JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Accuracy of Explicitly Correlated Local PNO-CCSD(T)

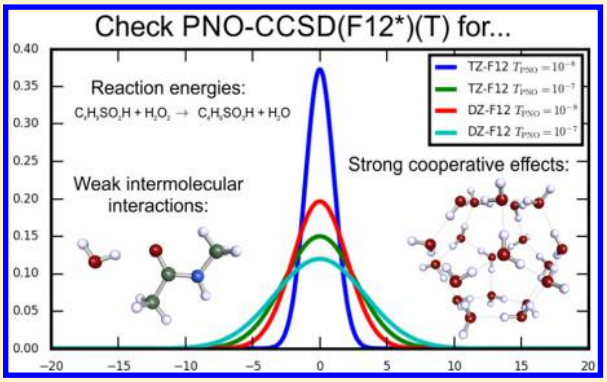
Gunnar Schmitz*,† and Christof Hättig‡

†Department of Chemistry, Aarhus Universitet, Aarhus 8000, Denmark

‡Arbeitsgruppe Quantenchemie, Ruhr-Universität, Bochum 44780, Germany

Supporting Information

ABSTRACT: In recent years PNO-based local correlation methods have gained popularity since they allow Coupled Cluster (CC) calculations with reduced computational costs, yet only a few systematic studies concerning their accuracy are available, in particular for the explicitly correlated versions. In this work we take a deeper look at the explicitly correlated local PNO-CCSD(F12*)(T0) and PNO-CCSD(F12*)(T) methods. The first variant uses the so-called semicanonical triples correction (T0) which neglects off-diagonal elements in the occupied block of the Fock matrix. In PNO-CCSD(F12*)(T) this approximation is avoided by means of Laplace transformation techniques and convergence to the canonical results in the limit of no PNO truncation is restored. We assess the accuracy of both methods using well established benchmark sets for reaction energies and weak molecular interactions and take a look at a system with strong cooperative many-body effects. For reaction energies a close agreement with canonical methods is observed, and chemical accuracy can be reached. Also for weak intermolecular interactions the accuracy is easily controlled, and the methods even allow for improving existing benchmark data.



INTRODUCTION

The goal of ab initio quantum chemistry is to predict energies and properties of molecular systems without relying on empirical parameters to solve approximately the (electronic) Schrödinger equation. For that purpose well-known hierarchies of methods have been developed. The Coupled Cluster (CC) methods are known to be highly accurate. They converge systematically to the Full CI limit by going higher in the excitation level. Unfortunately the CC methods suffer from a steep scaling with the system size N and, as all correlated wave function methods, from a slow basis set convergence. For example, the very accurate CCSD(T) method[1] shows a $O(N^7)$ scaling and requires quadruple- and quintuple-ζ basis sets to fully exploit the accuracy of CCSD(T). This limits this approach to systems with only a few atoms. Over the last decades solutions emerged for both problems. To reduce the scaling one usually exploits the local nature of dynamic electron correlation. The strategies to achieve this can be loosely divided into two categories. In the first category the system is partitioned into fragments for which conventional calculations are carried out. The total correlation energy is then obtained as a sum over the fragment energies.[2−12] In the second category the system is treated as a whole, but electron correlation is restricted to local domains or sparsity is rigorously exploited.[13−19] We follow an ansatz from the second category, in which for each pair of occupied orbitals a private compressed set of virtual orbitals is constructed by means of Pair Natural Orbitals (PNOs)[20−22] to describe the dynamic correlation. The required number of PNOs per pair quickly becomes independent of the system size and allows linear scaling algorithms.

Although PNOs allow low scaling algorithms, still the slow basis set convergence spoils the efficiency. This problem can be solved if contributions with an explicit dependence on the interelectronic distance r_{12} are included in the wave function ansatz. Nowadays the most widely used explicitly correlated methods are the F12 methods,[23−25] where the Slater determinants are augmented with geminals χ_{ij} which have an explicit dependence on r_{12} via a correlation factor $f(r_{12})$. The F12 approach enhances the convergence of the correlation energy typically such that with a basis set of cardinal number X results are obtained that have the same accuracy as a conventional calculation in a basis set with the cardinal number $X + 2$. At the CCSD level several approximations to CCSD-F12[26] have been proposed which are only around 20% more costly than conventional CCSD calculations in the same basis set.[27−29] Among them CCSD(F12*),[29] which is also referred to as CCSD-F12c by some authors, is the most complete approximation to CCSD-F12 which still can be applied in a routine fashion.

For efficient wave function based methods which allow accurate calculations at moderate costs it is mandatory to combine a local correlation treatment using for example — as in the current work — PNOs with F12 theory. Approaches like

Received: February 20, 2017
Published: May 17, 2017

Downloaded via BOCHUM LIBRARIES on January 16, 2019 at 09:30:24 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

© 2017 American Chemical Society

DOI: 10.1021/acs.jctc.7b00180
J. Chem. Theory Comput. 2017, 13, 2623−2633

PNO-CCSD(F12*)(T),[30] DLPNO-CCSD(T),[31,32] or PNO-LMP2-F12[33] are designed to be efficient and at the same time not to degrade the accuracy of the underlying canonical methods. During their development these approaches were typically checked against small test sets, but so far only a few extensive numerical studies like those from Liakos et al.,[34,35] Minenkov et al.,[36] and Friedrich[37] fathomed out the limits of these methods. Such studies are important to get insight into the strengths and weaknesses of PNO-based methods and to establish them as reliable black box tools, especially since PNO methods are already used to deliver benchmark data to evaluate the accuracy of DFT functionals[38] or as a workhorse in several application studies.[39−41]

In the current work we test a recently developed PNO-CCSD(F12*)(T) implementation[30] against well-known test sets for reaction energies and weak molecular interactions. This allows a more complete view of the strengths and weaknesses of PNO-CCSD(T) and helps to establish explicitly correlated PNO methods as a tool to study reaction and binding energies in larger systems. In the following we give first a brief description of the most important theoretical aspects. Thereafter, the accuracy of PNO-CCSD(F12*)(T) is assessed for reaction energies, followed by a benchmark against the S66 test set for weak intermolecular interactions and a study of interaction energies with strong cooperative many-body effects.

THEORY

The theoretical background of PNO-CCSD(F12*)(T) and related methods has been described in detail elsewhere.[30,42] Therefore, we describe here only the most important aspects and refer for more detailed presentations to the cited references. We start with a description of PNOs and F12-PNOs and outline the most important points about the triples correction (T0) and the full (T) correction within a PNO framework.

PNOs. PNOs are linear combinations of virtual orbitals $|a\rangle$ which are specific for a pair of occupied orbitals ij

$$|\bar{a}\rangle = \sum_a |a\rangle d^{ij}_{a\bar{a}} \tag{1}$$

They are chosen such that they diagonalize the contribution D^{ij}_{ab} of the pair ij to the virtual/virtual block of a MP2-like one-electron density matrix:

$$\sum_{ab} d^{ij}_{a\bar{a}} D^{ij}_{ab} d^{ij}_{b\bar{b}} = n^{ij}_{\bar{a}} \delta_{\bar{a}\bar{b}} \tag{2}$$

The eigenvalues $n^{ij}_{\bar{a}}$ of D^{ij}_{ab} are the occupation numbers of the PNOs. If sorted in descending order they decrease rapidly to zero. The fast decay allows for the truncation of the PNO expansion to PNOs with occupation numbers above a threshold T_{PNO} without sacrificing accuracy. For an efficient implementation it can be exploited that the decay is faster if localized occupied orbitals are used. In the current work we use Foster-Boys localization.[43] Deviating from this simple scheme the MP2-like one-electron density is usually built in an already truncated doubles space in order to to avoid the $O(N^5)$ scaling costs for a MP2 calculation.[42,44−47] This truncation is coupled to the threshold T_{PNO} to ensure that it does not diminish the overall accuracy.

F12-PNOs. In explicitly correlated calculations we exploit that parts of the virtuals can be described by the geminals to reduce further the number of required PNOs per pair. To account for this the PNOs for the virtual space in F12 calculations, F12-PNOs, are obtained from densities built from the differences[48]

$$\Delta^{ij}_{ab} = t^{ij}_{ab} - r^{\tilde{ij}}_{ab} \tag{3}$$

between amplitudes t^{ij}_{ab} for double excitations into products of virtual MOs and the part $r^{\tilde{ij}}_{ab}$ that can be described by the geminal χ_{ij}. Besides the F12-PNOs additional auxiliary PNOs (X-PNOs) have to be introduced in order to arrive at a pair-specific representation of the F12 strong orthogonality operator.[42] These X-PNOs are also truncated according to thresholds, which are automatically adjusted[48] to the value of T_{PNO}. For further details on the PNO construction for the explicitly correlated CCSD(F12*) variants we refer to refs 30 and 42.

Semicanonical Triples Correction. In the spirit of the PNO framework the triples correction is obtained in a triple specific TNO basis, which was introduced by Riplinger et al.[31] TNOs are constructed by diagonalizing the averaged pair densities of the contributing pairs. As for PNOs, also the TNO expansion is truncated using a threshold T_{TNO}. In ref 30 it was shown that it makes no sense to set T_{TNO} tighter than T_{PNO}. Therefore, we have set for the current work the TNO threshold always equal to the PNO threshold.

In a noncanonical orbitals basis the solution of the (T) triples equations would require the storage of a full set of triples amplitudes, which would severely limit the applicability of such an approach. To avoid this storage and I/O bottleneck, the (T0) approximation[18] neglects the off-diagonal occupied/occupied elements of the Fock matrix. The TNOs are transformed to a semicanonical basis where the virtual/virtual block of the Fock matrix is diagonal within the TNOs for the given LMO triple ijk. The equations for (T0) triples amplitudes assume then a form which is analogous to the canonical (T) expression

$$t^{ijk}_{\bar{a}\bar{b}\bar{c}} = (\varepsilon^{\bar{a}\bar{b}\bar{c}}_{ijk})^{-1} \hat{P}^{ijk}_{\bar{a}\bar{b}\bar{c}} \left(\sum_d g^{i\bar{d}}_{\bar{a}\bar{b}} t^{jk}_{\bar{d}\bar{c}} - \sum_l g^{ij}_{\bar{a}l} t^{lk}_{\bar{b}\bar{c}} \right) \tag{4}$$

where $(\varepsilon^{\bar{a}\bar{b}\bar{c}}_{ijk})^{-1}$ is the orbital energy denominator, $t^{ij}_{\bar{a}\bar{b}}$ the converged PNO-CCSD doubles amplitudes, and $g^{rs}_{pq} = (pr|qs)$ are two electron integrals. In many cases the T0 approximation works well, but a drawback is that even in the limit of no TNO and PNO truncation the canonical results cannot be recovered which spoils the accuracy of the underlying CCSD(T) method.

Laplace Triples Correction. An alternative way to avoid the storage and I/O bottleneck for the solution of coupled triples equations which is able to converge to the exact triples amplitudes is based on Laplace transformation techniques.[49] The energy denominator in the amplitude equation is replaced by an equivalent Laplace transformed expression which is evaluated numerically on a grid with n_L grid points t_z and weights ω_z:

$$t^{ijk}_{\bar{a}\bar{b}\bar{c}} = \hat{P}^{ijk}_{\bar{a}\bar{b}\bar{c}} \sum_{z=1}^{n_L} \omega_z \left(\sum_{\bar{d}} e^{(\varepsilon_j+\varepsilon_k-\varepsilon_c)t_z} t^{jk}_{\bar{d}\bar{c}} g^{i\bar{d}}_{\bar{a}\bar{b}} e^{(\varepsilon_i-\varepsilon_a-\varepsilon_b)t_z} - \sum_l e^{(\varepsilon_k-\varepsilon_b-\varepsilon_c)t_z} t^{lk}_{\bar{b}\bar{c}} g^{ij}_{\bar{a}l} e^{(\varepsilon_i+\varepsilon_j-\varepsilon_a)t_z} \right) \tag{5}$$

The exponential factors can be absorbed into the integrals which leads to grid point-dependent integrals. The resulting

DOI: 10.1021/acs.jctc.7b00180
J. Chem. Theory Comput. 2017, 13, 2623−2633

Laplace-transformed expression can be transformed from the canonical to any other MO basis without introducing a coupling between the amplitudes for different triples *ijk*. The computational price that has to be paid is the computational costs for the grid point-dependent integrals, but the systematic convergence to the canonical CCSD(T) limit outweighs this slight disadvantage. As shown in ref 30 in general already 3−4 grid points are sufficient to achieve converged energy differences.

Scaled Triples Correction. As in canonical implementations the explicitly correlated CCSD variants are combined with the conventional triples correction.[50] This causes sometimes the problem that convergence of the triples correction limits the overall basis set convergence of CCSD(F12*)(T). The (T*) correction proposed by Knizia[51] repairs this to some extent by scaling the correction with the ratio between the MP2-F12 and MP2 correlation energies. We follow for the PNO-CCSD(F12*)(T0) and PNO-CCSD(F12*)(T) calculations the same strategy and exploit that the PNO-MP2-F12 energy can be separated in an explicitly correlated and a conventional part

$$E^{\mathrm{corr}}_{\mathrm{PNO\text{-}MP2\text{-}F12}} = E^{\mathrm{corr}}_{\mathrm{PNO\text{-}MP2}} + \Delta E^{\mathrm{corr}}_{\mathrm{F12}} \tag{6}$$

where $E^{\mathrm{corr}}_{\mathrm{PNO\text{-}MP2}}$ is the MP2 energy evaluated with t^{ij}_{ab} amplitudes from the PNO-MP2-F12 calculation. For the (T*) variant the triples correction is then scaled by the factor

$$f_{\mathrm{T}*} = 1 + \frac{\Delta E^{\mathrm{corr}}_{\mathrm{F12}}}{E^{\mathrm{corr}}_{\mathrm{PNO\text{-}MP2}}} \tag{7}$$

In contrast to the canonical case the PNO-MP2 energy computed with amplitudes $\tilde{t}^{ij}_{ab}$ from PNO-MP2-F12 is slightly smaller than in a PNO-MP2 calculation without F12. The reason is that the virtual space is spanned by the more compact F12-PNOs. Since the F12-PNOs are also used for the (T) correction, this can partially correct for the slightly larger PNO-truncation error in the (T) correction with F12-PNOs. The correction can also be applied to the (T0) correction without special consideration.

COMPUTATIONAL DETAILS

The PNO-based calculations were carried out with a local development version of the TURBOMOLE program package. The details of the PNO implementation are described in refs 30, 42, 46, 48, and 52. For the canonical reference calculations the ccsdf12 module[29,53] from TURBOMOLE V6.6 was used. The HF calculation were done with the dscf[54] program. We employed the aug-cc-pVXZ basis set family,[55] which will be abbreviated in the following with aXZ, and the cc-pVXZ-F12 basis sets[56] for which we use the short hand notation XZ-F12. For some applications a combination of basis sets was used, where a cc-pVXZ basis was assigned to hydrogen and the aug-cc-pVXZ basis to all other atoms. We abbreviate this basis set as aXZ'. For the density-fitting and RI approximations we used the auxiliary basis sets from refs 57−59. The frozen core approximation was used throughout this work, and for the Laplace triples correction it was ensured that the error in the square root of the Laplace error function is below 10^{-2}, which resulted in typically 3−4 Laplace points.

RESULTS AND DISCUSSION

Reaction Energies. First, we study the performance of PNO-CCSD(F12*)(T) and related methods for reaction energies. For that purpose we use the test set of Friedrich and Hänchen which was introduced in ref 60 and later updated in ref 37 by a new scheme to calculate the reaction energies. The authors used a composite scheme

$$E_{\mathrm{ref}} = E^{\mathrm{QZ\text{-}F12}}_{\mathrm{HF+CABS}} + E^{\mathrm{QZ\text{-}F12}}_{\mathrm{MP2\text{-}F12}} + \Delta E^{\mathrm{TZ\text{-}F12}}_{\mathrm{CCSD(F12*)}} + E^{\mathrm{CBS(23)}}_{\mathrm{(T)}} \tag{8}$$

where $E^{\mathrm{QZ\text{-}F12}}_{\mathrm{HF+CABS}}$ is the HF energy including a CABS singles correction[51] for the remaining basis set error at the HF level computed in the QZ-F12 basis, $E^{\mathrm{QZ\text{-}F12}}_{\mathrm{MP2\text{-}F12}}$ is the MP2-F12/QZ-F12 correlation energy, $\Delta E^{\mathrm{TZ\text{-}F12}}_{\mathrm{CCSD(F12*)}}$ is a higher-order correction in the TZ-F12 basis defined as

$$\Delta E_{\mathrm{CCSD(F12*)}} = E_{\mathrm{CCSD(F12*)}} - E_{\mathrm{MP2\text{-}F12}} \tag{9}$$

and $E^{\mathrm{CBS(23)}}_{\mathrm{(T)}}$ is a Schwenke type extrapolation for the triples correction

$$E^{\mathrm{CBS(23)}}_{\mathrm{(T)}} = E^{\mathrm{DZ\text{-}F12}}_{\mathrm{(T)}} + F\cdot(E^{\mathrm{TZ\text{-}F12}}_{\mathrm{(T)}} - E^{\mathrm{DZ\text{-}F12}}_{\mathrm{(T)}}) \tag{10}$$

with $F = 1.431442$ to avoid that the triples correction limits the accuracy.

Our PNO based calculations are compared to this reference. A detailed list of the involved reactions and the energies can be found in the Supporting Information.

To analyze the accuracy we calculated the mean absolute error $|\bar{\Delta}|$ and the root-mean-square deviation Δ_{RMSD}. The results are listed in Table 1, where for comparison also the CCSD(F12*)(T)/TZ-F12 values from ref 37 are repeated. Additionally the errors are shown in Figure 1 as normal distributions. The errors decrease systematically if the PNO threshold is tightened. From the data one can conclude that PNO-CCSD(F12*)(T*) with $T_{\mathrm{PNO}} = 10^{-8}$ gives very accurate results with a RMSD of less than 1 kJ/mol from the estimated CCSD(T) basis set limits. It still enables the calculation of reaction energies within chemical accuracy. Also PNO-CCSD(F12*)(T) with these settings shows a good performance with a RMSD of 1.08 kJ/mol. With the (T0) correction the distribution of the errors is much broader than for methods with the (T) and (T*) correction. The accuracy of PNO-CCSD(F12*)(T0)/TZ-F12 is comparable to that of PNO-

Table 1. Statistical Measures of the Errors (kJ/mol) in the Test Set with Respect to the Reference Values from Ref 37

basis	triples	T_{PNO}	Δ_{RMSD}	$\lvert\bar{\Delta}\rvert$
TZ-F12	(T)	none[37]	0.28	0.38
TZ-F12	(T)	10^{-8}	1.08	0.73
TZ-F12	(T)	10^{-7}	2.70	1.96
TZ-F12	(T*)	10^{-8}	0.89	0.58
TZ-F12	(T*)	10^{-7}	2.35	1.70
TZ-F12	(T0)	10^{-8}	1.77	1.24
TZ-F12	(T0)	10^{-7}	3.23	2.39
TZ-F12	(T0*)	10^{-8}	1.49	1.05
TZ-F12	(T0*)	10^{-7}	2.86	2.13
DZ-F12	(T)	10^{-8}	2.01	1.49
DZ-F12	(T)	10^{-7}	3.32	2.54
DZ-F12	(T*)	10^{-8}	1.98	1.40
DZ-F12	(T*)	10^{-7}	2.96	2.16
DZ-F12	(T0)	10^{-8}	2.24	1.63
DZ-F12	(T0)	10^{-7}	3.69	2.84
DZ-F12	(T0*)	10^{-8}	1.87	1.30
DZ-F12	(T0*)	10^{-7}	3.20	2.43

DOI: 10.1021/acs.jctc.7b00180
J. Chem. Theory Comput. 2017, 13, 2623−2633

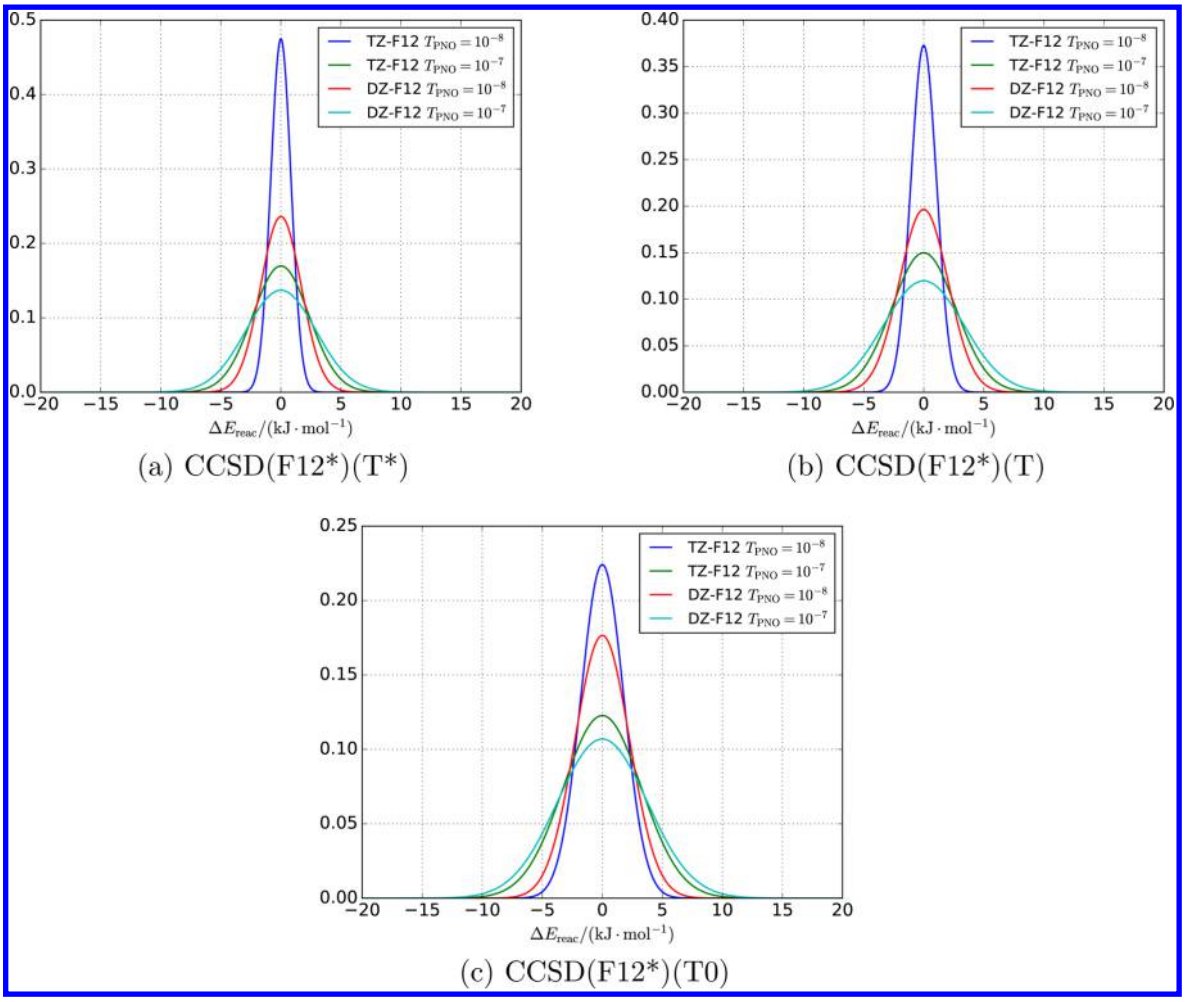


Figure 1. Deviations of PNO-CCSD(F12*)(T) calculations for reaction energies from the CCSD(T) basis set limits visualized as normal distributions.

CCSD(F12*)(T*) in the smaller DZ-F12 basis, which makes the (T0) correction unattractive as it does not significantly improve upon a (T*) calculation in a smaller basis. The scaled (T0*) correction improves upon the (T0) correction, but (T0*) is still inferior to either the (T) or (T*) correction.

Weak Intermolecular Interactions. To evaluate the accuracy of explicitly correlated local PNO-based methods for interaction energies we used the S66 test set developed by Hobza et al.[61] This test set contains 66 complexes of varying size. For instance it contains the water dimer as well as different uracil−uracil or benzene−uracil complexes. The test set includes examples dominated by dispersion interactions as well as hydrogen-bonded complexes. We assess the accuracy of the different PNO-CCSD(F12*)(T) variants in combination with the aDZ and aTZ basis sets.

Before looking at the S66 set as a whole, we use a subset with the dimers 1−22, 32−33, 50−51, 54, 59, 60, 64−66 of the S66 set to validate the PNO methods against their canonical counterparts. [Note that for PNO-CCSD(F12*)(T0) the canonical reference is also CCSD(F12*)(T) without (T0) approximation.] For that purpose calculations with the PNO and TNO thresholds set to $T_{PNO} = T_{TNO} = 10^{-7}$, 10^{-8}, and 10^{-9} have been carried out and compared to canonical reference results in the aTZ basis. Figure 2 shows the mean absolute deviations $|\bar{\Delta}|$ from the canonical reference results. As comparison also the results for PNO-CCSD(F12*) are

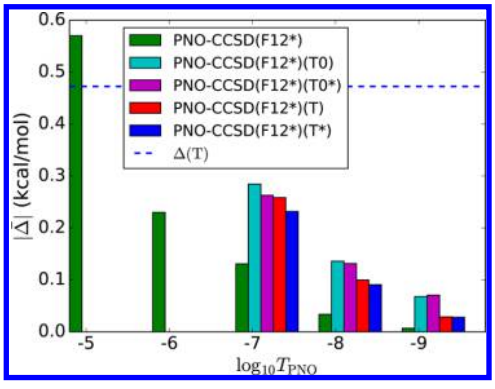


Figure 2. Mean absolute PNO and TNO truncation error in the counterpoise corrected interaction energies for the first 22 dimers in the S66 test set given in kcal/mol at the PNO-CCSD(F12*)(T) level of theory in the aTZ basis.

shown. Furthermore, the average size of the triples correction Δ(T) at the CCSD(F12*)(T) level is sketched as a horizontal line.

The error of PNO-CCSD(F12*)(T) is larger than for PNO-CCSD(F12*) but on average not more than a factor of 2 larger. Since for CCSD(F12*)(T) the remaining basis set error is a bit larger than for CCSD(F12*), also somewhat larger PNOs errors are acceptable. More important is that the overall PNO

DOI: 10.1021/acs.jctc.7b00180
J. Chem. Theory Comput. 2017, 13, 2623−2633

truncation error of PNO-CCSD(F12*)(T) is small compared to the typical size of the triples correction Δ(T) which is the case for $T_{PNO} = T_{TNO} \leq 10^{-7}$. PNO-CCSD(F12*)(T0) performs slightly not as good as PNO-CCSD(F12*)(T), although the effect for the studied subset with $T_{PNO} = T_{TNO} \geq 10^{-8}$ is small; but with very tight thresholds $T_{PNO} = T_{TNO} \geq 10^{-9}$, and the error from the (T0) approximation is about as large as the PNO and TNO truncation errors. The errors of PNO-CCSD(F12*)(T0*) and PNO-CCSD(F12*)(T*) show similar behaviors as their unscaled counterparts, although for loose thresholds their deviations from the canonical results are a bit smaller. The scaling factors in a PNO implementation are typically a bit larger than in the canonical case and converge with smaller T_{PNO} to the canonical values. Table 2 shows the

Table 2. Averaged Scaling Factors for (T*) and (T0*) for the Dimers in the aTZ Basis for Different T_{PNO} and the Canonical Limit

T_{PNO}			
10^{-7}	10^{-8}	10^{-9}	canonical
1.133	1.108	1.102	1.099

average scaling factors for the dimers in the studied subset for different thresholds. The larger scaling factors repair to some extent the TNO truncation error. For intermolecular interactions this additional correction is small but for thresholds like 10^{-7} still noteworthy. For all tested dimers, the binding energies obtained with the PNO implementation are a bit smaller (less negative). This is the case with the (T) as well as with the (T0) triples correction. For $T_{PNO} = T_{TNO} \geq 10^{-8}$ the energies are on average around 0.11 kcal/mol higher for the (T) correction and 0.14 kcal/mol for the (T0) correction.

New Basis Set Limits. After we evaluated the accuracy of PNO-CCSD(F12*)(T) for the reduced test set relative to canonical results, we applied PNO-CCSD(F12*)(T) to the whole set in order to establish improved CCSD(T) basis set limits. This is a typical example for applications of PNO-CCSD(F12*)(T) since the S66 test set contains several complexes which are currently out of reach for canonical CCSD(F12*)(T)/aTZ calculations. Based on the results from the previous paragraph we used $T_{PNO} = T_{TNO} = 10^{-8}$ to determine the improved basis set limits. The original reference data for post-MP2 methods in the S66 set was obtained by a composite scheme[61]

$$E(\mathrm{MCOR/CBS}) = E(\mathrm{HF}) + E^{\mathrm{corr}}(\mathrm{MP2}) + \Delta\mathrm{MCOR} \tag{11}$$

which combines the HF energy E(HF) in the aQZ basis with the extrapolated MP2 CBS limit E^{corr}(MP2) using the aTZ and aQZ basis and a higher order correction ΔMCOR in the aDZ basis:

$$\Delta\mathrm{MCOR} = E(\mathrm{MCOR}) - E(\mathrm{MP2}) \tag{12}$$

Due to the high computational costs of the canonical methods no larger basis sets could be used. The PNO approximation lifts this problem, and in combination with F12 theory it is possible to replace the composite scheme by a single calculation at the PNO-CCSD(F12*)(T)/aTZ level which also uses the HF energy in the aTZ basis in combination with the CABS singles[51] basis set correction.

To demonstrate that this setup enables us to improve upon the composite scheme in ref 61 we employ the canonical CCSD(F12*)(T) results for the studied subset and the existing MP2-F12/CBS limits from ref 46. In this study very large basis sets were used, so that the very accurate HF and MP2 CBS limits will lead to a better accuracy if they are used for the E(HF) and E^{corr}(MP2) contribution in the composite scheme. The higher order correction is then calculated with MP2-F12 and CCSD(F12*)(T) results extrapolated to the CBS limit from the aDZ and aTZ basis. We refer to the combination of the composite scheme with the extrapolated higher order correction as E^2_{lim}. The extrapolation of the required MP2-F12 and CCSD(F12*)(T) energies is done using an equation of Hill et al.[62]

$$E^{\infty}_{\mathrm{corr}}(X, Y) = \frac{X^{\alpha}E_{\mathrm{corr}}(X) - Y^{\alpha}E_{\mathrm{corr}}(Y)}{X^{\alpha} - Y^{\alpha}} \tag{13}$$

with optimized exponents. For the given basis sets the recommended values are 2.298829 for MP2, 2.483070 for CCSD, and 2.615472 for the triples correction. The separate extrapolation of the CCSD and triples part is recommended since the triples correction contains no explicitly correlated terms and therefore converges slower to the basis set limit, which can be accounted for by separate extrapolation. In cases where no results from ref 46 are available we use the extrapolated CCSD(T) correlation energies and the HF energy in the aTZ basis with CABS singles correction and refer to these limits as E^1_{lim}.

The extrapolated energies for both schemes as well as the PNO results with and without counterpoise correction[63] (CP) and the old S66 reference values are listed in Table 3. The mean absolute deviation $|\bar{\Delta}|$ of the non CP corrected PNO results is smaller than either for the S66 or the CP corrected PNO results. The mean absolute deviation $|\bar{\Delta}|$ from the E^1_{lim} limits is for the non CP corrected PNO results around a factor of 2 smaller than for the S66 values and for the very accurate E^2_{lim} limits $|\bar{\Delta}|$ even differs nearly by about an order of magnitude. The agreement of the CP corrected PNO energies is a bit worse. Overall the non CP corrected results show a better agreement with the extrapolated values. Such a behavior of explicitly correlated methods is not observed for the first time. Already Lange and Lane[64] showed for a small test set that F12 results without CP correction converge faster to the CBS limit. F12 calculations benefit from a cancellation of the basis set superposition error (BSSE) and the basis set incompleteness error (BSIE). The PNO approximation enhances this effect. It reduces the BSSE, and the PNO truncation error adds to the BSIE. From the data we can conclude that the PNO-CCSD(F12*)(T)/aTZ results without CP correction improve on the existing data, while the results with correction are slightly inferior.

At the PNO-CCSD(F12*)(T)/aTZ level of theory afterward calculations for the whole S66 set were carried out, and the original S66 values are compared to the new data, where for the analysis the deviations were calculated as $\Delta = E_{S66} - E_{PNO}$. Table 4 shows for PNO-MP2-F12, PNO-CCSD(F12*), and PNO-CCSD(F12*)(T) the root-mean-square deviation Δ_{RMSD}, the mean absolute $|\bar{\Delta}|$, and the mean deviation $\bar{\Delta}$ for the original S66 data with respect to the improved reference values obtained for the respective method with the composite scheme in eqs 11 and 12. For PNO-MP2-F12 we observe a nearly perfect agreement with the reference data. The Δ_{RMSD} is only 0.07 kcal/mol. The better agreement of the MP2 energies is due to the fact that for MP2 no composite scheme is applied. The extrapolation delivers more accurate limits, which better

DOI: 10.1021/acs.jctc.7b00180
J. Chem. Theory Comput. 2017, 13, 2623−2633

Table 3. Extrapolated CCSD(F12*)(T)/CBS Limits in Comparison with PNO-CCSD(F12*)(T)/aTZ with $T_{PNO} = T_{PNO}10^{-8}$ (E_{PNO}) and the Composite CCSD(T) Results from the Original S66 Benchmark[61] (E_{S66})[a]

dimer	E^1_{lim}	E^2_{lim}	E_{PNO}	E^{CP}_{PNO}	E_{S66}
01 water−water	−5.084	−5.019	−4.996	−4.874	−4.918
02 water−MeOH	−5.777	−5.709	−5.682	−5.526	−5.592
03 water−MeNH2	−7.096	−7.032	−7.018	−6.876	−6.908
04 water−peptide	−8.311		−8.192	−8.023	−8.103
05 MeOH−MeOH	−5.934	−5.856	−5.853	−5.693	−5.757
06 MeOH−MeNH2	−7.728	−7.654	−7.635	−7.477	−7.554
07 MeOH−peptide	−8.450		−8.304	−8.129	−8.230
08 MeOH−water	−5.166	−5.091	−5.111	−4.968	−5.009
09 MeNH2−MeOH	−3.160	−3.126	−3.113	−2.996	−3.059
10 MeNH2−MeNH2	−4.251	−4.219	−4.192	−4.076	−4.160
11 MeNH2−peptide	−5.539		−5.425	−5.266	−5.419
12 MeNH2−water	−7.458	−7.397	−7.389	−7.228	−7.266
13 peptide−MeOH	−6.380		−6.263	−6.073	−6.187
14 peptide−MeNH2	−7.634		−7.531	−7.375	−7.454
15 peptide−peptide	−8.843		−8.686	−8.477	−8.630
16 peptide−water	−5.283		−5.213	−5.074	−5.124
17 uracil−uracil_BP	−17.662		−17.323	−16.861	−17.182
18 water−pyridine	−7.026		−6.951	−6.827	−6.857
19 MeOH−pyridine	−7.573		−7.465	−7.335	−7.410
20 AcOH−AcOH	−19.595		−19.311	−18.989	−19.093
21 AcNH2−AcNH2	−16.682		−16.437	−16.135	−16.265
22 AcOH−uracil	−19.985		−19.664	−19.302	−19.491
32 uracil−ethyne	−3.692		−3.611	−3.481	−3.738
33 pyridine−ethene	−1.800		−1.765	−1.646	−1.872
50 benzene−ethyne_CH-π	−2.925		−2.934	−2.769	−2.867
51 ethyne−ethyne_TS	−1.528	−1.546	−1.536	−1.434	−1.524
54 benzene−water OH-π	−3.408		−3.331	−3.09	−3.277
59 ethyne−water CH−O	−2.972	−2.901	−2.967	−2.843	−2.85
60 ethyne−AcOH OH-π	−5.010	−4.910	−4.892	−4.766	−4.868
64 peptide−ethene	−3.046		−2.994	−2.835	−2.999
65 pyridine−ethyne	−4.133		−4.108	−4.023	−3.991
66 MeNH2−pyridine	−3.990		−3.909	−3.771	−3.968
$\bar{\Delta}(E^1_{lim})$			0.104	0.278	0.172
$\lvert\bar{\Delta}\rvert(E^1_{lim})$			0.105	0.278	0.179
$\bar{\Delta}(E^2_{lim})$			0.006	0.142	0.083
$\lvert\bar{\Delta}\rvert(E^2_{lim})$			0.021	0.142	0.083

[a]For E^1_{lim} the non-CP corrected energies in the aDZ and aTZ basis are used in the extrapolation scheme of Hill et al.[62] For E^2_{lim} the CBS limit is estimated using eq 11 with the MP2-F12/CBS limits from ref 46 and MP2-F12 and CCSD(F12*)(T) CBS(23) values for the higher order correction ΔMCOR. For comparison with ref 61 all values are given in kcal/mol.

Table 4. Statistical Parameters To Characterize the Deviation of the Original S66 Reference Data to Our Results Given in kcal/mol for the aTZ-Basis Using $T_{PNO} = T_{TNO} = 10^{-8}$[a]

method	Δ_{RMSD}	$\lvert\bar{\Delta}\rvert$	$\bar{\Delta}$
PNO-MP2-F12	0.074	0.070	0.070
PNO-CCSD(F12*)	0.110	0.083	−0.045
PNO-CCSD(F12*)(T)	0.100	0.075	0.014
PNO-CCSD(F12*)(T*)	0.109	0.089	0.048

[a]For the comparison always the data of the same level of theory was used.

coincide with our results. The deviations for the other methods are significantly larger but still show a good overall agreement of the S66 data to the PNO results.

The comparison with the old reference values shows that at the MP2 level the S66 values slightly underestimate the MP2 basis limits for the binding energies since no explicitly correlated terms are used. However, at the CCSD level the old S66 reference values overbind, which is related to the overshooting of the MP2 contribution which is only partially corrected when the higher order correction is computed in the small aDZ basis. At the CCSD(T) level the binding energy is again a bit underestimated. To emphasize this conclusion Table 6 shows the mean error of the different correlation contributions to the binding energies. Besides for CCSD the new binding energies are in most cases a bit larger.

Nevertheless the deviations of the old data and our results are in most cases small. Larger deviations mainly occur for complexes in the block S20−S33, where differences up to 0.6 kcal/mol are observed. This mainly occurs for examples with a large singles contribution, and in most of those cases the S66 values predict a stronger binding than the new reference data. In the composite scheme the MP2 contribution leads to an overshooting. Table 5 assembles examples with good and bad agreement of both data sets. For these examples additional PNO calculations in the aDZ and aTZ basis were used for an

DOI: 10.1021/acs.jctc.7b00180
J. Chem. Theory Comput. 2017, 13, 2623−2633

Table 5. Compilation of PNO-CCSD(F12*)(T) Interaction Energies in the aDZ and aTZ Basis As Well As the Extrapolated CBS Limits E_{lim}[a]

dimer	(kcal/mol)				
	E_{aDZ}	E_{aTZ}	E_{lim}	Δ_{lim}^{S66}	Δ_{lim}^{aTZ}
01 water−water	−5.069	−4.996	−4.982	0.064	−0.014
09 MeNH2−MeOH	−3.251	−3.113	−3.081	0.022	−0.032
10 MeNH2−MeNH2	−4.366	−4.192	−4.145	−0.015	−0.047
11 MeNH2−peptide	−5.657	−5.425	−5.355	−0.064	−0.070
14 peptide−MeNH2	−7.784	−7.531	−7.453	−0.001	−0.078
25 pyridine−pyridine $\pi-\pi$	−4.241	−3.688	−3.479	−0.416	−0.210
26 uracil−uracil $\pi-\pi$	−10.287	−9.582	−9.312	−0.517	−0.270
27 benzene−pyridine $\pi-\pi$	−3.806	−3.289	−3.111	−0.328	−0.178
28 benzene−uracil $\pi-\pi$	−6.213	−5.466	−5.191	−0.522	−0.275
29 pyridine−uracil $\pi-\pi$	−7.257	−6.523	−6.226	−0.593	−0.297

[a]Additionally the differences to the S66 reference data and the results on the PNO-CCSD(F12*)(T)/aTZ level of the theory are given.

Table 6. Mean and Absolute Deviation of the Old and New Reference Data for the Different Contributions to the Correlation Energy

	$\bar{\Delta}$	$\lvert\bar{\Delta}\rvert$
MP2 − HF	0.031	0.033
CCSD − MP2	−0.118	0.128
(T)	0.060	0.060

extrapolation to the CBS limit in order to verify again that the PNO calculations give the better accuracy. A table with all results is given in the Supporting Information. For the analysis the differences Δ_{lim}^{S66} and Δ_{lim}^{aTZ} to the extrapolated energies are calculated. The largest deviations occur in the block S25−S29. For the old S66 reference deviations up to 0.5 kcal/mol are observed. This is for example the case for the complexes S26 and S29, where the interaction energy is overestimated. The discussed energy difference are of course small and close to the intrinsic accuracy of the underlying method, but still it is important to mark this example in the S66 set since it is used for the determination of many empirical parameters. Therefore, every work to better quantify the error bar in the S66 data is of importance.

Cooperative Many-Body Effects. We now turn to a challenging case: the binding energy in a dodecahedron water cluster which was already the subject of previous studies.[65−67] The structure is taken from ref 66 and depicted in Figure 3. Calculations on this system suffer from a huge BSSE, which can lead to wrong conclusions. Anacker and Friedrich[67] were able to carry out conventional and explicitly correlated calculations in large basis sets with the method of increments. They computed explicitly correlated calculations in the DZ-F12, TZ-F12, and QZ-F12 basis and conventional calculations in the basis sets aDZ′, aTZ′, and aQZ′.

Again, tight PNO and TNO thresholds are necessary since the total binding energy and its error are size-extensive quantities. We therefore used again $T_{PNO} = T_{TNO} = 10^{-8}$.

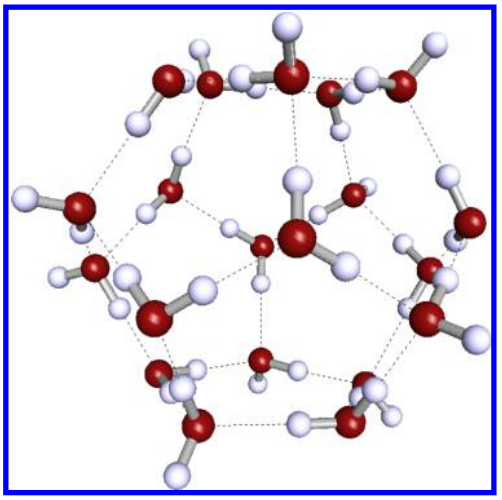


Figure 3. Representation of the $(H_2O)_{20}$ cluster with the shape of a dodecahedron.

The binding energy obtained as the difference of the cluster energy $E_{(H_2O)_{20}}$ and the energies of the geometry-relaxed monomers E_{H_2O}:

$$E_{bind} = E_{(H_2O)_{20}} - 20 \cdot E_{H_2O} \tag{14}$$

For the calculation without explicit correlation a CP correction was included which was calculated as the sum of the differences between the energies for the monomers i in the basis of the n-body cluster and the energies in the monomer basis computed for their geometries in the n-body cluster:

$$CP = \sum_{i}^{n} (E_i^{ijkl\ldots n} - E_i^i) \tag{15}$$

This many-body generalization of the Boys−Bernardi CP correction[63] for dimers is often referred to as the site−site function correction[68] (SSFC), but for simplicity we stay with the name "CP correction".

Anacker and Friedrich evaluated this CP correction only at the MP2 level and combined the MP2 CP correction also with CCSD(T) binding energies for the cluster. Since this approximation was only technically motivated we avoid it and evaluated the CP corrections at the same level as the binding energies. It can be expected that a CP correction at the MP2 level will overshoot for CCSD and CCSD(T) since MP2 usually overshoots the correlation energy.

For the aXZ' CP corrected result we carried extrapolations to the CBS limits. For the HF contribution we applied the two point extrapolation of Petersson and co-workers[69]

$$E_{HF}^{\infty}(X, Y) = \frac{e^{-a\sqrt{Y}} E_{HF}(X) - e^{-a\sqrt{X}} E_{HF}(Y)}{e^{-a\sqrt{Y}} - e^{-a\sqrt{X}}} \tag{16}$$

where $X < Y$, and a is set to 6.60 as recommended in ref 69. The extrapolation of the correlation energy was done with the formula proposed by Halkier et al.:[70]

$$E_{corr}^{\infty}(X, Y) = \frac{X^3 E_{corr}(X) - Y^3 E_{corr}(Y)}{X^3 - Y^3} \tag{17}$$

The CP corrected, uncorrected, and extrapolated binding energies are listed in Table 7. The CP correction has a large influence on the binding energy: for the aDZ′ basis around 31 kcal/mol, for aTZ′ around 15 kcal/mol, and for the aQZ′ basis

DOI: 10.1021/acs.jctc.7b00180
J. Chem. Theory Comput. 2017, 13, 2623−2633

Table 7. Binding Energies for the Dodecahedron Water Cluster (kcal/mol) Obtained from PNO Methods with $T_{PNO} = 10^{-8}$ and Different Basis Sets As Well As the Extrapolated CBS Limits[a]

	aDZ′		aTZ′		aQZ′			
method	E_{int}	E_{int}^{CP}	E_{int}	E_{int}^{CP}	E_{int}	E_{int}^{CP}	CBS(23)	CBS(34)
HF	−130.45	−120.39	−122.45	−120.00	−121.37	−120.34	−119.94	−120.42
MP2	−206.16	−175.69	−203.01	−188.24	−200.90	−193.97	−194.11	−197.98
PNO-MP2	−205.78	−175.40	−202.53	−187.84	−200.43	−193.53	−194.43	−198.22
PNO-CCSD	−190.00	−161.12	−187.54	−174.23	−185.31	−180.05	−181.10	−184.84
PNO-CCSD(T0)	−202.13	−169.73	−199.85	−185.45	−197.62	−191.99	−193.42	−197.30
PNO-CCSD(T)	−202.97	−170.38	−200.61	−186.13			−194.12	−198.10
PNO-CCSD(T)$\frac{1}{2}$CP	−202.97	−186.68	−200.61	−193.37			−197.93	−197.69
inc-CCSD(T)[67]							−195.4	−198.0
inc-CCSD(T)/CBS(45)[67]		−197.4						

[a]For PNO-CCSD(T) additionally an extrapolation with only half the CP correction is included. The PNO-CCSD(T)/CBS(34) limit was estimated based on the assumption that the relative difference in the triples correction for (T0) and (T) in the aQZ basis is comparable to that in the aTZ basis.

Table 8. Binding Energies for the Conventional PNO Method Using a $T_{PNO} = 10^{-7}$ for the Dodecahedron Water Cluster (kcal/mol) for Different Basis Sets As Well As the Extrapolated CBS Limits

	aDZ′		aTZ′		aQZ′			
method	E_{int}	E_{int}^{CP}	E_{int}	E_{int}^{CP}	E_{int}	E_{int}^{CP}	CBS(23)	CBS(34)
PNO-MP2	−204.96	−174.67	−201.84	−187.12	−199.67	−192.77	−192.94	−196.72
PNO-CCSD	−189.10	−160.29	−186.90	−173.52	−184.79	−179.53	−179.67	−183.74
PNO-CCSD(T0)	−198.95	−166.70	−196.26	−181.86	−194.17	−188.53	−188.83	−193.23
PNO-CCSD(T)	−199.60	−167.16	−196.88	−182.38	−194.95	−189.27	−189.37	−193.84

still around 6 kcal/mol. This corresponds for the aDZ′ basis to a CP correction of about 1 kcal/mol per hydrogen bond. Compared to the CP correction of 0.74 kcal/mol in the same basis for the water dimer of the S66 set this is an increase of ≈36% due to cooperative effects in the $(H_2O)_{20}$ cluster. For quantitative predictions the CP correction is therefore necessary but makes a comparison of the canonical and PNO based methods more difficult since the PNO truncation removes already for medium tight thresholds (e.g., $T_{PNO} = 10^{-7}$ and $T_{PNO} = 10^{-8}$) most of the determinants that contribute to the BSSE and to recover these determinants even tighter thresholds are required, which would substantially increase the computational costs.

Nevertheless we observe a good agreement with the extrapolated CCSD(T) results obtained with the methods of increments. The binding energies extrapolated from the aTZ′ and aQZ′ values, CBS(34), deviate only 0.10 kcal/mol from the reference results. As in ref 67 (there for the MP2 level) we observe that the arithmetic average of the uncorrected and the CP corrected results, which we denote in the following as $\frac{1}{2}$CP, converges faster with the basis set than the uncorrected or fully corrected results. The CBS(34) energies obtained from the $\frac{1}{2}$CP results are in close agreement with the CBS(45) limits from ref 67.

Before we turn to the explicitly correlated results we want to demonstrate the importance of a tight PNO threshold for weak intermolecular interactions. In the literature $T_{PNO} = 10^{-7}$ is often considered as a balanced choice concerning accuracy and computational costs.[19,42] First test calculations with these settings for the $(H_2O)_{20}$ cluster lead to unacceptable large deviations. Table 8 summarizes some results. The extrapolated limits deviate by up to 4.3 kcal/mol from those obtained with $T_{PNO} = 10^{-8}$. This is clearly outside the bounds for chemical accuracy of 1 kcal/mol. It demonstrates that this system is challenging for PNO methods, but it is also a difficult case for other low scaling approaches and canonical methods due to the slow convergence with the basis set. There is, however, a large difference between the methods. For MP2 and CCSD the PNO truncation errors are with $T_{PNO} = 10^{-7}$ still below 1 kcal/mol; but when the triples correction is included, the truncation error increases. The triples correction is sensitive to the truncation of the CCSD doubles amplitudes.

The explicitly correlated results are listed in Table 9. They have been evaluated without CP correction to compare them to the reference values and since the F12 theory usually enhances the basis set convergence such that the CP correction does not improve the result significantly.[64] A comparison of the PNO results for CCSD(F12*)(T) with those from the incremental scheme shows on the first sight significant deviations. One

Table 9. Binding Energies with the Explicitly Correlated PNO Methods Using $T_{PNO} = 10^{-8}$ for the Dodecahedron Water Cluster (kcal/mol)

	DZ-F12	TZ-F12	CBS(23)
method	E_{int}	E_{int}	E_{int}
HF+CABS	−122.12	−121.22	−121.08
MP2-F12	−201.76	−200.76	−200.58
PNO-MP2-F12	−200.52	−199.49	−199.31
PNO-CCSD(F12*)	−184.82	−184.50	−184.59
PNO-CCSD(F12*)(T0)	−195.68	−196.00	−196.43
PNO-CCSD(F12*)(T)	−196.27	−196.71	−197.20
PNO-CCSD(F12*)(T0*)	−197.80	−196.93	
PNO-CCSD(F12*)(T*)	−198.51	−197.71	
inc-CCSD(F12*)(T)[67]	−198.5	−199.3	
inc-CCSD(F12*)(T)/QZ-F12[67]	−198.6		

DOI: 10.1021/acs.jctc.7b00180
J. Chem. Theory Comput. 2017, 13, 2623−2633

could of course enhance the agreement by further tightening the PNO threshold, but it is still important to analyze the origin of the unexpected deviations. We already observed for the conventional calculations that they suffer from a huge BSSE, and by calculating the CP correction for the F12 calculations we realize that even with explicit electron correlation there is a nonvanishing BSSE around 4.6 kcal/mol at the MP2-F12 level in the TZ-F12 basis as shown in Table 10. In a PNO implementation the determinants, which mainly cause the BSSE, are removed even with tight thresholds. For a rough estimate the CP correction was scaled by the ratio of the correction on the MP2-F12 and PNO-MP2-F12 levels. The larger BSSE in the canonical case explains to some part the deviation between the PNO and reference values.

Table 10. CP Correction for the Different Methods[a]

	DZ-F12		TZ-F12	
method	CP	CP*	CP	CP*
MP2-F12	4.63	4.63	2.21	2.21
PNO-MP2-F12	4.36	4.63	1.97	2.21
PNO-CCSD(F12*)	4.72	5.01	2.24	2.50
PNO-CCSD(F12*)(T0)	6.98	7.42	3.04	3.39
PNO-CCSD(F12*)(T)	7.15	7.60	3.10	3.46

[a]CP* indicates an estimated correction for the canonical case.

Since using even tighter thresholds is computationally not economic and still a large BSSE would remain, we use an extrapolation scheme to obtain more accurate results. The correlation energy is extrapolated using eq 13. For the CCSD(F12*) and (T) part again separate extrapolations were done since both parts show a different convergence behavior. For the extrapolation we used the recommended exponents of 3.144518 and 2.615472.[62] In accordance with work of Brauer et al.[71] we only used the non CP corrected energies since this leads to a better convergence. For the extrapolation of the HF part we used eq 16. The results are listed in Table 9. With −197.20 kcal/mol the binding energy is consistent with the results in Table 7.

Still, there remains an unexpected difference of 2.6 kcal/mol between the MP2/CBS(34) result of −197.98 kcal/mol and the explicitly correlated MP2-F12/CBS(23) value of −200.58 kcal/mol. This indicates that the remaining basis set error might still be significant. The slow convergence of the MP2 and MP2-F12 values to the common basis set limit was already noted in ref 67 where the MP2-F12/aQZ-F12 value of −199.8 kcal/mol and the MP2/CBS(45) results of −198.8 kcal/mol still deviate by 1 kcal/mol. At the MP2 level it is possible to carry out calculations in larger basis sets. In Table 11 we have compiled the results from MP2 calculations in the a5Z′ and a6Z′ basis sets and MP2-F12 in the a5Z′ basis. Eventually we obtained a good agreement between the MP2-F12/a5Z′ and MP2/CBS(56) results. We conclude that the MP2 basis set limit is −199.3 ± 0.2 kcal/mol. It is in the middle between the MP2-F12/aQZ-F12 and MP2/CBS(45) results from ref 67. The F12 results converge from above and the non-F12 results from below to this limit.

Table 11. MP2 Binding Energies for the Dodecahedron Water Cluster in the a5Z′ and a6Z′ Basis As Well As the MP2-F12 Energy in the a5Z′ Basis[a]

	E_{int}			
method	a5Z′	a6Z′	CBS(56)	a5Z′/F12
HF	−120.62	−120.45	−120.39	−120.33
MP2	−199.71	−199.45	−199.28	−199.30

[a]No CP correction was applied.

We can use this accurate estimate for the MP2 basis set limit to further improve the estimate for the CCSD(T) basis set limit as

$$E_{CCSD(T)/CBS} = E_{MP2\text{-}F12/a5Z'} + \Delta E_{CCSD(T)} \quad (18)$$

For the higher order correction we also consider the CBS(23) results on the PNO-CCSD(F12*)(T) and PNO-MP2-F12 level of theory. The results are given in Table 12. For the two canonical calculations a good agreement of the binding energies can be observed, so that a limit of −198.0 ± 0.4 kcal/mol can be concluded for CCSD(T). This value is again in the middle between the F12 and non-F12 results in ref 67. The results using the higher order correction calculated with PNO methods still differs by around 0.8 kcal/mol if the CBS(23) values are used and around 0.45 kcal/mol if the PNO-CCSD(F12*)-(T*)/TZ-F12 energies are used, which indicates that here also calculations in larger basis sets would be required. These results demonstrate that even for a challenging system such as the $(H_2O)_{20}$ cluster PNO-CCSD(F12*)(T*) provides at comparatively moderate computational costs results close to the highly accurate benchmark.

Table 12. CCSD(T) CBS Limits Using a Composite Scheme[a]

	E_{int}			
method	Δa5Z′	ΔQZ-F12	ΔPNO(23)	ΔPNO(T*)
CCSD(T)/CBS	−197.99	−198.01	−197.17	−197.52

[a]For the MP2 contribution the MP2-F12 energy in the a5Z′ basis is used. For the higher order correction four choices are presented: Δa5Z′ is calculated from MP2/a5Z′ and inc-CCSD(T)/a5Z′ (ref 67), ΔQZ-F12 is calculated from inc-MP2-F12/QZ-F12 and inc-CCSD-(F12*)(T)/QZ-F12 (ref 67), ΔPNO(23) is calculated from the PNO-CCSD(F12*)(T) and PNO-MP2-F12 CBS(23), and ΔPNO(T*) is calculated from PNO-CCSD(F12*)(T*)/TZ-F12 and PNO-MP2-F12/TZ-F12.

CONCLUSION

Benchmarks for test sets for reaction energies and weak intermolecular interactions show that the PNO and TNO truncation errors of PNO-CCSD(F12*)(T) and related methods decrease smoothly with the truncation thresholds. The accuracy of the triples correction increases in the order (T0) → (T0*) → (T) → (T*).

For reactions energies the RMS deviation from the CCSD(T) basis set limit values is for PNO-CCSD(F12*) in combination with all four variants of the triples correction already in the cc-pVDZ-F12 basis and $T_{PNO} = T_{TNO} = 10^{-7}$ below 1 kcal/mol; but only with the cc-pVTZ-F12 basis, the (T) and (T*) triples corrections, and $T_{PNO} = T_{TNO} = 10^{-8}$ also the maximum absolute errors are below the target mark for chemical accuracy of 1 kcal/mol. With PNO-CCSD(F12*)-(T*)/cc-pVTZ-F12 and $T_{PNO} = 10^{-8}$ the RMS deviation dropped even below 1 kJ/mol. With the (T0) and (T0*) triples corrections the RMS deviation is still in an acceptable range although significantly larger than for (T) and (T*), but for many cases its accuracy is only comparable to that of PNO-

DOI: 10.1021/acs.jctc.7b00180
J. Chem. Theory Comput. 2017, 13, 2623−2633

CCSD(F12*)(T*)/cc-pVDZ-F12 which is computationally cheaper.

For weak intermolecular interactions PNO and TNO thresholds of 10^{-8} are required to keep the truncation errors small. In combination with triple-ζ basis sets these settings give benchmark quality results. We could show that the accuracy of PNO-CCSD(F12*)(T)/aTZ challenges that of the existing S66 reference data. In most cases the new results agree very well with the original S66 data, but in the block 24−29 larger deviations are observed. These could be traced back to an overshooting of the MP2 basis set correction in the composite scheme that was used for the original S66 data. Compared to reaction energies we observe for binding energies of weak intermolecular complexes only minor differences between the (T) and (T0) triples corrections.

In addition to these test sets we studied the binding energy of a water dodecamer which is a challenging system for local correlation methods due to large a BSSE that leads to a slow convergence with the basis set. Also for this system tight PNO and TNO thresholds of 10^{-8} had to be used to achieve accurate results. Thresholds of 10^{-7} lead to errors up to 4 kcal/mol. It is possible to obtain results for CCSD(T) (−197.2 kcal/mol) and MP2 (−199.31 kcal/mol) which are in close agreement to recent benchmark values of Anacker and Friedrich,[67] but deviations to the explicitly correlated results still indicate that the CBS limit is not reached yet. At the MP2 level we were able to extrapolate from the a5Z′ and a6Z′ basis set results to a value, which is finally in close agreement to MP2-F12 results in the a5Z′ basis. The new best estimate for the MP2 basis set limit is −199.3 ± 0.2 kcal/mol and in the middle of the so far reported conventional and explicitly correlated results. With this improved MP2 basis set limit also the estimate for the CCSD(T) limit can be further improved to −198.0 ± 0.4 kcal/mol.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jctc.7b00180.

Detailed information on the obtained reaction and interaction energies for both test sets are assembled (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: Gunnar.Schmitz@rub.de.

ORCID

Gunnar Schmitz: 0000-0003-3677-4748

Christof Hättig: 0000-0002-5752-2710

Funding

Financial support by the Deutsche Forschungsgemeinschaft through grant no. HA 2588/7 is gratefully acknowledged.

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. *Chem. Phys. Lett.* **1989**, *157*, 479−483.
(2) Stoll, H. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1992**, *46*, 6700−6704.
(3) Paulus, B.; Fulde, P.; Stoll, H. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1995**, *51*, 10572−10578.
(4) Friedrich, J. *J. Chem. Theory Comput.* **2012**, *8*, 1597−1607.
(5) Friedrich, J.; Dolg, M. *J. Chem. Phys.* **2008**, *129*, 244105.
(6) Friedrich, J.; Hanrath, M.; Dolg, M. *J. Chem. Phys.* **2007**, *126*, 154110.
(7) Li, S.; Ma, J.; Jiang, Y. *J. Comput. Chem.* **2002**, *23*, 237−244.
(8) Li, W.; Piecuch, P. *J. Phys. Chem. A* **2010**, *114*, 8644−8657.
(9) Rolik, Z.; Kállay, M. *J. Chem. Phys.* **2011**, *135*, 104111.
(10) Kristensen, K.; Høyvik, I.-M.; Jansik, B.; Jørgensen, P.; Kjærgaard, T.; Reine, S.; Jakowski, J. *Phys. Chem. Chem. Phys.* **2012**, *14*, 15706−15714.
(11) Kristensen, K.; Ziółkowski, M.; Jansík, B.; Kjærgaard, T.; Jørgensen, P. *J. Chem. Theory Comput.* **2011**, *7*, 1677−1694.
(12) Kristensen, K.; Jørgensen, P.; Jansik, B.; Kjærgaard, T.; Reine, S. *J. Chem. Phys.* **2012**, *137*, 114102.
(13) Scuseria, G. E.; Ayala, P. Y. *J. Chem. Phys.* **1999**, *111*, 8330−8343.
(14) Lambrecht, D. S.; Doser, B.; Ochsenfeld, C. *J. Chem. Phys.* **2005**, *123*, 184102.
(15) Doser, B.; Lambrecht, D. S.; Ochsenfeld, C. *Phys. Chem. Chem. Phys.* **2008**, *10*, 3335−3344.
(16) Schütz, M.; Hetzer, G.; Werner, H.-J. *J. Chem. Phys.* **1999**, *111*, 5691.
(17) Schütz, M.; Werner, H.-J. *J. Chem. Phys.* **2001**, *114*, 661.
(18) Schütz, M. *J. Chem. Phys.* **2000**, *113*, 9986−10001.
(19) Neese, F.; Wennmohs, F.; Hansen, A. *J. Chem. Phys.* **2009**, *130*, 114108.
(20) Meyer, W. *Int. J. Quantum Chem.* **1971**, *5*, 341−348.
(21) Ahlrichs, R.; Lischka, H.; Staemmler, V.; Kutzelnigg, W. *J. Chem. Phys.* **1975**, *62*, 1225.
(22) Ahlrichs, R.; Driessler, F.; Lischka, H.; Staemmler, V.; Kutzelnigg, W. *J. Chem. Phys.* **1975**, *62*, 1235.
(23) Kutzelnigg, W. *Theoretica chimica acta* **1985**, *68*, 445−469.
(24) Klopper, W.; Kutzelnigg, W. *J. Phys. Chem.* **1990**, *94*, 5625−5630.
(25) Klopper, W.; Samson, C. C. M. *J. Chem. Phys.* **2002**, *116*, 6397−6410.
(26) Noga, J.; Kedžuch, S.; Simunek, J.; Ten-no, S. *J. Chem. Phys.* **2008**, *128*, 174103.
(27) Adler, T. B.; Knizia, G.; Werner, H.-J. *J. Chem. Phys.* **2007**, *127*, 221106.
(28) Torheyden, M.; Valeev, E. F. *Phys. Chem. Chem. Phys.* **2008**, *10*, 3410−3420.
(29) Hättig, C.; Tew, D. P.; Köhn, A. *J. Chem. Phys.* **2010**, *132*, 231102.
(30) Schmitz, G.; Hättig, C. *J. Chem. Phys.* **2016**, *145*, 234107.
(31) Riplinger, C.; Sandhoefer, B.; Hansen, A.; Neese, F. *J. Chem. Phys.* **2013**, *139*, 134101.
(32) Riplinger, C.; Pinski, P.; Becker, U.; Valeev, E. F.; Neese, F. *J. Chem. Phys.* **2016**, *144*, 024109.
(33) Ma, Q.; Werner, H.-J. *J. Chem. Theory Comput.* **2015**, *11*, 5291−5304. PMID: 26574323.
(34) Liakos, D. G.; Sparta, M.; Kesharwani, M. K.; Martin, J. M. L.; Neese, F. *J. Chem. Theory Comput.* **2015**, *11*, 1525−1539.
(35) Liakos, D. G.; Neese, F. *J. Chem. Theory Comput.* **2015**, *11*, 4054−4063. PMID: 26575901.
(36) Minenkov, Y.; Chermak, E.; Cavallo, L. *J. Chem. Theory Comput.* **2015**, *11*, 4664−4676. PMID: 26574257.
(37) Friedrich, J. *J. Chem. Theory Comput.* **2015**, *11*, 3596−3609. PMID: 26574443.
(38) Calbo, J.; Orti, E.; Sancho-Garcia, J. C.; Arago, J. *J. Chem. Theory Comput.* **2015**, *11*, 932−939. PMID: 26579747.
(39) Sparta, M.; Neese, F. *Chem. Soc. Rev.* **2014**, *43*, 5032−5041.
(40) Hansen, A.; Bannwarth, C.; Grimme, S.; Petrović, P.; Werlé, C.; Djukic, J.-P. *ChemistryOpen* **2014**, *3*, 177−189.
(41) Liakos, D. G.; Neese, F. *J. Chem. Theory Comput.* **2015**, *11*, 2137−2143. PMID: 26574416.
(42) Schmitz, G.; Hättig, C.; Tew, D. P. *Phys. Chem. Chem. Phys.* **2014**, *16*, 22167−22178.
(43) Boys, S. F. *Rev. Mod. Phys.* **1960**, *32*, 296−299.

DOI: 10.1021/acs.jctc.7b00180
J. Chem. Theory Comput. 2017, 13, 2623−2633

(44) Hansen, A.; Liakos, D. G.; Neese, F. *J. Chem. Phys.* **2011**, *135*, 214102.
(45) Riplinger, C.; Neese, F. *J. Chem. Phys.* **2013**, *138*, 034106.
(46) Hättig, C.; Tew, D. P.; Helmich, B. *J. Chem. Phys.* **2012**, *136*, 204105.
(47) Werner, H.-J.; Knizia, G.; Krause, C.; Schwilk, M.; Dornbach, M. *J. Chem. Theory Comput.* **2015**, *11*, 484−507.
(48) Tew, D. P.; Hättig, C. *Int. J. Quantum Chem.* **2013**, *113*, 224−229.
(49) Häser, M.; Almlöf, J. *J. Chem. Phys.* **1992**, *96*, 489.
(50) Fliegl, H.; Hättig, C.; Klopper, W. *Int. J. Quantum Chem.* **2006**, *106*, 2306−2317.
(51) Knizia, G.; Adler, T. B.; Werner, H.-J. *J. Chem. Phys.* **2009**, *130*, 054104.
(52) Schmitz, G.; Helmich, B.; Hättig, C. *Mol. Phys.* **2013**, *111*, 2463−2476.
(53) Bachorz, R. A.; Bischoff, F. A.; Glöß, A.; Hättig, C.; Höffner, S.; Klopper, W.; Tew, D. P. *J. Comput. Chem.* **2011**, *32*, 2492−2513.
(54) Ahlrichs, R.; Bär, M.; Horn, H.; Kölmel, C. *Chem. Phys. Lett.* **1989**, *162*, 165−169.
(55) Dunning, T. H. J. *J. Chem. Phys.* **1989**, *90*, 1007.
(56) Peterson, K. A.; Adler, T. B.; Werner, H.-J. *J. Chem. Phys.* **2008**, *128*, 084102.
(57) Weigend, F.; Köhn, A.; Hättig, C. *J. Chem. Phys.* **2002**, *116*, 3175−3183.
(58) Weigend, F.; Häser, M.; Patzelt, H.; Ahlrichs, R. *Chem. Phys. Lett.* **1998**, *294*, 143.
(59) Yousaf, K. E.; Peterson, K. A. *J. Chem. Phys.* **2008**, *129*, 184108.
(60) Friedrich, J.; Hänchen, J. *J. Chem. Theory Comput.* **2013**, *9*, 5381−5394.
(61) Rezac, J.; Riley, K. E.; Hobza, P. *J. Chem. Theory Comput.* **2011**, *7*, 2427−2438.
(62) Hill, J. G.; Peterson, K. A.; Knizia, G.; Werner, H.-J. *J. Chem. Phys.* **2009**, *131*, 194105.
(63) Boys, S.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553−566.
(64) de Lange, K. M.; Lane, J. R. *J. Chem. Phys.* **2011**, *134*, 034301.
(65) Fanourgakis, G. S.; Aprà, E.; Xantheas, S. S. *J. Chem. Phys.* **2004**, *121*, 2655−2663.
(66) Bryantsev, V. S.; Diallo, M. S.; van Duin, A. C. T.; Goddard, W. A. *J. Chem. Theory Comput.* **2009**, *5*, 1016−1026.
(67) Anacker, T.; Friedrich, J. *J. Comput. Chem.* **2014**, *35*, 634−643.
(68) Wells, B. H.; Wilson, S. *Chem. Phys. Lett.* **1983**, *101*, 429−434.
(69) Zhong, S.; Barnes, E. C.; Petersson, G. A. *J. Chem. Phys.* **2008**, *129*, 184116.
(70) Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Olsen, J. *Chem. Phys. Lett.* **1999**, *302*, 437−446.
(71) Brauer, B.; Kesharwani, M. K.; Martin, J. M. L. *J. Chem. Theory Comput.* **2014**, *10*, 3791−3799.

DOI: 10.1021/acs.jctc.7b00180
J. Chem. Theory Comput. 2017, 13, 2623−2633