Applications of a perturbation-aware local correlation method to coupled cluster linear response properties

Ruhee D'Cunha[a] and T. Daniel Crawford[a,b]

[a]Department of Chemistry, Virginia Tech, Blacksburg, VA 24061, USA; [b]Molecular Sciences Software Institute, 1880 Pratt Drive, Suite 1100, Blacksburg, VA 24060, USA

ARTICLE HISTORY
Compiled August 4, 2022

ABSTRACT
We have investigated the efficacy of two recently proposed variations of the pair-natural-orbital approach to reducing the scaling of coupled cluster property calculations. In particular, we have extended our implementations of the PNO++ and combined PNO++ methods, which make use of field-aware pair-densities to define the virtual-orbital spaces used to describe electron correlation effects, in order to test their accuracy, efficiency, and robustness on larger molecular systems than previously investigated. For fluoroalkane chains up to 1-fluoroheptane we find that the PNO++ and combined PNO++ methods yield smaller truncation errors in response properties than PNO for similarly compact virtual spaces, and, while the PNO method performs better than the PNO++ method for correlation energies, the combined PNO++ method recovers similar accuracy for correlation energies to the PNO method. For more three-dimensional molecular structures such as α- and $\beta-$pinenes, the PNO, PNO++, and combined PNO++ methods all yield similar errors for response properties, whereas for (S)-1-phenylethanol, the PNO method performs slightly better than the other two approaches. We also investigate the use of a product density to define the virtual space, as well as two candidates for defining weak-pair contributions.

KEYWORDS
coupled cluster theory; linear response; local correlation

1. Introduction

Coupled-cluster theory is one of the most popular methods used for electronic structure calculations, due to its expected accuracy and systematic improvability[1, 2], which arise though its exponential form of the wave function and convergence to the exact, nonrelativistic solution to the electronic Schrödinger equation within a finite basis set. However, this performance comes at a substantial price, as coupled cluster suffers from steep scaling of its computational cost with system size. For example, coupled cluster wave functions truncated at the single- and double-excitation (CCSD) level scales as $\mathcal{O}(N)^6$, while including triples excitations (CCSDT) increases the scaling to $\mathcal{O}(N)^8$, where N is a measure of the size of the system.[3]

The high-degree polynomial scaling problem has been tackled by the development of numerous methods over the last several decades, including fragmentation

CONTACT T. Daniel Crawford. Email: crawdad@vt.edu

approaches,[4–7] tensor decompositions,[8–13], localization schemes[14–19] and more, each with its own advantages and disadvantages. The main motivation for localization methods, in particular, is that the dynamic electron correlation effects that coupled cluster aims to capture are relatively short-ranged, and thus a localized basis would better represent the sparsity present in the wave function than the usual delocalized molecular orbital (MO) basis. Local-correlation schemes to reduce the size of the wave function parameter space, such as projected atomic orbitals (PAOs)[14, 15, 20], pair natural orbitals (PNOs)[16, 21–25], and orbital specific virtuals (OSVs)[19] have been successfully applied to coupled cluster ground state energies, and several methods have been developed in order to apply them to excited state energies[26–32]. Recently, PNOs have been utilized in methods to compute excited state energies of complex molecular systems.[33–37] However, their application to response properties has been limited[38–43] due to the observed sensitivity of such properties — especially mixed electric-field/magnetic-field responses — to the definition of the correlation space[44].

Russ and Crawford investigated an alternative domain-selection scheme for PAOs that incorporated the perturbation into the virtual-orbital domain definition[45, 46] and found that the computational crossover points between canonical and local schemes lay at much larger molecules than for ground-state energies. Similarly, McAlexander and Crawford[47], in a first application of PNOs to coupled-cluster linear-response properties, found that, due to the reduced sparsity of the field-perturbed wave function in the localized orbital basis, the sensitivity of response properties to truncation of the wave function basis was much higher than for properties requiring only the unperturbed state. Kumar and Crawford[48], exploring the effect of truncation of a natural orbital space on linear response properties, saw an inverse relationship between the diffuseness of the orbitals and their occupation number, and thus the elimination of low occupation number orbitals in such schemes leads to the removal of more diffuse orbitals, which are essential for accurate modeling of response properties. Consideration of the character of target electronically excited states led to the creation of transition-specific natural orbitals by Høyvik, Myhre and Koch[49], and the use of a similar approach by Baudin and Kristensen, in combination with the local framework to reduce the scaling of CC2-level excitation energy calculations without significant loss of accuracy[50, 51]. Höfener and Klopper formed effective natural transition orbitals using ground and excited state densities[52], while Mester, Nagy, and Kállay combined MP2 and CIS(D) densities to produce state-specific natural orbitals for excitation energy calculations[53, 54]. Variants of natural orbitals have also been used in order to reduce the scaling of equation-of-motion coupled cluster methods.[55–57] More recently, frozen natural spinors have been applied in a similar way to four-component relativistic coupled cluster theory.[58]

In our group's most recent work, the incorporation of external perturbations into the construction of the virtual-orbital space for the PNO method led us to the creation and exploration of the PNO++ approach for coupled cluster response properties.[59, 60] The PNO++ method described in Ref 60 has shown encouraging results for small, localizable molecular systems. In addition, the "combined PNO++" method, which includes a fixed number of PNOs in the final orbital space, yields smaller truncation errors in the correlation energy than the PNO++ method alone — as well as smaller errors in the response properties studied than the PNO method alone. In our pilot, Python-based implementation of these new methods, we used a canonical-MO formulation of the CC response equations, but then simulated the effect of the local truncations by filtering out non-local contributions in each iterative step. (We have described such simulations in detail in previous work.[47, 60]) However, in order for

either method to be implemented at production-level, validation is necessary on larger systems than are feasible in our original Python-based code.

The sensitive nature of response properties means that a larger proportion of the orbital space needs to be kept to maintain accuracy in the value of the property. However, a key efficiency characteristic of the PNO method is aggressive truncation of the virtual space in order to avoid increasing the computational cost of the initial integral transformation into the PNO basis. Thus, a production-level code requires careful consideration as to an optimal method, including the ability to truncate aggressively for larger, more delocalized systems. The simulation code described above, on the other hand, is essentially identical in cost to a canonical coupled cluster calculation, because it only incorporates a single extra transformation (to filter out non-local contributions) at each iteration. A simulation code implemented in an optimized, canonical-MO coupled cluster linear response code, such as the one present in Psi4,[61] provides an effective testing environment for further approximations as well as improvements to the method for specific properties, as described below.

In this work, we apply the PNO++ and combined PNO++ methods as implemented as a simulation code in Psi4 to larger organic molecules in order to validate the methods and compare them to the performance of the PNO method for the same systems. We also explore refinements to the PNO++ method, including a new formulation of the PNO++ density specifically for optical rotation calculations, and the application of energy- and perturbation-based weak pair approximations to reduce the computational expense of the methods.

2. Theory and Computational Details

2.1. Larger Benchmark Calculations

Large-molecule benchmark calculations were carried out using an implementation of a local correlation simulation code in a development version of the Psi4 program package.[61] In this code, non-local contributions to a given residual vector were filtered out at each iteration in the coupled cluster amplitude, left-hand amplitude, and perturbed wave function equations. Calculations on small molecule systems were carried out using a Python-based simulation code. (For further details of the simulation see Ref 60.) The combined PNO++ method had its unperturbed T_{cutPNO} threshold fixed to 10^{-6}, which was determined in our previous study to be a reasonable value when combined with the PNO++ space. The perturbed $T_{cutPNO++}$ threshold was varied in order to study the effect of the truncation. All molecular geometries were optimized at the B3LYP[62–64]/aug-cc-pVDZ[65, 66] level using Psi4's optking module. The molecules studied in this work are given in Fig 1 and include two hydrogen molecule helices, H_2O_2, 1,3-dimethylallene (DMA), a set of fluoroalkane chains, a phenyl-substituted alcohol, and two bicyclic alkenes. Dipole polarizabilities and specific rotations were computed using CCSD linear response[67, 68] and an external field wavelength of 589 nm, with and without the local simulation applied. The specific rotation was computed using both the length and the velocity gauge representation of the dipole moment operator, with a shift to zero field frequency to obtain the modified velocity gauge value of the specific rotation.[69, 70] All response calculations used the augmented correlation-consistent double zeta basis set of Dunning and coworkers, aug-cc-pVDZ.[71]

In order to estimate the computational savings obtained via truncation of the PNO,

PNO++ and combined PNO++ spaces, we use a T_2 ratio. This is given as the ratio of the number of truncated T_2 amplitudes to the full untruncated number. This gives us an upper bound to the storage requirements as well as to the cost of the highest-scaling term in the CCSD amplitude equations, an $\mathcal{O}(N^6)$ operation containing T_2 amplitudes.

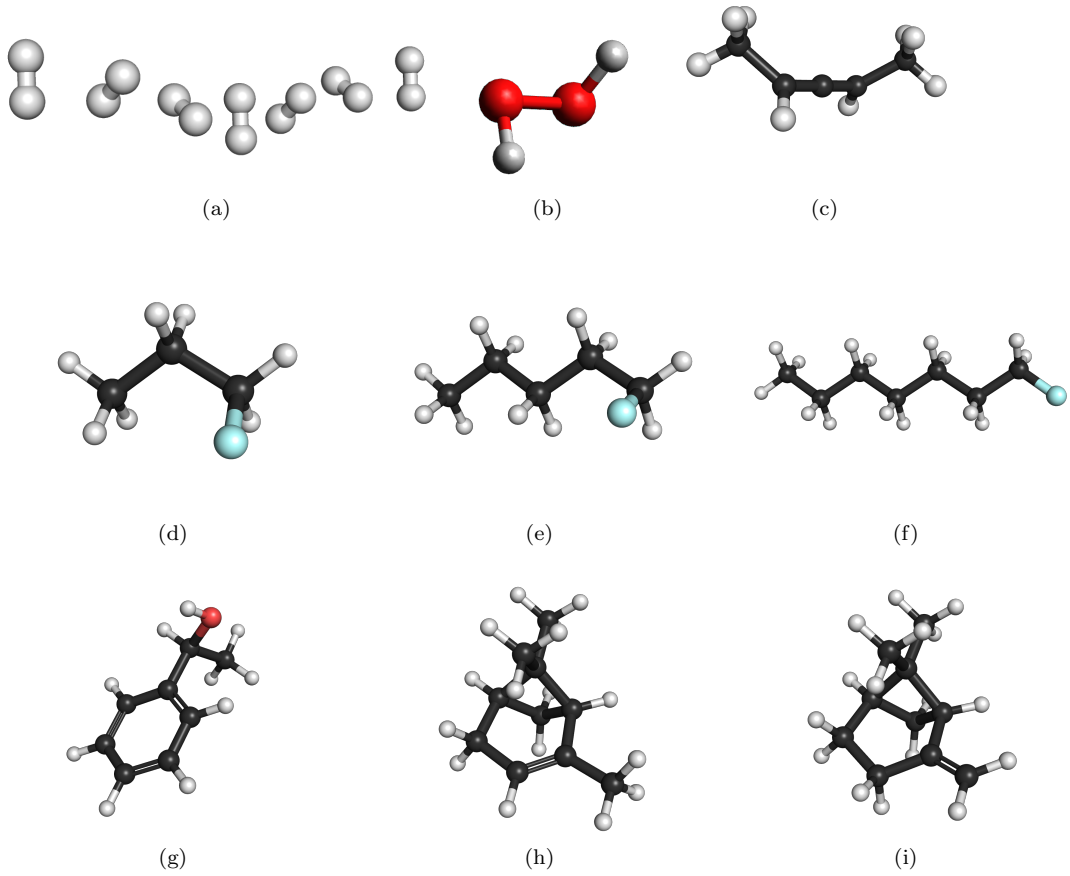


Figure 1. Molecular systems studied in this work: (a) an $(H_2)_7$ helix, (b) H_2O_2, (c) (*P*)-1,3-dimethylallene (DMA) (d) (*M*)-1-fluoropropane, (e) (*M*)-1-fluoropentane, (f) (*M*)-1-fluoroheptane, (g) (*S*)-1-phenylethanol, (h) (1*R*,5*R*)-α-pinene, and (i) (1*R*,5*R*)-β-pinene. All optimized geometries can be found in the Supporting Information.

2.2. Product-based densities

Pair densities in the PNO and PNO++ approaches are formed using MP2-level amplitudes in the form of the MP2 reduced density matrix. Spaces of virtual natural orbitals are then formed by diagonalization of the pair densities, with the PNO++ density including some measure of the external perturbation in order to improve the compactness of the space for response properties. However, the PNO++ density, written as

$$\boldsymbol{D}^{ij}(B,\omega) = \frac{2}{1+\delta_{ij}}\left(\boldsymbol{X}_B^{ij}\widetilde{\boldsymbol{X}}_B^{ij\dagger} + \boldsymbol{X}_B^{ij\dagger}\widetilde{\boldsymbol{X}}_B^{ij}\right) \tag{1}$$

only contains dependence on a single perturbation B at a given frequency ω, e.g. the electric dipole operator $\hat{\mu}$ or the magnetic dipole operator $\hat{m}$. Here, $\boldsymbol{X}_B^{ij}$ is an MP2-level approximation to the first-order perturbed T_2 amplitudes corresponding to the external perturbation B:

$$X_{ab}^{ij}(B,\omega) = \frac{\bar{B}}{\bar{H}_{ii} + \bar{H}_{jj} - \bar{H}_{aa} - \bar{H}_{bb} + \omega}. \tag{2}$$

The optical rotation tensor $\beta(\omega)$, on the other hand, requires both of these operators:

$$\beta(\omega) = \mathrm{Im}\langle\langle\hat{\mu};\hat{m}\rangle\rangle_\omega. \tag{3}$$

The linear response function can further be formulated in terms of the first-order perturbed density as

$$\langle\langle\hat{\mu};\hat{m}\rangle\rangle_\omega = \frac{1}{2}\hat{P}(\hat{\mu},\hat{m})\left[\sum_{pq}\hat{\mu}_{pq}\left[D_{pq}^{\hat{m}^\omega}\right]^{(1)}\right], \tag{4}$$

where the permutation operator $\hat{P}$ ensures that the linear response function contains both $[D_{pq}^{\hat{\mu}^\omega}]^{(1)}$ and $[D_{pq}^{\hat{m}^\omega}]^{(1)}$. Further, these perturbed coupled cluster densities are multiplied by the opposite operator, i.e. the density derived from $\hat{\mu}$ is multiplied by $[D_{pq}^{\hat{m}^\omega}]^{(1)}$ and vice versa. The MP2-level perturbed pair densities, as reasonable approximations to the perturbed coupled cluster densities, can thus be used in an element-wise (Hadamard) product in order to create a space optimized for the mixed linear-response function, i.e.,

$$\boldsymbol{D}^{ij} = \boldsymbol{D}^{ij}(\hat{\mu},\omega) \circ \boldsymbol{D}^{ij}(\hat{m},\omega). \tag{5}$$

Virtual orbitals created using this product density should, in principle, contain more information about the response of the system in the case of the mixed response function in Eq (3) above, leading to a more compact space for the specific rotation.

2.3. Weak Pairs

The sparsity of the localized occupied-MO basis can be exploited by treating at a lower level or neglecting certain sets of PNOs. For example, those PNOs that correspond to the pairs of occupied orbitals ij whose approximated contributions to the quantity of interest are smaller than a predetermined cutoff can be neglected or their contributions approximated. These "weak pairs." are generally treated at a lower level,[15, 20, 72] however, we have chosen to neglect them to see more clearly the effect of their truncation on correlation energies and properties. For the PNO method, this criterion takes the form

$$|\epsilon_{ij}| < T_{cutPairs}, \tag{6}$$

where ϵ_{ij} is defined as the pair correlation energy as typically computed at the MP2 level of theory. The method works by neglecting or otherwise treating with a less

expensive method the pairs whose integrals,

$$\int \phi_i^*(r_1)\phi_j^*(r_2)\frac{1}{r_{12}}\phi_a(r_1)\phi_b(r_2)dr_1dr_2, \tag{7}$$

are negligible, and thus whose contribution to the energy is also expected to be negligible. Here, the pairs ij are occupied, with the pairs ab being virtual orbitals. This is true when the orbitals are sufficiently distant spatially to have minimal overlap. This works well for localized occupied orbitals in large molecules, and thus the number of pairs and therefore the total number of wave function parameters to be computed and stored can be reduced to a much more computationally efficient level.

Previous studies neglecting weak pairs using the energy criterion found that adding a second truncation threshold in addition to the virtual-space truncation led to larger errors in response properties at similar fractions of the space kept.[47] A similar perturbation-including criterion can be formulated as

$$|\bar{\mu}_{ij}| < T_{cutPairs}, \tag{8}$$

where

$$\bar{\mu}_{ij} = \sum_{ab} \bar{\mu}_{ij}^{ab}, \tag{9}$$

with $\bar{\mu}_{ij}^{ab}$ being the similarity-transformed perturbation using MP2-level doubles amplitudes,

$$\bar{\mu}_{ij} = e^{-T_2}\mu e^{T_2}. \tag{10}$$

In this work, we retain the full PNO space while neglecting contributions from the weak pairs, determined either using the MP2 pair correlation energy or the MP2-level similarity transformed perturbation, in order to quantify the loss in accuracy due to truncation of the weak pairs alone. The goal is to combine this perturbation-including weak pair criterion with the PNO++ method in order to truncate the space even further than in previous work and attempt to maintain accuracy in the correlation energy as well as response properties.

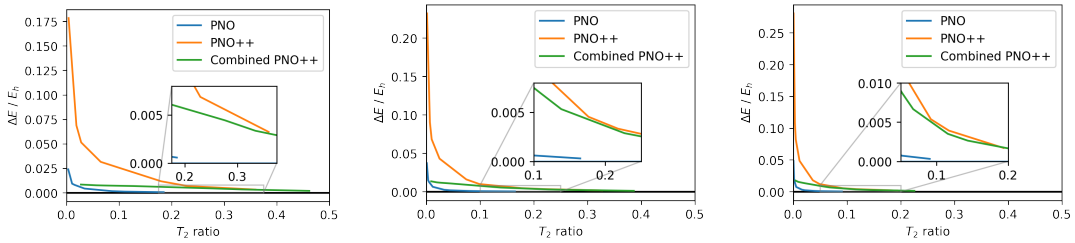
3. Results and Discussion

3.1. Larger Benchmark Calculations

3.1.1. Fluoroalkane Chains

Figures 2-4 illustrate the effect of truncation of the PNO, PNO++ and combined spaces on the errors in correlation energy, dipole polarizabilities and specific rotations, respectively, for 1-fluoropropane, -pentane and -heptane. The systems are linear and can provide information about the effect of an increasing linear system size on the truncation errors in energy and properties. In each case, the truncation errors are plotted as a function of the T_2 ratio, as described in Section 2.1, and can be used as a measure of the potential computational savings expected by the corresponding truncation of the space.

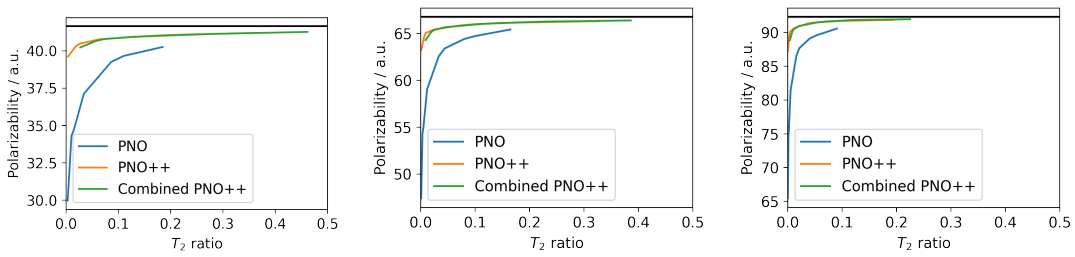
Across the substituted propane, pentane, and heptane molecular species, we see that the truncation errors in correlation energy are larger for the PNO++ method, with the PNO method achieving values within chemical accuracy of the CCSD reference at T_2 ratios of 0.11, 0.04 and 0.02 respectively, while the PNO++ requires T_2 ratios of more than 0.36, 0.22 and 0.19 for the same result. The combined method, however, with a set threshold for unperturbed PNOs to be included in the space, regains a significant portion of the accuracy of the PNO method at all truncations.



(a) (b) (c)

Figure 2. Truncation errors in CCSD correlation energy in Hartree for (a) (*M*)-1-fluoropropane, (b) (*M*)-1-fluoropentane and (c) (*M*)-1-fluoroheptane systems, computed using the aug-cc-pVDZ basis set.

Figure 3 depicts the convergence in the values of the dipole polarizabilities with respect to the T_2 ratio, with the error for the PNO method being larger than that for the PNO++ and combined methods at each truncation. The PNO++ and combined methods show a slowly-tapering error, reaching within 5% of the reference value at a T_2 ratio of 0.03 but naturally requiring a larger amount of the space for tighter convergence to the reference value.

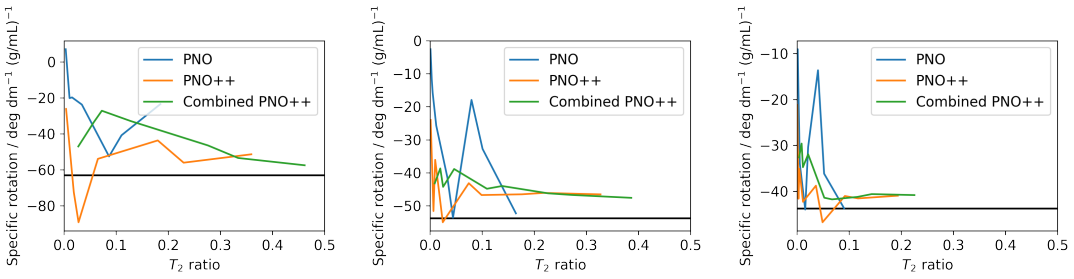


(a) (b) (c)

Figure 3. CCSD linear response dynamic polarizabilities at 589 nm in a.u. for (a) (*M*)-1-fluoropropae, (b) (*M*)-1-fluoropentane and (c) (*M*)-1-fluoroheptane systems, computed using the aug-cc-pVDZ basis set.

Specific rotations computed for the fluoroalkane systems, which are optically active due to a twist in the methyl group containing the fluorine away from the molecular mirror plane, can be seen in Figure 4. The data exhibit better convergence for the PNO++ and combined PNO++ methods as compared to the PNO method, at T_2 ratios below 0.2 for the smaller 1-fluoropropane and 1-fluoropentane systems, and below 0.1 for the larger 1-fluoroheptane system. The large fluctuations in specific rotation values seen with PNOs become smaller fluctuations with the PNO++ and combined PNO++ methods, an observation that aligns with the results of our previous study[60]. However, rotations computed at T_2 ratios below 0.5 are slow to converge to the reference value for both the PNO++ and the combined PNO++ methods. Rotations computed

using the length representation of the electric dipole moment operator show faster convergence to the reference value, results which are specific to this set of molecules and are not reproduced for any other system studied (cf. Supporting Information). While we do not observe full convergence of the linear response properties to the ref-



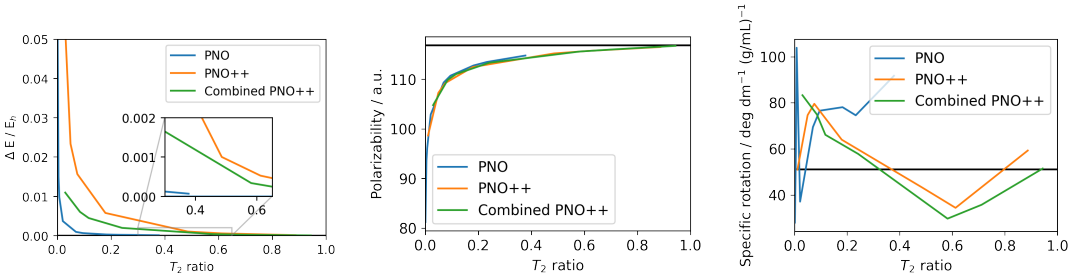
(a) (b) (c)

Figure 4. CCSD specific rotations in deg dm^{-1} (g/mL)$^{-1}$ at 589 nm for (a) (*M*)-1-fluoropropane, (b) (*M*)-1-fluoropentane and (c) (*M*)-1-fluoroheptane systems, computed using the aug-cc-pVDZ basis set.

erence CCSD value at T_2 ratios below 0.5, the results indicate that the PNO++ and combined PNO++ methods still perform significantly better in terms of reducing the cost of property calculations than the PNO method for this set of linear systems.

3.1.2. α- and β-pinene

Figures 5 and 6 depict the effect of wave function truncation with the PNO, PNO++ and combined PNO++ methods on correlation energies and linear response properties computed for the systems of, respectively, (1*R*,5*R*)-α-pinene and (1*R*,5*R*)-β-pinene (Fig. 1). Errors in correlation energies [Fig. 5(a) and 6(a)] follow the same trend as seen for the fluoroalkane systems earlier, with the PNO method having much smaller errors in comparison with the PNO++ at large truncations, and the combined method recovering similar accuracy at the same truncations.

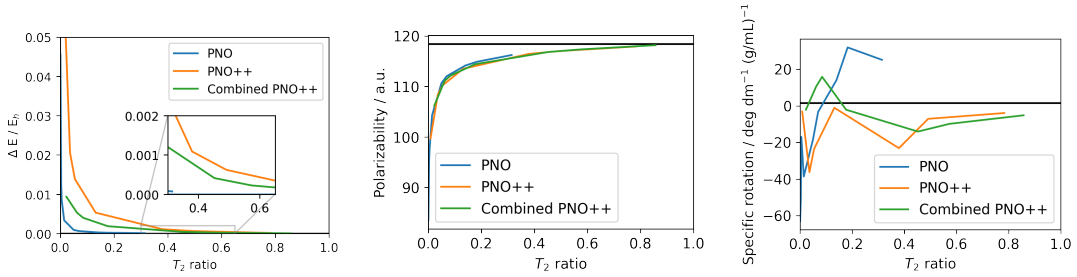


(a) (b) (c)

Figure 5. (a) Truncation errors in CCSD correlation energy in Hartree, (b) CCSD dynamic polarizabilities at 589 nm in a.u., and (c) CCSD specific rotations in deg dm^{-1} (g/mL)$^{-1}$ at 589 nm for (1*R*,5*R*)-α-pinene, computed using the aug-cc-pVDZ basis set.

Interestingly, dipole polarizabilities for both pinene systems [panel (b) in each figure] calculated using the PNO, PNO++ and combined methods have similar errors at similar T_2 ratios. This is in contrast to the results seen with the fluoroalkanes and indicates that all three spaces show similar levels of compactness for the polarizability

for these bicyclic molecules.



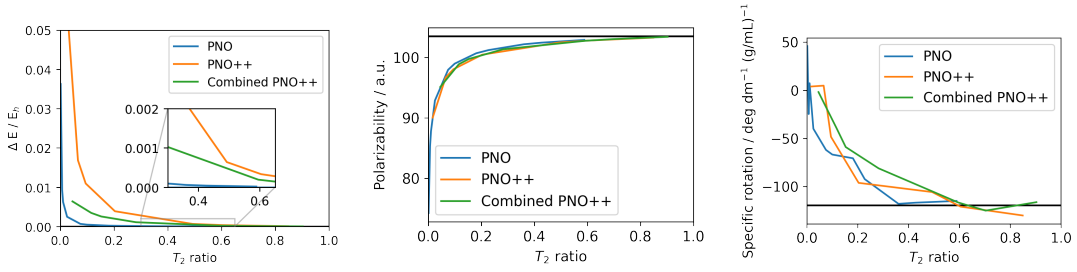
(a) (b) (c)

Figure 6. (a) Truncation errors in CCSD correlation energy in Hartree, (b) CCSD dynamic polarizabilities at 589 nm in a.u., and (c) CCSD specific rotations in deg dm^{-1} (g/mL)$^{-1}$ at 589 nm for (1*R*,5*R*)-*β*-pinene, computed using the aug-cc-pVDZ basis set.

Figures 5(c) and 6(c) contain specific rotations as a function of the T_2 ratio. The pinene systems differ only by the placement of the double bond, but their specific rotations calculated using CCSD linear response differ significantly, with the modified velocity gauge value for *α*-pinene being 51.19 deg dm^{-1} (g/mL)$^{-1}$, and the same for *β*-pinene being 1.72 deg dm^{-1} (g/mL)$^{-1}$. We see that for *α*-pinene, all three methods show oscillations around the reference value, with the error being fairly large up to the T_2 ratio of 1.0, indicating that even small truncations of the space introduce a large error (of up to 33% at a T_2 ratio of 0.6) in the specific rotation value. The small value of the specific rotation of *β*-pinene introduces additional problems, as a negative truncation error at small T_2 ratios means that the sign of the specific rotation for the PNO and PNO++ methods is incorrect at all T_2 ratios considered. We see this in Figure 6(c). For both pinene systems, while the error using the PNO method is larger than that using the PNO++ method at several T_2 ratios, neither the PNO++ nor the combined method can be used to reliably maintain accuracy in the specific rotation value. This observation is consistent with our earlier findings[47] regarding the exquisite sensitivity of chiroptical properties such as optical rotations with respect to the quality of the virtual space, such that even a small error in the Rosenfeld tensor translates to 10^7 times the error in the rotation value as compared to errors in the polarizability tensor. We have confirmed that even tighter PNO cutoffs do yield the canonical result, but, even for the PNO++ method, individual virtual orbitals can still make a non-negligible contribution to the total rotation.

3.1.3. (S)-1-phenylethanol

Figure 7 plots the truncation errors in correlation energy as well as the dynamic polarizabilities and specific rotations for all three methods for (*S*)-1-phenylethanol, a challenging system for reduced scaling methods exploiting spatial locality due to the presence of the aromatic ring substituent. The PNO method, along with exhibiting lower errors in correlation energy relative to the PNO++ and combined methods, also shows similar errors for the polarizability (within 2 a.u.) to the PNO++ and combined methods. The convergence behavior of the specific rotation is better for the PNO method for this system than the PNO++ method, with a point at a T_2 ratio of 0.36 falling within 3 deg dm^{-1} (g/mL)$^{-1}$ of the reference value. In comparison, the PNO++ and combined methods require T_2 ratios of between 0.6-0.7 to achieve



(a) (b) (c)

Figure 7. (a) Truncation errors in CCSD correlation energy in Hartree, (b) CCSD dynamic polarizabilities at 589 nm in a.u., and (c) CCSD specific rotations in deg dm^{-1} $(g/mL)^{-1}$ at 589 nm for (*S*)-1-phenylethanol, computed using the aug-cc-pVDZ basis set.

similarly close values. The system presents a challenge to the local correlation methods tested here, with very large portions of the space required in order to minimize the error in specific rotation value.

3.2. Product Densities

3.2.1. Correlation Energies

Figure 8 contrasts the PNO++ method explored in Ref 60 with the a similar method employing the product density described in Section 2.2 for four small test systems: two hydrogen molecule helices, H_2O_2 and 1,3-dimethylallene (DMA) (see Ref60 for geometry details). At the same T_2 ratio, the truncation error in the correlation energy is higher for the product-based density than for the regular PNO++ density using only the electric dipole moment operator. The product-based density requires T_2 ratios greater than 0.5 in order to maintain reasonable accuracy in the correlation energy. Since the PNO++ method itself shows higher truncation errors than the PNO method for correlation energies at small T_2 ratios, this indicates that the product-based density is not a good method for obtaining correlation energies at truncated system sizes.

3.2.2. Optical Rotation

Figure 9 shows optical rotations computed using the PNO++ method with regular and product densities. The product density is created using the length gauge representation of the electric dipole moment operator, and thus, for mixed electric/magnetic perturbations, the elements of the density do potentially vary with the choice of coordinate origin. However, for relatively small molecules with the origin placed at the center of mass, this is expected to have little impact on the final specific rotations, which utilize the modified velocity gauge, as described above. While convergence is seen for both methods, it can be seen from the figures that the product-based density offers no improvement to the convergence behavior shown by the PNO++ method for the four test systems. At T_2 ratios below 0.25, we see large magnitude errors in the rotation value for the product density space, with the specific rotation for DMA at very large truncation being the wrong sign. Thus the space created by the product density is not more optimized for the mixed property of rotation than the density containing only a single external perturbation operator. One possible reason for the

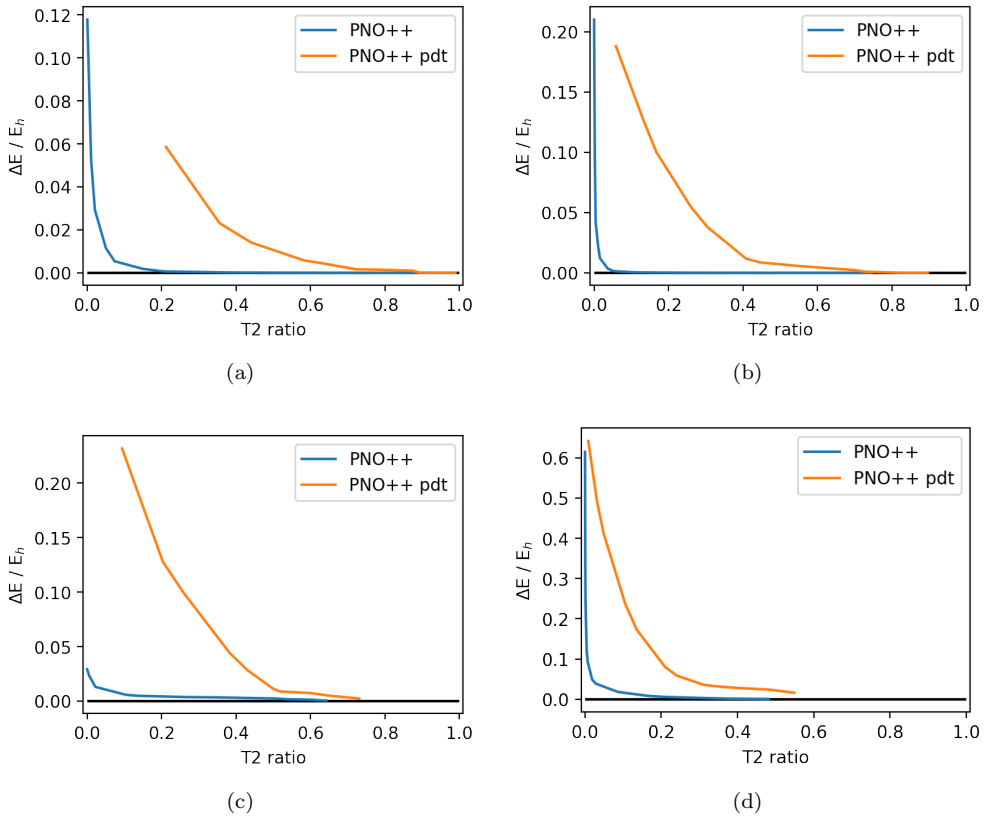


Figure 8. Truncation errors in CCSD correlation energy in Hartree for (a) $(H_2)_4$, (b) $(H_2)_7$, (c) H_2O_2 and (d) DMA systems, computed using the aug-cc-pVDZ basis set for the PNO++ methods with the regular perturbed (PNO++, blue) and product (PNO++ pdt, orange) densities.

poor behavior of the product density versus the density using only the electric dipole moment operator may be the slower convergence of the magnetic dipole operator with respect to basis set completeness (cf. the Supporting Information of Ref 60).

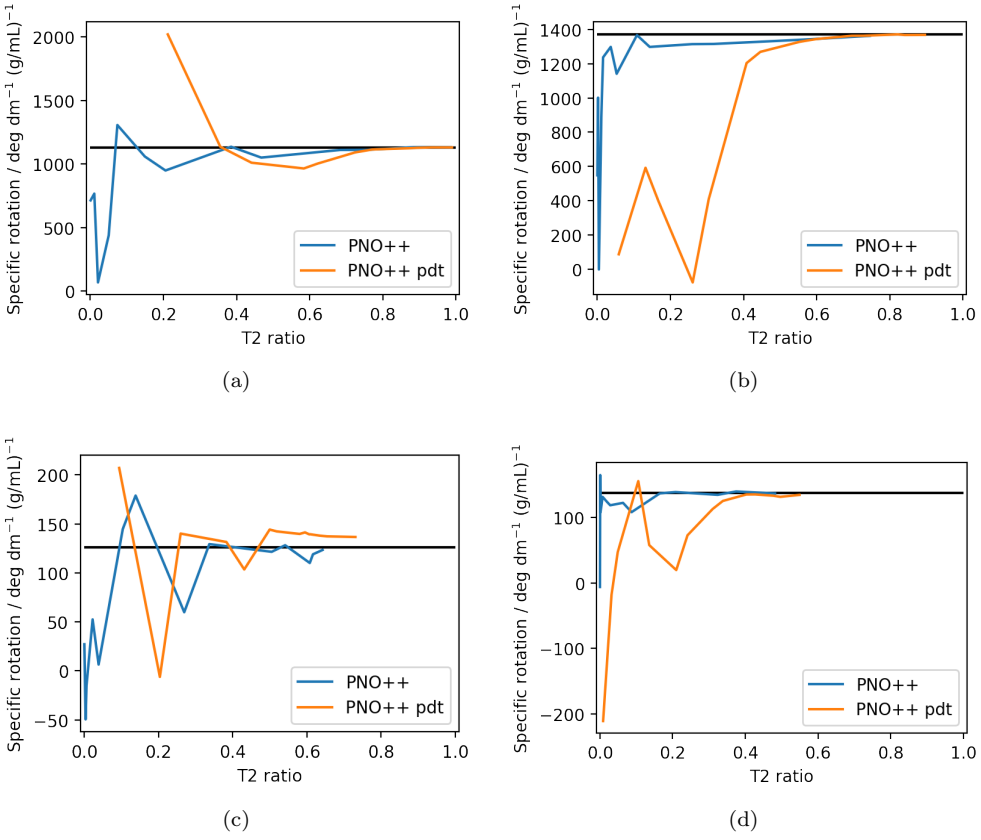


Figure 9. CCSD specific rotations in deg dm^{-1} (g/mL)$^{-1}$ at 589 nm for (a) $(H_2)_4$, (b) $(H_2)_7$, (c) H_2O_2 and (d) DMA systems, computed using the aug-cc-pVDZ basis set for the PNO++ methods with the regular perturbed (PNO++, blue) and product (PNO++ pdt, orange) densities.

3.3. Weak Pairs

Table 1 contains T_2 ratios and errors in correlation energies for the pair energy criterion ϵ as well as the perturbation-including criterion $\bar{\mu}$ at a given weak pair threshold $T_{cutPairs}$ for the 1-fluoropropane, -pentane, and -heptane systems. Clearly pair selection thresholds above 10^{-6} do not yield useful results, but we include them here to demonstrate the rate of convergence with respect to the cutoff parameters. As the system size increases, we see the T_2 ratio becoming uniformly smaller at a given threshold for the ϵ criterion indicating that fewer pairs have contributions to the energy larger in magnitude than the chosen cutoff. This is also true for the $\bar{\mu}$ criterion and aligns with the discussion in Refs 46 and 73 that on increasing molecular sizes the computational efficiency of local correlation methods improves.

For both of the considered criteria, errors in the correlation energy require T_2 ratios above 0.83 in order to remain below chemical accuracy or 1.6 mE_h, thus keeping a significant portion of the space. However, the errors in correlation energy seen using the $\bar{\mu}$ criterion are larger at a given T_2 ratio than the errors using the ϵ criterion. As

an example, for the very similar T_2 ratios of 0.805 and 0.795 for 1-fluoropentane, the error in correlation energy was less than 1 mE_h for the pair energy criterion, while it was above 13 mE_h for the perturbation-including criterion. The pair energy criterion thus provides a better way to identify weak pairs specifically for maintaining accuracy in the correlation energy.

Molecule	Threshold	ϵ_{ij}		$\bar{\mu}_{ij}$	
		T_2 ratio	ΔE	T_2 ratio	ΔE
1-fluoropropane	10^{-3}	0.211	51.1838	0.457	136.677
	10^{-4}	0.612	2.75454	0.827	2.64050
	10^{-5}	0.709	1.57325	0.958	0.22850
	10^{-6}	0.882	0.11981	1.000	0.00000
1-fluoropentane	10^{-3}	0.146	83.3302	0.413	70.5225
	10^{-4}	0.456	6.20241	0.795	13.6891
	10^{-5}	0.568	3.20850	0.955	0.05368
	10^{-6}	0.805	0.15424	0.987	0.00912
1-fluoroheptane	10^{-3}	0.111	112.815	0.328	87.2679
	10^{-4}	0.361	9.77172	0.722	16.5131
	10^{-5}	0.469	4.85166	0.935	2.22753
	10^{-6}	0.726	0.18945	0.991	0.01013

Table 1. T_2 ratios and errors in correlation energy (mE_h) computed at the CCSD level for the three 1-fluoroalkane systems as a function of the $T_{cutPairs}$ threshold, using the aug-cc-pVDZ basis set.

Table 2 presents polarizabilities and specific rotations computed using the modified velocity gauge as a function of the threshold for both criteria for the 1-fluoropropane, -pentane, and -heptane systems. For all three fluoroalkanes and all values of the threshold the polarizability value remains within 1 a.u. of the reference, even at the smallest T_2 ratios. Thus the polarizability is not strongly affected by the neglect of weak pairs using either criterion. The specific rotation, on the other hand, is affected by the truncation and is seen to require a T_2 ratio between 0.5-0.7 for the ϵ criterion and 0.4-0.8 for the $\bar{\mu}$ criterion to remain within 10% of the reference CCSD value.

4. Conclusions

In this work, we have compared the performance of the PNO++ and combined PNO++ methods to the conventional PNO method for larger molecules than considered in previous work using a new implementation within the Psi4 package. We have also tested a new product-based density in the creation of the PNO++ space, as well as the accuracy limits of both an energy- and perturbation-based weak-pair threshold.

For the series of 1-fluoroalkane systems tested, we see better convergence behavior and lower truncation errors for the PNO++ and combined PNO++ methods than the PNO for the linear response properties tested. While the truncation errors in correlation energy were large for the PNO++ method, we recover similar accuracy to the PNO method by incorporating a number of the original pair natural orbitals in the combined PNO++ space. The bicyclic α- and β-pinene molecules appear to be more difficult test cases than the linear alkane chains for all three methods tested,

Molecule	Threshold	ϵ_{ij}			$\bar{\mu}_{ij}$		
		T_2 ratio	α	MVG	T_2 ratio	α	MVG
1-fluoropropane	10^{-3}	0.211	42.02	-83.06	0.457	42.69	-58.10
	10^{-4}	0.612	41.69	-77.05	0.827	41.67	-60.30
	10^{-5}	0.709	41.66	-64.56	0.958	41.66	-57.66
	10^{-6}	0.882	41.65	-63.12	1.000	41.65	-62.89
1-fluoropentane	10^{-3}	0.146	67.26	-77.99	0.413	67.61	-71.85
	10^{-4}	0.456	66.64	-65.14	0.795	66.95	-57.39
	10^{-5}	0.568	66.77	-57.89	0.955	66.77	-54.78
	10^{-6}	0.805	66.77	-54.96	0.987	66.77	-54.11
1-fluoroheptane	10^{-3}	0.111	92.73	-62.17	0.328	92.99	-50.76
	10^{-4}	0.361	91.89	-53.42	0.722	92.47	-43.19
	10^{-5}	0.469	92.21	-47.57	0.935	92.30	-42.51
	10^{-6}	0.726	92.30	-44.82	0.991	92.30	-44.21

Table 2. T_2 ratios, dynamic polarizabilities (a.u.), and specific rotations (deg dm^{-1} $(\mathrm{g/mL})^{-1}$) at 589 nm computed at the CCSD level for the three 1-fluoroalkane systems as a function of the $T_{cutPairs}$ threshold, using the aug-cc-pVDZ basis set. α_{Ref} 1-fluoropropane: 41.65, 1-fluoropentane: 66.77, 1-fluoroheptane: 92.30. $\mathrm{MVG}_{\mathrm{Ref}}$ 1-fluoropropane: -63.03, 1-fluoropentane: -53.72, 1-fluoroheptane: -43.75.

requiring large amounts of the virtual space to be retained in order to obtain accuracy in the dynamic polarizability, and showing large errors in the specific rotation value even at small truncations. A surprising result is the very similar convergence behavior of the polarizability for all three methods, with all three having approximately the same error at a given truncation. Finally, truncation errors in both the correlation energy and the specific rotation are smaller for the PNO method than the PNO++ or combined methods for the system of (*S*)-1-phenylethanol.

The use of a product perturbed density, in which mixed perturbations are incorporated simultaneously, was tested for its ability to produce a more compact space for specific rotations, having combined two perturbed densities via a Hadamard product. However, the results suggest that the density using a single electric-field perturbation is more effective than the product density and that further testing — particularly for the basis-set completeness of the magnetic dipole operator — is required in order to understand the impact of using a mixed density in order to create the PNO++ space.

The two weak pair criteria examined here underscore several observations from the larger benchmark calculations with the PNO, PNO++, and combined PNO++ methods. First, T_2 ratios are larger at a given threshold for the $\bar{\mu}$ than for the ϵ criterion, suggesting that this criterion naturally contains larger magnitude contributions than the pair energy criterion. Second, the ϵ criterion provides a better way to identify weak pairs specifically for maintaining accuracy in the correlation energy, while the polarizability is not affected by aggressive truncation, and thus either criterion could be used to truncate the space. For the specific rotation, the ability to truncate at a lower threshold for a given criterion is system-dependent, with the linear fluoroalkane chains requiring slightly less of the space with the $\bar{\mu}$ than with the ϵ criterion. This mirrors observations of the truncation error with the PNO method versus the PNO++ method for the same systems.

Through this study we have studied the effect of truncation on CCSD correlation energies, dynamic polarizabilities, and specific rotations using three local correlation methods, the PNO, PNO++ and combined PNO++ methods. While the results for

the linear alkane chains are encouraging in their better convergence behavior, it is clear that the PNO++ method is not yet a panacea for the local-correlation computation of linear-response properties. The reasons for the inconsistent behavior of all of the PNO-based methods — from conventional PNO, to the various flavors of the PNO++ method considered here – require further investigation.

5. Acknowledgements

This research was supported by the U.S. National Science Foundation (grant CHE-1900420). The authors are grateful to Advanced Research Computing at Virginia Tech for providing computational resources and technical support that have contributed to the results reported within the paper.

References

[1] T.D. Crawford and H.F. Schaefer, in *Rev. Comput. Chem.*, edited by Donald B. Boyd Kenny B. Lipkowitz, Chap. 2 (John Wiley and Sons, Inc., New York, 2000), pp. 33–136.
[2] R.J. Bartlett and M. Musiał, Rev. Mod. Phys. **79** (1), 291–352 (2007).
[3] T. Helgaker, W. Klopper and D.P. Tew, Mol. Phys. **106** (16-18), 2107–2143 (2008).
[4] M.S. Gordon, D.G. Fedorov, S.R. Pruitt and L.V. Slipchenko, Chem. Rev. **112** (1), 632–672 (2012).
[5] E. Epifanovsky, D. Zuev, X. Feng, K. Khistyaev, Y. Shao and A.I. Krylov, J. Chem. Phys. **139** (13) (2013).
[6] W. Li and S. Li, J. Chem. Phys. **121** (14), 6649–6657 (2004).
[7] W. Li, P. Piecuch, J.R. Gour and S. Li, J. Chem. Phys. **131** (11) (2009).
[8] T. Kinoshita, O. Hino and R.J. Bartlett, J. Chem. Phys. **119** (15), 7756–7762 (2003).
[9] H. Koch, A. Sánchez De Merás and T.B. Pedersen, J. Chem. Phys. **118** (21), 9481–9484 (2003).
[10] E.G. Hohenstein, R.M. Parrish and T.J. Martínez, J. Chem. Phys. **137** (4), 1085 (2012).
[11] R. Schutski, J. Zhao, T.M. Henderson and G.E. Scuseria, J. Chem. Phys. **147** (18) (2017).
[12] R.M. Parrish, Y. Zhao, E.G. Hohenstein and T.J. Martínez, J. Chem. Phys. **150** (16) (2019).
[13] F. Pawłowski, J. Olsen and P. Jørgensen, J. Chem. Phys. **150** (13) (2019).
[14] P. Pulay, Chem. Phys. Lett. **100** (2), 151–154 (1983).
[15] C. Hampel and H.J. Werner, J. Chem. Phys. **104** (16), 6286–6297 (1996).
[16] F. Neese, F. Wennmohs and A. Hansen, J. Chem. Phys. **130** (136), 114108–134101 (2009).
[17] A.G. Taube and R.J. Bartlett, J. Chem. Phys. **128** (16), 164101 (2008).
[18] M. Schütz and H.J. Werner, J. Chem. Phys. **114** (2), 661–681 (2001).
[19] J. Yang, G.K.L. Chan, F.R. Manby, M. Schütz and H.J. Werner, J. Chem. Phys. **136** (14), 144105 (2012).
[20] S. Saebø and P. Pulay, Annu. Rev. Phys. Chem **44**, 213–36 (1993).
[21] W. Meyer, J. Chem. Phys. **58** (3), 1017–1035 (1973).
[22] C. Edmiston and M. Krauss, J. Chem. Phys. **45** (5), 1833–1839 (1966).
[23] C. Edmiston and M. Krauss, J. Chem. Phys. **49** (1), 192–205 (1968).
[24] R. Ahlrichs and F. Driessler, Theor. Chim. Acta **36** (4), 275–287 (1975).
[25] F. Neese, A. Hansen and D.G. Liakos, J. Chem. Phys. **131** (132), 64103–114108 (2009).
[26] T. Crawford and R.A. King, Chem. Phys. Lett. **366** (5-6), 611–622 (2002).
[27] T. Korona and H.J. Werner, J. Chem. Phys. **118** (7), 3006–3019 (2003).
[28] B. Helmich and C. Hättig, J. Chem. Phys. **135** (21) (2011).
[29] A.K. Dutta, F. Neese and R. Izsák, J. Chem. Phys. **145** (3) (2016).

[30] C. Peng, M.C. Clement and E.F. Valeev, J. Chem. Theory Comput. **14** (11), 5597–5607 (2018).
[31] M.S. Frank and C. Hättig, J. Chem. Phys. **148** (13) (2018).
[32] R.H. Myhre and H. Koch, J. Chem. Phys. **145** (4) (2016).
[33] M. Saitow and F. Neese, J. Chem. Phys. **149** (3), 034104 (2018).
[34] J. Brabec, J. Lang, M. Saitow, J. Pittner, F. Neese and O. Demel, J. Chem. Theory Comput. **14** (3), 1370–1382 (2018).
[35] Q. Ma and H.J. Werner, J. Chem. Theory Comput. **16** (5), 3135–3151 (2020).
[36] Q. Ma and H.J. Werner, J. Chem. Theory Comput. **17** (2), 902–926 (2021).
[37] D. Kats and H.J. Werner, J. Chem. Phys. **150** (21), 214107 (2019).
[38] J. Gauss and H.J. Werner, Phys. Chem. Chem. Phys. **2** (10), 2083–2090 (2000).
[39] T. Korona, K. Pflüger and H.J. Werner, Phys. Chem. Chem. Phys **6**, 2059–2065 (2004).
[40] D. Kats, T. Korona and M. Schütz, J. Chem. Phys. **127** (6) (2007).
[41] K. Ledermüller, D. Kats and M. Schütz, J. Chem. Phys. **139** (8) (2013).
[42] M. Schütz, J. Chem. Phys. **142** (21) (2015).
[43] G.L. Stoychev, A.A. Auer, J. Gauss and F. Neese, J. Chem. Phys. **154** (16), 164110 (2021).
[44] T.D. Crawford, A. Kumar, A.P. Bazanté and R. Di Remigio, Wiley Interdiscip. Rev. Comput. Mol. Sci. **9** (4), 1–25 (2019).
[45] N.J. Russ and T.D. Crawford, Chem. Phys. Lett. **400** (1-3), 104–111 (2004).
[46] N.J. Russ and T.D. Crawford, Phys. Chem. Chem. Phys. **10** (23), 3345 (2008).
[47] H.R. McAlexander and T.D. Crawford, J. Chem. Theory Comput. **12** (1), 209–222 (2016).
[48] A. Kumar and T.D. Crawford, J. Phys. Chem. A **121** (3), 708–716 (2017).
[49] I.M. Høyvik, R.H. Myhre and H. Koch, J. Chem. Phys. **146** (14) (2017).
[50] P. Baudin and K. Kristensen, J. Chem. Phys. **144** (22) (2016).
[51] P. Baudin and K. Kristensen, J. Chem. Phys. **146** (21) (2017).
[52] S. Höfener and W. Klopper, Chem. Phys. Lett. **679**, 52–59 (2017).
[53] D. Mester, P.R. Nagy and M. Kállay, J. Chem. Phys. **146** (19) (2017).
[54] D. Mester, P.R. Nagy and M. Kállay, J. Chem. Phys. **148** (9) (2019).
[55] A. Landau, K. Khistyaev, S. Dolgikh and A.I. Krylov, J. Chem. Phys. **132** (1), 014109 (2010).
[56] A.K. Dutta, M. Saitow, C. Riplinger, F. Neese and R. Izsák, J Chem. Phys. **148** (24), 244101 (2018).
[57] P. Pokhilko, D. Izmodenov and A.I. Krylov, J. Chem. Phys. **152** (3), 034105 (2020).
[58] S. Chamoli, K. Surjuse, B. Jangid, M.K. Nayak and A.K. Dutta, J. Chem. Phys. **156** (20), 204120 (2022).
[59] T.D. Crawford, A. Kumar, A.P. Bazanté and R. Di Remigio, Wiley Interdiscip. Rev. Comput. Mol. Sci. **9** (4), 1–25 (2019).
[60] R. D'Cunha and T.D. Crawford, J. Chem. Theory Comput. **17** (1), 290–301 (2021).
[61] D.G. Smith, L.A. Burns, A.C. Simmonett, R.M. Parrish, M.C. Schieber, R. Galvelis, P. Kraus, H. Kruse, R. Di Remigio, A. Alenaizan, A.M. James, S. Lehtola, J.P. Misiewicz, M. Scheurer, R.A. Shaw, J.B. Schriber, Y. Xie, Z.L. Glick, D.A. Sirianni, J.S. O'Brien, J.M. Waldrop, A. Kumar, E.G. Hohenstein, B.P. Pritchard, B.R. Brooks, H.F. Schaefer, A.Y. Sokolov, K. Patkowski, A.E. Deprince, U. Bozkaya, R.A. King, F.A. Evangelista, J.M. Turney, T.D. Crawford and C.D. Sherrill, J. Chem. Phys. **152** (18) (2020).
[62] A.D. Becke, J. Chem. Phys. **98** (7), 5648–5652 (1993).
[63] C. Lee, W. Yang and R.G. Parr, Phys. Rev. B **37** (2), 785–789 (1988).
[64] P.J. Stephens, F.J. Devlin, C.F. Chabalowski and M.J. Frisch, J. Phys. Chem. **98** (45), 11623–11627 (1994).
[65] T.H. Dunning, J. Chem. Phys. **90** (2), 1007–1023 (1989).
[66] R.A. Kendall, T.H. Dunning and R.J. Harrison, J. Chem. Phys. **96** (9), 6796–6806 (1992).
[67] T.B. Pedersen and H. Koch, J. Chem. Phys. **106** (19), 8059 (1997).
[68] H. Koch and P. Jørgensen, J. Chem. Phys. **93** (5), 3333–3344 (1990).
[69] T.B. Pedersen and H. Koch, J. Chem. Phys. **106** (19), 8059–8072 (1997).

[70] T.B. Pedersen, H. Koch, L. Boman and A.M. Sánchez de Merás, Chem. Phys. Lett. **393** (4-6), 319–326 (2004).
[71] D.E. Woon and T.H. Dunning, J. Chem. Phys. **100** (4), 2975–2988 (1994).
[72] O. Masur, D. Usvyat and M. Schütz, J. Chem. Phys. **139** (16), 164116 (2013).
[73] H.R. McAlexander, T.J. Mach and T.D. Crawford, Phys. Chem. Chem. Phys. **14** (21), 7830–6 (2012).