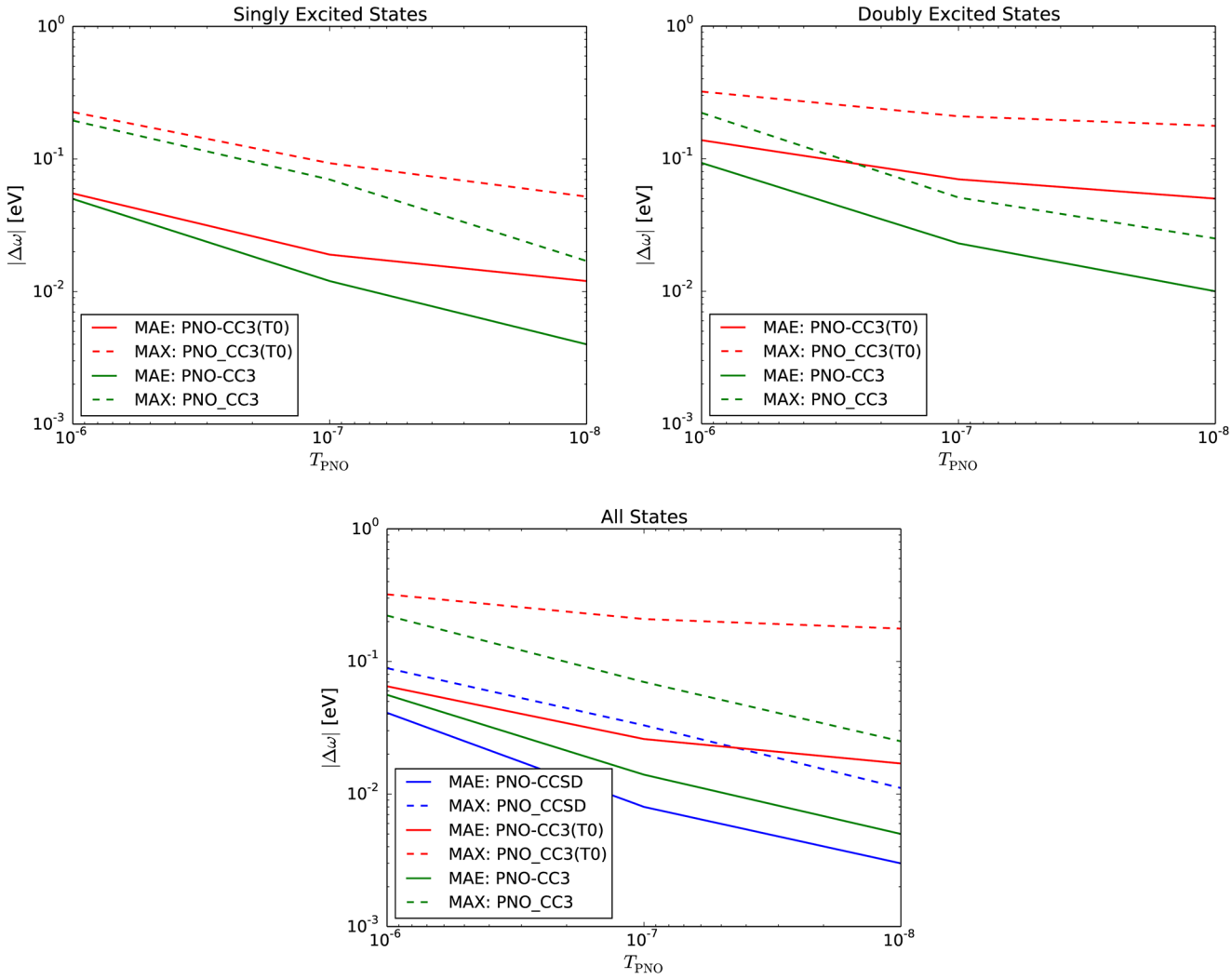


FIG. 5. Mean absolute and maximum errors for PNO-CC3 and PNO-CC3(T0) with respect to the canonical method obtained by a statistical analysis over the excited states included in Schreiber's test set. Additionally, singly excited states ($\%T_1 > 80\%$) and states with a significant doubles contribution ($\%T_1 \leq 80\%$) are analyzed separately in the upper figures.

truncation threshold, no LMO pairs (and, hence, no excited state LMO triples) are discarded.

E. Computational efficiency of the implementation

1. Scaling of the excited state TNO spaces

To investigate how the number of excited-state TNOs scales with the chromophore size and how this compares with the number of ground-state TNOs and ground- and excited-state PNOs and OSVs, we studied the alkene chains (C_nH_{n+2}) and carboxylic acids ($C_nH_{2n+1}COOH$) in the TZVP basis as model systems. For both model systems, we studied the first excited singlet state and used a PNO threshold of 10^{-7}. The total number of elements in the different sets of virtual orbitals is plotted in Fig. 8 as a function of the system size. The results for the alkene chains are shown in Table IV and on the left-hand side of Fig. 8. For larger systems, the curves behave very similarly for all types of local virtual orbitals and approach, in the asymptotic limit, a linear scaling with the system size (see Table IV). The number of TNOs converges slower toward the linear scaling limit than the number of PNOs and OSVs. In particular, the scaling of the number of excited-state TNOs converges slower than for the ground-state TNOs, while the number of ground- and excited-state PNOs and OSVs shows more similar convergence behaviors. This difference is amplified if for the excited state also the triples with only two pair relations between the LMO indices are included because this leads to a larger number of triples and, hence, to a larger number of TNOs in total.

Published under license by AIP Publishing

TABLE III. Statistical measures for the errors in PNO-CC3(T0) and PNO-CC3 excitation energies in the TZVP basis. The abbreviations ME, MAE, and MAX stand for mean error, mean absolute error, and maximum error, respectively, and STD denotes the standard deviation with respect to the ME.

		PNO-CC3(T0)				PNO-CC3			
		ME	MAE	MAX	STD	ME	MAE	MAX	STD
All states	$T_{PNO} = 10^{-6}$	0.064	0.065	0.321	0.053	0.055	0.056	0.222	0.039
	$T_{PNO} = 10^{-7}$	0.022	0.026	0.209	0.031	0.012	0.014	0.070	0.014
	$T_{PNO} = 10^{-8}$	0.010	0.017	0.177	0.025	<0.001	0.005	0.025	0.007
Singly excited states	$T_{PNO} = 10^{-6}$	0.053	0.055	0.226	0.040	0.049	0.050	0.195	0.033
	$T_{PNO} = 10^{-7}$	0.015	0.019	0.093	0.021	0.010	0.012	0.070	0.013
	$T_{PNO} = 10^{-8}$	0.005	0.012	0.052	0.015	−0.001	0.004	0.017	0.005
Doubly excited states	$T_{PNO} = 10^{-6}$	0.138	0.138	0.321	0.072	0.093	0.093	0.222	0.055
	$T_{PNO} = 10^{-7}$	0.070	0.070	0.209	0.045	0.021	0.023	0.051	0.018
	$T_{PNO} = 10^{-8}$	0.048	0.050	0.177	0.042	0.002	0.010	0.025	0.012

For the carboxylic acids, the lowest excitation is localized on the π-system of the carboxy group. The results for the increase in the virtual spaces with the length of the alkyl chain are shown in Fig. 8 (right). While the number of ground-state virtual orbitals approaches, as for the alkene chains, in the asymptotic limit a linear scaling, the number of excited-state TNOs and PNOs converges toward a constant number because they are now restricted to the chromophore region (only LMO triples that fulfill at least one excited state pair relation are included), which in this case is independent of the system size. The excited-state OSVs do not share this feature since we currently do not select the LMOs contributing to the excitations and the excited-state OSV spaces include the ground-state OSVs.

2. Scaling of the Jacobian right transformation

In Sec. IV E 1, it was shown that the number of excited-state TNOs scales linearly with the size of the chromophore as it is required for a linear scaling number of triples amplitudes. To study how the actual wall-time for the Jacobian right transformation scales with the system size and how the Laplace transformation for PNO-CC3 affects the efficiency, we performed calculations on increasingly large alkene chains in the TZVP basis. A PNO threshold of 10^{-7} was chosen, and for the PNO-CC3 calculations, three Laplace grid points ($n_L = 3$) were used. All calculations were done on a machine with two Intel Xeon Gold 6230 processors (20 cores, 2.1 GHz each), 192 GB of DDR4 RAM, and two 6 TB HDDs (combined to a RAID-0).

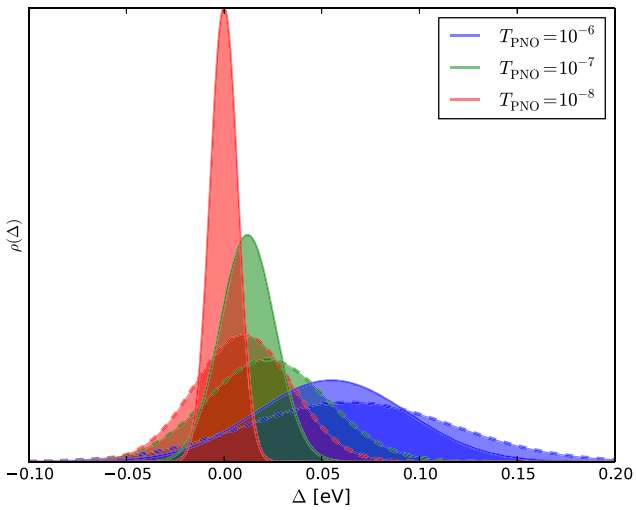


FIG. 6. Normal distributions of the deviations from the canonical reference results centered around the ME. The dashed lines indicate the PNO-CC3(T0) results, while the solid lines denote the PNO-CC3 results.

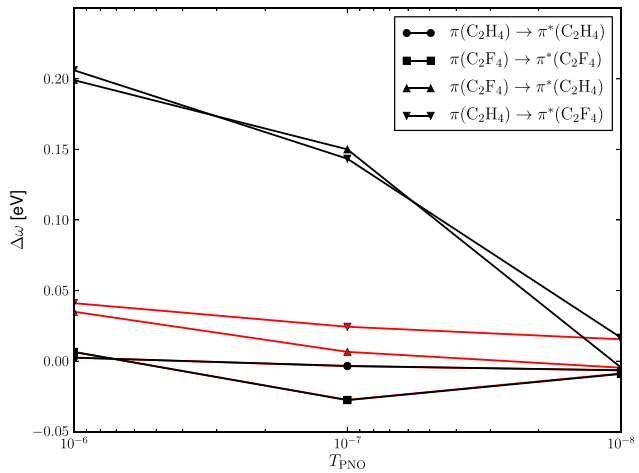


FIG. 7. Comparison of the errors for charge-transfer compared to local excitations for the model system C_2H_4–C_2F_4 in the cc-pVDZ basis set. Black lines denote the results obtained with the excited state triples selection based on three pair relations, and red lines refer to the results from the excited state triples selection based on only two pair relations. The red lines for the local $\pi\pi^*$ excitations are not visible since they are congruent with the black lines.

Published under license by AIP Publishing

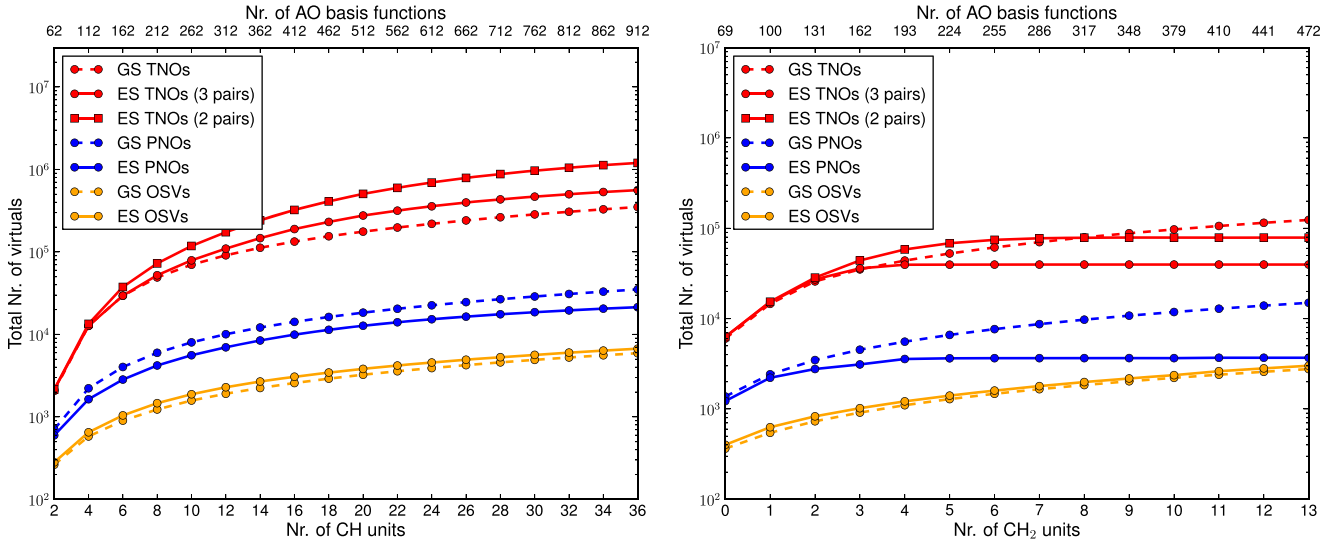


FIG. 8. Increase in the total number of different ground-state (GS) and excited-state (ES) virtual orbitals with respect to the system size. (Left) The results obtained for alkene chains (C_nH_{n+2}). (Right) The results for the carboxylic acids ($C_nH_{2n+1}COOH$).

For canonical CC3, we were not able to do calculations for excitation energies for molecules larger than $C_{14}H_{16}$. Therefore, we estimated the wall-time for the next two larger alkene molecules from the ratio between the wall-time needed for the residual of the ground-state cluster equations and the Jacobian right transformation for $C_{14}H_{16}$ and the wall-time for the ground-state residuals for $C_{16}H_{18}$ and $C_{18}H_{20}$. Linear regression analysis of the timings for $C_{14}H_{16}$ and $C_{18}H_{20}$ show that for such medium-sized molecules, the scaling of the canonical implementation with the system size is approximately $\mathcal{O}(\mathcal{N}^{6.15})$. Thus, the scaling of canonical CC3 is in this regime not yet fully dominated by the $\mathcal{O}(\mathcal{N}^7)$-scaling steps.

For PNO-CC3 and PNO-CC3(T0), the effective cost-scaling for the Jacobian right transformation is in the range of $\mathcal{O}(\mathcal{N}^2)$ and, hence, significantly larger systems are accessible. Even for $C_{36}H_{38}$, the PNO-CC3 calculation is not much more expensive than the canonical CC3 calculation for $C_{14}H_{16}$. Figure 9 shows that PNO-CC3 calculations with $n_L = 3$ take approximately three times longer than PNO-CC3(T0). For both approaches, the first half-transformation of three-index integrals from the AO to the

TABLE IV. Power laws for the dependence of the number of different ground-state (GS) and excited-state (ES) virtual orbitals on the system/chromophore size for the alkene chains obtained from a linear regression analysis ($C_{32}H_{34}$–$C_{36}H_{38}$).

Orbital type	Scaling
ES TNOs (two pairs)	$10^{6.06} \cdot \mathcal{N}^{1.16}$
ES TNOs (three pairs)	$10^{6.52} \cdot \mathcal{N}^{0.99}$
GS TNOs	$10^{4.97} \cdot \mathcal{N}^{1.14}$
ES PNOs	$10^{4.73} \cdot \mathcal{N}^{0.77}$
GS PNOs	$10^{3.01} \cdot \mathcal{N}^{1.09}$
ES OSVs	$10^{2.42} \cdot \mathcal{N}^{0.94}$
GS OSVs	$10^{1.65} \cdot \mathcal{N}^{1.03}$

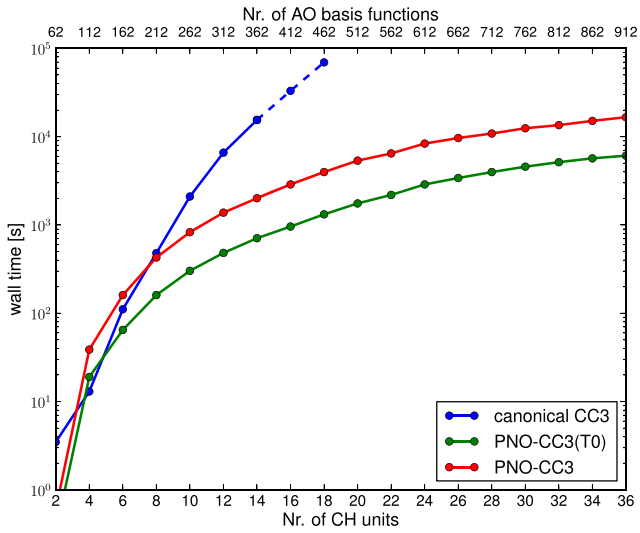


FIG. 9. The total wall-time as a function of the system size obtained from canonical CC3, PNO-CC3(T0), and PNO-CC3 calculations on alkene chains. For canonical CC3, the last two points were obtained by considering the fraction between the wall-time needed for the computation of the ground state vector function T_{GS} and the wall-time needed for the Jacobian right transformation T_{ES} for $C_{14}H_{16}$: $\frac{T_{ES}}{T_{GS}} = 1.76$. Since, for canonical CC3, the computation of the Jacobian right transformation for $C_{16}H_{18}$ and $C_{18}H_{20}$ was not possible, these wall-times were estimated as $T_{GS} * 1.76 = T_{ES}$.

Published under license by AIP Publishing

TABLE V. Power laws for the computational time of Jacobian right transformations for canonical CC3, PNO-CC3(T0), and PNO-CC3 obtained from a linear regression of the curves shown in Fig. 9. For details, see text and Fig. 9.

Method	Scaling
Canonical CC3	$10^{-26.60} \cdot \mathcal{N}^{6.15}$
PNO-CC3(T0)	$10^{-1.19} \cdot \mathcal{N}^{1.45}$
PNO-CC3	$10^{-2.36} \cdot \mathcal{N}^{1.78}$

TNO basis, which in our current implementation scales asymptotically as $\mathcal{O}(\mathcal{N}^2)$, is the most expensive step. For PNO-CC3, this step has to be repeated for each grid point of the numerical Laplace transformation. For PNO-CC3(T0), the wall-time needed to compute these integrals is less dominant compared to the linear scaling contributions to the Jacobian-transformed vectors, which explains the differences in the scaling behaviors shown in Table V.

Even though we have observed an effective scaling of the Jacobian right transformation for PNO-CC3 and PNO-CC3(T0), which is in the range of $\mathcal{O}(\mathcal{N}^2)$ for the system sizes considered here, it should be stressed that the underlying PNO-CCSD implementation formally still scales as $\mathcal{O}(\mathcal{N}^3)$, as described in Ref. 76. This scaling behavior is not seen here because of the large pre-factor for the triples contribution.

V. LARGE-SCALE APPLICATION: S_1 IN BERENIL

As a first application for the PNO-CC3 implementation, we chose the lowest excited singlet state of berenil, an aromatic triazene, at the S_1 minimum geometry found in Ref. 102 with ADC(2). The authors of the latter work argued[102,103] that the experimentally observed ultrafast radiationless decay from S_1 to S_0 is mediated by a small gap or even a conical intersection close to the S_1 minimum geometry. At the ADC(2) level, the energy of the S_1 state falls at its minimum structure below that of the ground state (S_0), which at this level is described with MP2. The negative excitation energy at the ADC(2) level, which is an artifact of shortcomings of this computational level, was in Ref. 102 interpreted as an indication for a very small S_0–S_1 gap or a conical intersection. Because of the size of this molecule, it was not feasible to check with a higher-order order method if this interpretation is correct or if ADC(2) gave a completely wrong result because of a breakdown of the underlying single-reference second-order perturbation theory treatment.

With the PNO-CC3 implementation, we have now a reliable tool at hand to investigate this question. We computed the energy for the first excited singlet state of berenil at the S_1 minimum geometry found at the ADC(2) level. For these calculations, we used the method hierarchy PNO-CC2, PNO-CCSD, and PNO-CC3 and the basis sets aug-cc-pVDZ (aDZ, 636 functions) and aug-cc-pVTZ (aTZ, 1357 functions). In addition to vacuum calculations, we also did calculations where we included solvent effects for an aqueous solution within the conductor-like screening model (COSMO)[104] and the perturbation theory with self-consistent energy (PTE) reaction field scheme with the frozen solvent approximation for the excitation process, in the following denoted as PTE-COSMO.[105,106] For the vacuum as well as for the PTE-COSMO calculations, we used the geometry that was optimized in vacuum [see Fig. 10(a)]. Since the excited state of interest has some charge-transfer character [see the occupied and virtual NTOs depicted in Figs. 10(b) and 10(c)], we performed PNO-CC3 calculations with triple selections based on two and three pair relations. The differences in the excitation energies were less than 0.02 eV in the aDZ basis and, hence, almost negligible. The small difference is probably explained by the short-range intramolecular character of the charge-transfer. This hints to the fact that a triples selection based on only two pair-relations is only necessary for long-range charge-transfer states.

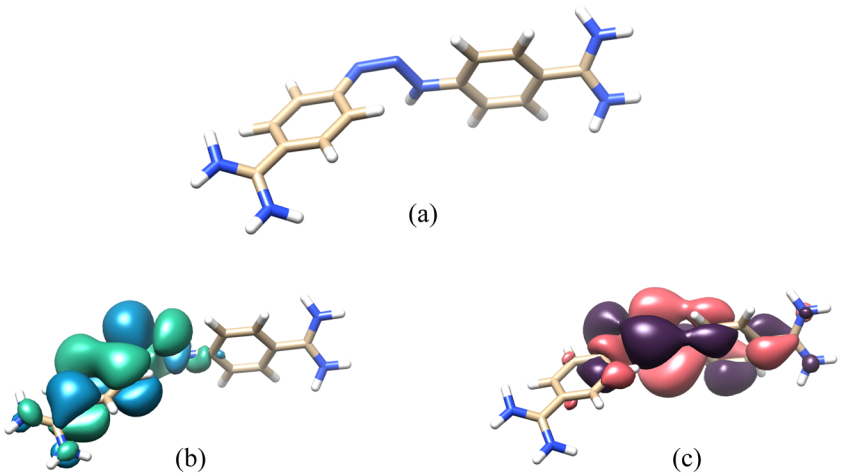


FIG. 10. Molecular structure of the S_1 minimum geometry of berenil optimized in vacuum at the ADC(2) level of theory (a) and the occupied (b) and virtual NTOs (c) obtained from a PNO-CC3/aTZ calculation in vacuum.

Published under license by AIP Publishing

TABLE VI. Excitation energies for S_1 of berenil at the S_1 minimum geometry with and without solvation effects. The diagnostic $\mathcal{T}_1$ is defined as $\mathcal{T}_1 = \frac{|\mathbf{t}_1|}{N_{act}}$, where N_{act} = 106 is the number of correlated electrons (here, we only give the $\mathcal{T}_1$ for the vacuum calculations). For the COSMO calculations, we used for the dielectric constant ϵ = 80.4 and for the refractive index n = 1.3. For the PNO-based calculations, we used $T_{PNO} = T_{TNO} = 10^{-7}$. The values in parentheses refer to the post-SCF-COSMO reaction field scheme.[107]

Basis		ADC(2)	CC2	PNO-CC2	PNO-CCSD	PNO-CC3
aug-cc-pVDZ	ΔE (eV) (vacuum)	-0.20	0.29	0.21	0.39	0.25
	ΔE (eV) (water)	0.17 (0.16)	0.51	0.49	0.73	0.57
	$\mathcal{T}_1$	...	0.0037	0.0023	0.0018	0.0019
aug-cc-pVTZ	ΔE (eV) (vacuum)	-0.22	0.28	0.20	0.41	0.16
	ΔE (eV) (water)	0.14 (0.14)	0.50	0.47	0.74	0.49
	$\mathcal{T}_1$	...	0.0037	0.0023	0.0018	0.0018

As seen from the results collected in Table VI, the excitation energy for the S_1 state increases by the inclusion of solvent effects for all methods and basis sets by ≈0.3 eV. The results are very stable with the basis set. Furthermore, the post-SCF treatment of the solvent at the COSMO-ADC(2) level of theory basically yields the same excitation energies as the PTE-COSMO calculations, indicating that the most important solvent effects are already captured at the PTE-COSMO level. For PNO-CC2 and PNO-CCSD, the changes from aDZ to aTZ are with ±0.02 eV essentially negligible. The strongest basis set effect is observed for PNO-CC3 where the excitation energy is decreased by ≈0.08 eV from the aDZ to the aTZ basis.

To obtain an estimate of the errors introduced by the local approximations, we also performed canonical CC2 calculations. In solution, the CC2 and PNO-CC2 results are in very good agreement (deviations of about 0.02 eV–0.03 eV). However, the PNO errors in vacuum are somewhat larger (up to 0.08 eV). An explanation might be that the ground state singles amplitudes in vacuum are much more important than the ground state singles amplitudes in the PTE-COSMO treatment, which also explains the breakdown of the ADC(2) method in vacuum. Since we do not account for the ground state singles amplitudes in our PNO construction, the errors of PNO-based methods also increase as the importance of the ground state singles amplitudes increases.

However, our results show that in difference to ADC(2), the CC models give positive excitation energies, although they are low: for all models in gas phase and aqueous solution below 1 eV. From CC2 to CCSD, the excitation energy increases by about 0.2 eV. With the inclusion of triple excitations in PNO-CC3, they decrease and are almost back to the PNO-CC2 results, a trend that is often observed for single excitations in organic chromophores. In the aDZ basis, the PNO-CC3 results are in between the PNO-CC2 and PNO-CCSD values; in the aTZ basis, the PNO-CC3 calculation predicts in vacuum the smallest S_0–S_1 gap of 0.16 eV; in water, the result is almost identical with the PNO-CC2 value.

Generally, the PNO-CC2 and PNO-CC3 results are in strikingly good agreement for both basis set and with and without solvent effects. This indicates that the description of the S_1 potential energy surface (PES) in berenil with second-order methods, as done in Refs. 102 and 103, gives at least a qualitatively correct picture in so far as at the minimum structure found with ADC(2) the S_0–S_1 gap is indeed very small and there might be a conical intersection close by, in particular, in vacuum. The inclusion of triple excitations does not significantly change the results relative to CC2. The PNO-CC3 results corroborate the conclusions drawn in Ref. 102. In addition, the $\mathcal{T}_1$-diagnostics are clearly far below any proposed threshold, which would indicate the necessity of a multireference treatment ($\mathcal{T}_1 \approx 0.02$). However, for accurate excitation energies in the region of the small S_0–S_1 gap, it is important to go beyond ADC(2) and include the ground-state singles amplitudes. Beyond CC2, the results are stable and the small gap between S_1 and S_0 is not caused by an inappropriate account of electron correlation effects.

The calculation we presented here are, to the best of our knowledge, by far the largest CC3 calculations for excitation energies ever performed (more than 1300 basis function in the aTZ basis).

VI. SUMMARY AND CONCLUSIONS

We presented the first implementation of a local correlation method for excitation energies with connected triples. To circumvent also in the local implementation the storage of triples amplitudes, we employ a Laplace transformation of the energy denominators for the ground- and the excited-state triples. The PNO-based implementation of CC3 excitation energies shows an effective scaling of $\mathcal{O}(\mathcal{N}^2)$ with the system size for the range of molecules (alkene chains) for which calculations are currently feasible. For PNO-CC3, the pre-factor is roughly by the number of Laplace grid points larger than for PNO-CC3(T0). The number of excited-state TNOs scales linearly with the chromophore size, although the size of excited-state TNO basis per selected triples is larger than for the ground-state due to the combination of ground and excited-state intermediates in the triples amplitudes.

Based on a multipole analysis of the contributions to the excited-state triples amplitudes, the criteria for the selection of the LMO triples *ijk* for local and long-ranging excitations are derived. For local excitations, the criteria are similar as for the electronic ground state. For long-ranging, e.g., charge-transfer, excitations relaxed criteria with only two pair relations between the

Published under license by AIP Publishing

LMO indices *i*, *j*, and *k* have to be used. These formal findings have been verified for local and charge-transfer excitations in the intermolecular dimer C_2H_4–C_2F_4 that has been used as a model system.

Benchmark calculations for a large and diverse test set with 146 singlet excitations showed that PNO-CC3 gives already, with moderate PNO and TNO thresholds of 10^{-7}–10^{-8}, mean absolute errors of about 0.01 eV, which are much below the intrinsic error of the CC3 model. For single excitations also, the computationally cheaper PNO-CC3(T0) method gives relatively accurate results. For states with significant double excitation contributions, the errors of the T0 approximation are somewhat larger, but still small compared to the intrinsic error of CC3, which for such states are also larger.

Finally, we applied our PNO-CC3 implementation to the first excited state of berenil at the S_1 minimum geometry to check if the results obtained previously with the second order methods CC2 and ADC(2) in Refs. 102 and 103 were qualitatively correct. Our calculations showed that PNO-CC2 and PNO-CC3 excitation energies are in very good agreement for this system and state. Hence, it can be concluded that the second-order treatment of this system is justified and the first excited state of berenil at the S_1 minimum geometry has a small S_0–S_1 gap, which explains the ultrafast radiationless decay found in the experiments.

The results presented here pave the way for the routine application of coupled cluster models including triple excitations for the calculation of excitation energies for medium-sized molecules. The calculations presented here for berenil in the aTZ basis with more than 1300 basis functions were out of reach, or at least extremely costly, with the available canonical CC3 implementations.

ACKNOWLEDGMENTS

This work was supported by the Deutsche Forschungsgemeinschaft (DFG) under Grant No. HA 2588/7. The authors would like to thank Alireza Marefat Khah for fruitful discussions on berenil and for providing the S_1 structure.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1] M. Ziolkwski, B. Jansik, T. Kjaergaard, and P. Jørgensen, J. Chem. Phys. **133**, 014107 (2010).
[2] K. Kristensen, M. Ziółkowski, B. Jansík, T. Kjærgaard, and P. Jørgensen, J. Chem. Theory Comput. **7**, 1677 (2011).
[3] I. M. Hoyvik, K. Kristensen, B. Jansik, and P. Jørgensen, J. Chem. Phys. **136**, 014105 (2012).
[4] K. Kristensen, T. Kjærgaard, I.-M. Høyvik, P. Ettenhuber, P. Jørgensen, B. Jansik, S. Reine, and J. Jakowski, Mol. Phys. **111**, 1196–1210 (2013).
[5] S. Li, J. Ma, and Y. Jiang, J. Comput. Chem. **23**, 237 (2002).
[6] S. Li, J. Shen, W. Li, and Y. Jiang, J. Chem. Phys. **125**, 074109 (2006).
[7] W. Li, P. Piecuch, J. R. Gour, and S. Li, J. Chem. Phys. **131**, 114109 (2009).
[8] W. Li and P. Piecuch, J. Phys. Chem. A **114**, 8644 (2010).
[9] W. Li and P. Piecuch, J. Phys. Chem. A **114**, 6721 (2010).
[10] Y. Guo, W. Li, and S. Li, J. Phys. Chem. A **118**, 8996 (2014).
[11] W. Li and S. Li, Sci. China Chem. **57**, 78 (2014).
[12] G. Rauhut, P. Pulay, and H.-J. Werner, J. Comput. Chem. **19**, 1241 (1998).
[13] M. Schütz, G. Hetzer, and H.-J. Werner, J. Chem. Phys. **111**, 5691 (1999).
[14] C. Hampel and H. J. Werner, J. Chem. Phys. **104**, 6286 (1996).
[15] M. Schütz, Phys. Chem. Chem. Phys. **4**, 3941 (2002).
[16] M. Schütz and F. R. Manby, Phys. Chem. Chem. Phys. **5**, 3349 (2003).
[17] M. Schütz and H.-J. Werner, Chem. Phys. Lett. **318**, 370 (2000).
[18] M. Schütz, J. Chem. Phys. **113**, 9986 (2000).
[19] H. Stoll, J. Chem. Phys. **97**, 8449 (1992).
[20] H. Stoll, Chem. Phys. Lett. **191**, 548 (1992).
[21] H. Stoll, Phys. Rev. B **46**, 6700 (1992).
[22] J. Friedrich, M. Hanrath, and M. Dolg, J. Chem. Phys. **126**, 154110 (2007).
[23] J. Friedrich and M. Dolg, J. Chem. Phys. **129**, 244105 (2008).
[24] J. Friedrich and M. Dolg, J. Chem. Theory Comput. **5**, 287 (2009).
[25] J. Friedrich, J. Chem. Theory Comput. **8**, 1597 (2012).
[26] J. Friedrich and K. Walczak, J. Chem. Theory Comput. **9**, 408 (2013).
[27] J. Friedrich and J. Hänchen, J. Chem. Theory Comput. **9**, 5381 (2013).
[28] W. Meyer, Int. J. Quantum Chem. **5**(S5), 341 (1971).
[29] W. Meyer, J. Chem. Phys. **58**, 1017 (1973).
[30] W. Meyer, Theor. Chem. Acc. **35**, 277 (1974).
[31] W. Meyer, *Methods of Electronic Structure Theory*, 3rd ed. (Plenum Press, New York, 1977).
[32] R. Ahlrichs, F. Driessler, H. Lischka, V. Staemmler, and W. Kutzelnigg, J. Chem. Phys. **62**, 1235 (1975).
[33] R. Ahlrichs and F. Driessler, Theor. Chem. Acc. **36**, 275 (1975).
[34] V. Staemmler and R. Jaquet, Theor. Chem. Acc. **59**, 487 (1981).
[35] F. Neese, A. Hansen, and D. G. Liakos, J. Chem. Phys. **131**, 064103 (2009).
[36] F. Neese, A. Hansen, F. Wennmohs, and S. Grimme, Acc. Chem. Res. **42**, 641 (2009).
[37] F. Neese, F. Wennmohs, and A. Hansen, J. Chem. Phys. **130**, 114108 (2009).
[38] D. G. Liakos, A. Hansen, and F. Neese, J. Chem. Theory Comput. **7**, 76 (2011).
[39] A. Hansen, D. G. Liakos, and F. Neese, J. Chem. Phys. **135**, 214102 (2011).
[40] C. Riplinger and F. Neese, J. Chem. Phys. **138**, 034106 (2013).
[41] J. L. Whitten, J. Chem. Phys. **58**, 4496 (1973).
[42] B. I. Dunlap, J. W. D. Conolly, and J. R. Sabin, J. Chem. Phys. **71**, 3396 (1979).
[43] D. P. Tew and C. Hättig, Int. J. Quantum Chem. **113**, 224 (2013).
[44] C. Hättig, D. P. Tew, and B. Helmich, J. Chem. Phys. **136**, 204105 (2012).
[45] G. Schmitz, C. Hättig, and D. P. Tew, Phys. Chem. Chem. Phys. **16**, 22167 (2014).
[46] G. Schmitz and C. Hättig, J. Chem. Phys. **145**, 234107 (2016).
[47] C. Krause and H.-J. Werner, Phys. Chem. Chem. Phys. **14**, 7591 (2012).
[48] G. Schmitz, B. Helmich, and C. Hättig, Mol. Phys. **111**, 2463 (2013).
[49] J. Yang, G. K.-L. Chan, F. R. Manby, M. Schütz, and H.-J. Werner, J. Chem. Phys. **136**, 144105 (2012).
[50] Y. Guo, K. Sivalingam, E. F. Valeev, and F. Neese, J. Chem. Phys. **144**, 094111 (2016).
[51] F. Menezes, D. Kats, and H.-J. Werner, J. Chem. Phys. **145**, 124115 (2016).
[52] P. Pulay, Chem. Phys. Lett. **100**, 151 (1983).
[53] S. Saebø and P. Pulay, Chem. Phys. Lett. **113**, 13 (1985).
[54] S. Saebo and P. Pulay, Annu. Rev. Phys. Chem. **44**, 213 (1993).
[55] P. Pulay and S. Saebø, Theor. Chim. Acta **69**, 357 (1986).
[56] S. Saebø and P. Pulay, J. Chem. Phys. **86**, 914 (1987).
[57] S. Saebo and P. Pulay, J. Chem. Phys. **88**, 1884 (1988).
[58] M. S. Frank, G. Schmitz, and C. Hättig, Mol. Phys. **115**, 343 (2017).
[59] P. Pinski and F. Neese, J. Chem. Phys. **148**, 031101 (2018).
[60] P. Pinski and F. Neese, J. Chem. Phys. **150**, 164102 (2019).
[61] O. Christiansen, H. Koch, and P. Jørgensen, Chem. Phys. Lett. **243**, 409 (1995).
[62] D. Kats, T. Korona, and M. Schütz, J. Chem. Phys. **125**, 104106 (2006).
[63] T. Korona and H.-J. Werner, J. Chem. Phys. **118**, 3006 (2003).
[64] T. D. Crawford and R. A. King, Chem. Phys. Lett. **366**, 611 (2002).

Published under license by AIP Publishing

[65]D. Mester, P. R. Nagy, and M. Kállay, J. Chem. Phys. **146**, 194102 (2017).
[66]P. Baudin and K. Kristensen, J. Chem. Phys. **144**, 224106 (2016).
[67]A. K. Dutta, F. Neese, and R. Izsák, J. Chem. Phys. **145**, 034102 (2016).
[68]A. K. Dutta, M. Saitow, C. Riplinger, F. Neese, and R. Izsák, J. Chem. Phys. **148**, 244101 (2018).
[69]A. K. Dutta, M. Saitow, B. Demoulin, F. Neese, and R. Izsák, J. Chem. Phys. **150**, 164123 (2019).
[70]M. Nooijen and R. J. Bartlett, J. Chem. Phys. **107**(17), 6812 (1997).
[71]O. Christiansen, H. Koch, and P. Jørgensen, J. Chem. Phys. **105**, 1451 (1996).
[72]H. Koch, O. Christiansen, P. Jørgensen, A. M. Sanchez de Merás, and T. Helgaker, J. Chem. Phys. **106**, 1808 (1997).
[73]C. Peng, M. C. Clement, and E. F. Valeev, J. Chem. Theory Comput. **14**, 5597 (2018).
[74]B. Helmich and C. Hättig, J. Chem. Phys. **139**, 084114 (2013).
[75]B. Helmich and C. Hättig, Comput. Theor. Chem. **1040-1041**, 35 (2014).
[76]M. S. Frank and C. Hättig, J. Chem. Phys. **148**, 134102 (2018).
[77]O. Christiansen, H. Koch, and P. Jørgensen, J. Chem. Phys. **103**, 7429 (1995).
[78]B. Helmich and C. Hättig, J. Chem. Phys. **135**, 214106 (2011).
[79]H. Koch, O. Christiansen, R. Kobayashi, P. Jørgensen, and T. Helgaker, Chem. Phys. Lett. **228**, 233 (1994).
[80]O. Christiansen, H. Koch, A. Halkier, P. Jørgensen, T. Helgaker, and A. Sánchez de Merás, J. Chem. Phys. **105**, 6921 (1996).
[81]J. Yang, Y. Kurashige, F. R. Manby, and G. K. L. Chan, J. Chem. Phys. **134**(4), 044123 (2011).
[82]C. Riplinger, B. Sandhoefer, A. Hansen, and F. Neese, J. Chem. Phys. **139**, 134101 (2013).
[83]K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon, Chem. Phys. Lett. **157**, 479 (1989).
[84]G. Hetzer, P. Pulay, and H.-J. Werner, Chem. Phys. Lett. **290**, 143 (1998).
[85]G. Schmitz and C. Hättig, J. Chem. Theory Comput. **13**, 2623 (2017).
[86]Y. Guo, C. Riplinger, U. Becker, D. G. Liakos, Y. Minenkov, L. Cavallo, and F. Neese, J. Chem. Phys. **148**, 011101 (2018).
[87]J. Almlöf, Chem. Phys. Lett. **181**, 319 (1991).
[88]M. Häser and J. Almlöf, J. Chem. Phys. **96**, 489 (1992).
[89]A. K. Wilson and J. Almlöf, Theor. Chim. Acta **95**, 49 (1997).
[90]C. Hättig and F. Weigend, J. Chem. Phys. **113**, 5154 (2000).
[91]R. Ahlrichs, M. Bär, M. Häser, H. Horn, and C. Kölmel, Chem. Phys. Lett. **162**, 165 (1989).
[92]M. Häser and R. Ahlrichs, J. Comput. Chem. **10**, 104 (1989).
[93]M. Schreiber, M. R. Silva-Junior, S. P. A. Sauer, and W. Thiel, J. Chem. Phys. **128**, 134110 (2008).
[94]A. Schäfer, C. Huber, and R. Ahlrichs, J. Chem. Phys. **100**, 5829 (1994).
[95]T. H. Dunning, Jr., J. Chem. Phys. **90**, 1007 (1989).
[96]R. A. Kendall, T. H. Dunning, Jr., and R. J. Harrison, J. Chem. Phys. **96**, 6796 (1992).
[97]F. Weigend, A. Köhn, and C. Hättig, J. Chem. Phys. **116**, 3175 (2002).
[98]C. Hättig, Phys. Chem. Chem. Phys. **7**, 59 (2005).
[99]G. Knizia, J. Chem. Theory Comput. **9**, 4834 (2013).
[100]A. Hellweg, C. Hättig, S. Höfener, and W. Klopper, Theor. Chem. Acc. **117**, 587 (2007).
[101]D. Kánnár and P. G. Szalay, J. Chem. Theory Comput. **10**(9), 3757 (2014).
[102]L. Grimmelsmann, A. Marefat Khah, C. Spies, C. Hättig, and P. Nuernberger, J. Phys. Chem. Lett. **8**, 1986 (2017).
[103]A. Marefat Khah, L. Grimmelsmann, J. Knorr, P. Nuernberger, and C. Hättig, Phys. Chem. Chem. Phys. **20**, 28075 (2018).
[104]A. Klamt and G. Schüürmann, J. Chem. Soc., Perkin Trans. 2 **2**, 799 (1993).
[105]S. Corni, R. Cammi, B. Mennucci, and J. Tomasi, J. Chem. Phys. **123**, 134512 (2005).
[106]B. Lunkenheimer and A. Köhn, J. Chem. Theory Comput. **9**, 977 (2013).
[107]S. Karbalaei Khani, A. Marefat Khah, and C. Hättig, Phys. Chem. Chem. Phys. **20**, 16354 (2018).

Published under license by AIP Publishing