

## Molecular Physics

An International Journal at the Interface Between Chemistry and Physics

ISSN: (Print) (Online) Journal homepage: <https://www.tandfonline.com/loi/tmph20>

# Pursuing the basis set limit of CCSD(T) non-covalent interaction energies for medium-sized complexes: case study on the S66 compilation

Péter R. Nagy, László Gyevi-Nagy, Balázs D. Lőrincz & Mihály Kállay

To cite this article: Péter R. Nagy, László Gyevi-Nagy, Balázs D. Lőrincz & Mihály Kállay (2022): Pursuing the basis set limit of CCSD(T) non-covalent interaction energies for medium-sized complexes: case study on the S66 compilation, Molecular Physics, DOI: [10.1080/00268976.2022.2109526](https://doi.org/10.1080/00268976.2022.2109526)

To link to this article: <https://doi.org/10.1080/00268976.2022.2109526>



View supplementary material [↗](#)



Published online: 08 Aug 2022.



Submit your article to this journal [↗](#)



Article views: 84



View related articles [↗](#)



View Crossmark data [↗](#)



# Pursuing the basis set limit of CCSD(T) non-covalent interaction energies for medium-sized complexes: case study on the S66 compilation

Péter R. Nagy <sup>a,b</sup>, László Gyevi-Nagy <sup>a,b</sup>, Balázs D. Lőrincz <sup>a,b</sup> and Mihály Kállay <sup>a,b</sup>

<sup>a</sup>Faculty of Chemical Technology and Biotechnology, Department of Physical Chemistry and Materials Science, Budapest University of Technology and Economics, Budapest, Hungary; <sup>b</sup>ELKH-BME Quantum Chemistry Research Group, Budapest, Hungary

## ABSTRACT

Our recently developed conventional, explicitly correlated, and local natural orbital (LNO) based coupled cluster with single, double, and perturbative triple excitations [CCSD(T)] methods are utilised to closely approach the basis set limit of interaction energies between medium-sized molecular complexes. The study is partly motivated by our previous finding that state-of-the-art diffusion Monte Carlo and CCSD(T) interaction energies are not in complete agreement even for some of the dimers in the popular S66 compilation [Al-Hamdani and Nagy et al. Nat. Commun. (2021) 12:3927]. We improve upon the previous S66 CCSD(T) reference by using up to the quadruple- $\zeta$  (QZ) level CCSD(T) contributions, utilising our recent (T+) approach and suggesting a more robust explicitly correlated CCSD reference. Our efficient LNO-CCSD(T) method allowed us to also compute tightly converged, quintuple- $\zeta$  level interaction energies for the S66 set. Moreover, the new CCSD(T)/QZ references enabled the so far most rigorous benchmark of the LNO and other popular local correlation approaches on the S66 set. The accuracy of local approximations was found in the few 10 to few 100 cal/mol interaction energy error range, assuming sufficiently tight settings, such as the Tight and the Normal settings of LNO-CCSD(T) threshold combinations, respectively.

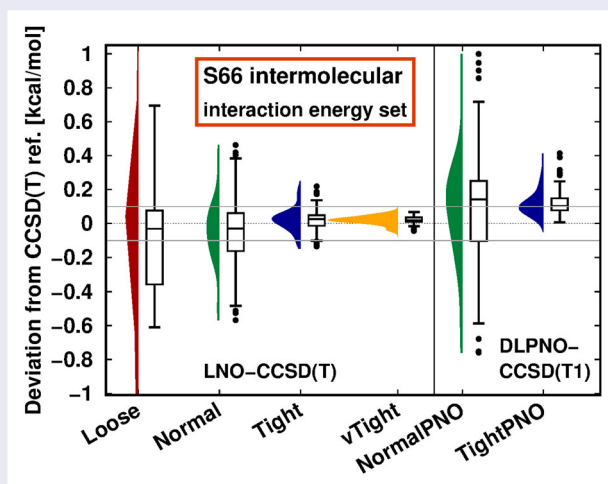
## ARTICLE HISTORY

Received 13 June 2022

Accepted 28 July 2022

## KEYWORDS

Non-covalent interactions; coupled-cluster theory; CCSD(T); local and explicit correlation methods



## 1. Introduction

Non-covalent intermolecular interactions are ubiquitous and play a key role in a large number of (bio)chemical systems. The precise and efficient modelling of intermolecular interactions is still one of the great challenges

of computational modelling since accurate description is required for a large number of individually small but cumulatively important interaction components. Since electron correlation effects have to be taken into account due to the required sub-kcal/mol accuracy, a

**CONTACT** Péter R. Nagy nagy.peter@vbk.bme.hu Faculty of Chemical Technology and Biotechnology, Department of Physical Chemistry and Materials Science, Budapest University of Technology and Economics, Műegyetem rkp. 3., Budapest H-1111, Hungary; ELKH-BME Quantum Chemistry Research Group, Műegyetem rkp. 3., Budapest H-1111, Hungary

Supplemental data for this article can be accessed here. <https://doi.org/10.1080/00268976.2022.2109526>

number of wave function based methods have been made available for intermolecular interaction computations [1–4]. Among others, the application of the symmetry-adapted perturbation theory (SAPT) [5–7], diffusion Monte Carlo (DMC) [8–11], random phase approximation (RPA) [12, 13], and coupled-cluster (CC) methods [14], in particular the gold standard CC model with single and double excitations (CCSD), augmented with perturbative triples correction [CCSD(T)] [15] have been particularly successful in this context.

For the performance assessment and comparison of the various computational methods, a number of test compilations have been suggested including molecular complexes from a few atoms to extensive supramolecular or biomolecular examples representative of all major types of non-covalent interactions [16–23]. Among those the S66 list of Řezáč, Riley, and Hobza [17] has become one of the standard benchmark sets. The S66 compilation contains 66 dimers composed from 14 monomer molecules and represents a wide range of interaction patterns occurring in organic and biomolecules, such as hydrogen bonding or dispersion interactions between both aliphatic and aromatic compounds. The largest dimers of S66 reach up to 34 atoms and 116 electrons, which pose a considerable difficulty for computationally demanding methods, such as conventional CCSD(T). While the previous CCSD(T) benchmarks of S66 appear to be well-established [17, 24–26], we found recently that highly-converged CCSD(T) and DMC interaction energies (IEs) are not completely consistent for a few  $\pi$ - $\pi$  stacking complexes among the most challenging S66 dimers [27]. The deviation of at least 0.1 kcal/mol between the CCSD(T) and DMC IEs was outside of the methods' uncertainty estimates. This finding provided an additional motivation to improve on the basis set convergence of the previous S66 IEs, to explore and eliminate basis set incompleteness as a potential source of deviation between DMC and CCSD(T) IEs.

Since the evaluation of conventional CCSD(T) energies scale with the seventh power of the system size, the largest molecules accessible with sufficiently large atomic orbital (AO) basis sets are also in the range of 20–40 atoms, even with recent massively parallel implementations [28–38]. Though the use of convergent basis set hierarchies [39], complete basis set (CBS) extrapolation techniques [40], and counterpoise (CP) corrections [41, 42] is invaluable, basis sets including diffuse and high angular momentum orbitals are still required to sufficiently approach the CBS limit of the CC IEs. Therefore, the original S66 IEs reference [17] and its revisited variants [24–26] had to rely on composite or focal point type approaches in combination with

second-order Møller–Plesset (MP2) based basis set corrections [43, 44]. However, the well-known challenges of MP2, e.g. in accurately describing aromatic interactions provide a reason for minimising the role of MP2-based corrections when aiming at improved CCSD(T)/CBS estimates.

To that end, the basis set used for the CCSD(T) part, and thus the computational effort, has to be increased. To accelerate such extensive CCSD(T) computations, we combined the t1-transformed CCSD expressions and the density-fitting (DF) approximation [45] with our integral-direct and efficiently parallelised CCSD and (T) algorithms into a highly-optimised, parallel, and data-economic CCSD(T) code, scalable up to over 30 atoms representing the largest quadruple- $\zeta$  level CCSD(T) computation ever presented [37]. Combined with the rapid development in computational hardware, our new CCSD(T) code enabled us here to update the original [17] double- $\zeta$  and the latest [26] triple- $\zeta$  CCSD(T) components of the S66 IEs, presented 11 and 5 years ago, respectively, by quadruple- $\zeta$  level CCSD(T) contributions.

Alternatively, explicitly correlated CCSD approaches, which explicitly contain the function of interelectronic distances (F12), can significantly accelerate convergence of CCSD toward its CBS limit [46–50]. The most cost-efficient variants currently include the CCSD-F12a/b schemes of Knizia, Adler, and Werner [51, 52], the CCSD(2)<sub>F12</sub> approach of Valeev and Crawford [53, 54], and the CCSD(F12\*) method of Hättig, Tew, and Köhn [55], the latter of which is employed here with the corresponding MP2-F12 approach [56]. These F12-based CCSD methods, especially with at least triple- $\zeta$  basis sets, exhibit similarly good performance benchmarked extensively also for non-covalent interactions [57–61].

Since an affordable explicitly correlated treatment of the (T) corrections is not yet available, in practise MP2-based scaling approaches are employed to the (T) terms [52, 62]. Most recently, we showed that the first, not size-consistent variant of Knizia *et al.* [52], that is, the (T\*) approach, can be effectively replaced by a similar, but size-consistent scheme termed (T+) [62]. While the (T\*) and (T+) results are often similar when the basis set convergence patterns match for the dimer and the monomers, (T\*) can exhibit considerable size-consistency errors when the monomers of a dimer are markedly different. We showed that this issue is effectively resolved by the (T+) alternative [62], and thus we will employ here the resulting CCSD(F12\*)(T+) method for the S66 dimers too.

To further increase the efficiency of CC methods, a number of successful local correlation approaches have been introduced exploiting the relatively quick distance

decay or local nature of electron correlation [63–92]. Currently, strategies that combine these local approximations with natural orbital [93–96] (NO)-based compressions stand out [87, 89, 91, 97]. Here, we will utilise and benchmark some of these state-of-the-art local CCSD(T) approaches, including the pair NO (PNO) based method of Hättig, Tew, and Schmitz [70, 97, 98], the PNO-L method of Werner and Ma [89, 99], the domain-based local PNO (DLPNO) method of Neese, Valeev, and co-workers [87, 100, 101], as well as the local natural orbital (LNO) approach developed by Nagy and Kállay [36, 91].

The efficiency of these local CCSD(T) methods has already been utilised to investigate the non-covalent interactions in large molecular systems beyond the capabilities of conventional CCSD(T) [4, 20, 27, 87, 89, 91, 102–111]. For instance, we reported some of the largest local CCSD(T) computations for the up to 101-atom complexes of the L7 compilation [18] and a 132-atom supramolecular fullerene dimer using very tightly converged LNO-CCSD(T) with augmented quintuple- $\zeta$  basis sets [27]. Such challenging complexes with extended  $\pi$ - $\pi$  interactions limit the applicability range of all local CCSD(T) codes aiming at robust accuracy. Systems with more rapidly decaying long-range interactions and smaller volumetric density of electrons, such as biomolecules can be treated more effectively. We illustrated this recently by computing the quadruple- $\zeta$  level LNO-CCSD(T) interaction energy of a 1023-atom protein–ligand complex [91].

Clearly, these large-scale applications are far from the range of the at most 20–30-atom molecules and triple- $\zeta$  basis sets, where the accuracy of various local CCSD(T) approaches are usually documented. Therefore, it is important to show that local approximations have well-controlled and homogeneous accuracy for a wide range of molecules, across multiple large basis set choices, CBS extrapolation, and CP corrections, etc. Thus, we will also utilise the new, conventional CCSD(T) results to assess the LNO and other, PNO-based local correlation schemes in a complexity range that was not accessible before. Currently, more and more of such benchmarks would be needed considering the increasing popularity and wide applicability domain of local CCSD(T) approaches. Moreover, comparisons treating different local CCSD(T) methods on an equal footing [27, 91, 110, 112–115] remain scarce. Additionally, by utilising the unique capabilities of our LNO-CCSD(T) implementation [36, 91], we approach the CBS limit even closer by reporting very well converged, quintuple- $\zeta$  level LNO-CCSD(T) IEs for the first time for the entire S66 compilation.

## 2. Methods: algorithmic considerations and development

A brief summary is given of our previous and present algorithmic developments and programming efforts regarding the relevant CCSD(T) methods allowing for a more detailed analysis of the current capabilities and limitations.

### 2.1. Conventional and explicitly correlated CCSD(T)

In conventional or explicitly correlated CCSD(T), both the computational time and the data storage bottlenecks are often relevant since these scale with the seventh- and fourth power of the number of correlated occupied ( $n_o$ ) and virtual ( $n_v$ ) molecular orbitals (MOs), respectively. For a single CCSD iteration and the (T) parts, respectively, the rate determining steps require about  $2(n_o^2 n_v^4 / 4 + 4n_o^3 n_v^3)$  and  $2n_o^3 n_v^3 (n_o + n_v)$  floating point operations (FLOP). For instance, an extensive computation with  $n_o = 30$  and  $n_v = 1000$ , the FLOP requirement of 12 CCSD iterations and the (T) step amounts to about 96 and 55.6 petaFLOP ( $10^{15}$  FLOP), respectively. It has become possible to meet this high demand with the latest parallel CCSD(T) codes exploiting the appearance of many-core CPUs with teraFLOPS, i.e.  $10^{12}$  FLOP per second capabilities. However, conventional CCSD(T) applications reaching the 1000-orbital range are still scarce [28, 30–37].

Using our new integral-direct and highly optimised CCSD(T) code and 224 CPU cores for 3 days, we reported the largest CCSD(T) computation so far at the quadruple- $\zeta$  level involving 31 atoms and 1569 orbitals [37]. Consequently, the outstanding peak-performance utilisation and parallel-efficiency of our CCSD(T) code made the next step accessible in the basis set hierarchy for the entire S66 compilation. Namely, using the highly-reliable haug-cc-pVQZ (heavy-augmented correlation consistent valence quadruple- $\zeta$ ) basis set, the largest S66 dimers containing 36 atoms and 1520 orbitals are now within our reach.

When using such a large number of orbitals, we also have to consider the data storage bottlenecks of CCSD(T). Many CCSD and (T) implementations require the storage of at least a part of the four-external and three-external four-centre electron repulsion integral (ERI) tensors, respectively. With the above dimensions of  $n_o = 30$  and  $n_v = 1000$ , the corresponding arrays alone would require 1863 and 224 GB storage space, which is far from practical without an integral direct implementation. Consequently, various algorithms have been implemented to deal with these data size bottlenecks [28,

33–35, 37, 45, 116]. We introduced the first completely integral-direct approach eliminating the storage of all four-centre ERI tensors with at least two-external indices. Both of our CCSD and (T) algorithms affordably assemble these four-centre ERIs when needed from three-centre ERIs exploiting DF and t1-transformed expressions, making the memory and disk input/output (I/O) demand affordable even well above 1000 orbitals [37].

Considering the slow convergence of the CC correlation energies with the AO basis set, the use of F12-based CC methods is highly advantageous, at least at the CCSD level. Multiple variants of explicitly correlated CCSD approaches [52, 55, 117, 118] exhibit improved accuracy per cost performance compared to conventional CCSD [57]. Recently, we also combined the CCSD(F12\*) formulation of Hättig and co-workers [55, 117] with our efficient DF-CCSD(T) algorithm discussed above [62], although its data-storage and parallel scaling capabilities are not yet fully optimised. Nevertheless, dozens of CCSD(F12\*)(T+) computations presented here with the haug-cc-pVTZ-F12 basis [119] exceeded the 1000 AO mark, illustrating that the implementation is already highly capable. Since the latest DF- and parallelisation-based advancements are not yet available for F12-based CCSD codes, except for the CCSD(2)<sub>F12</sub> implementation of Valeev et al. [33], our S66 computations with 1184 AOs are close to the capability limits of current F12-based CCSD codes without additional approximations.

Unfortunately, current F12-based CC methods, in practise, cannot directly accelerate the basis set convergence of the (T) contribution. A frequently employed heuristic approach called (T\*) scales the conventional, that is, F12 contribution-free (T) correction, ( $E^{(T)}$ ) by the ratio of the MP2-F12 and MP2 correlation energies ( $E^{\text{MP2-F12}}$  and  $E^{\text{MP2}}$ , respectively) [52] as

$$E^{(T*)} = \frac{E^{\text{MP2-F12}}}{E^{\text{MP2}}} E^{(T)}, \quad (1)$$

which is, however, not size-consistent. Inspired by the practicality of (T\*), we recently proposed a similar scaling approach denoted by (T+), which is, however, size-consistent [62]. In order to introduce the (T+) approach, let us rewrite the CCSD(T) correlation energy as

$$E^{\text{CCSD(T)}} = \sum_i \delta E_i^{\text{CCSD}} + \sum_i \delta E_i^{(T)}, \quad (2)$$

where one of the occupied summation indices,  $i$  was separated to define orbital-specific correlation energy contributions. Then, we utilise an analogous reformulation of the MP2 and MP2-F12 energy expressions, leading to their orbital-specific correlation energy contributions,  $\delta E_i^{\text{MP2}}$  and  $\delta E_i^{\text{MP2-F12}}$ , respectively. Finally, one can show

that the (T+) energy expression of

$$E^{(T+)} = \sum_i \frac{\delta E_i^{\text{MP2-F12}}}{\delta E_i^{\text{MP2}}} \delta E_i^{(T)}, \quad (3)$$

is size-consistent and also improves upon the basis set convergence of the (T) energy contribution [62]. Sometimes a size consistent (sc) variants of (T\*), that is ( $T^*_{\text{sc}}$ ) and ( $T_{\text{sc}}$ ), are also utilised, where the  $\frac{E^{\text{MP2-F12}}}{E^{\text{MP2}}}$  or the analogous  $\frac{E^{\text{CCSD(F12*)}}}{E^{\text{CCSD}}}$  scaling factors obtained for the dimer are also used for all monomers [25, 26].

## 2.2. Local natural orbital CC methods

It is clear from the above analysis that it is very challenging to extend the limits of the seventh-power-scaling CCSD(T) with parallelisation and F12 methods alone. Therefore, the aim of the numerous approximate local correlation approaches is to accelerate CC methods while retaining their reliability. It is common to consider orbital or orbital pair specific correlation energy contributions, such as in Equation (2), in order to introduce orbital (pair) specific approximations [36, 71, 78, 81, 87, 89, 91, 120, 121].

Let us briefly summarise the main ideas of the LNO family of local correlation approaches [36, 90–92, 122–126] in order to provide some background for the introduction of a methodological development proposed here. All of our local correlation methods, starting from the MP2 level through CCSD(T) and CCSDT(Q) up to general order CC, employ orbital specific approximations utilising the following energy expression:

$$E^{\text{LNO-M}} = \sum_I \left[ \delta E_I^{\text{LNO-M}} + \Delta E_I^{\text{MP2}} + \frac{1}{2} \sum_J^{\text{distant pair}} \delta E_{IJ}^{\text{MP2}} \right]. \quad (4)$$

Here,  $I$  and  $J$  denote localised MOs (LMOs) and all LMO pairs are assigned to either the strongly correlated or the distant orbital pair group. The individually small distant pair contributions,  $\delta E_{IJ}^{\text{MP2}}$ , are accounted for by an efficient, approximate MP2-level expression [124]. Then domain approximations are employed for the dominant, strong pair contributions. For each LMO  $I$ , an orbital domain, that is a set of atomic orbitals, strong pair LMOs, domain specific virtual MOs, etc., is defined to evaluate their correlation energy contribution at the MP2 level [36, 91]. In a subsequent step, domain- or LMO-specific LNOs are constructed to compress the orbital spaces of the domain and to evaluate the  $\delta E_I^{\text{LNO-M}}$  correlation energy contribution of the domain at the  $M = \text{CCSD(T)}$ ,



CCSDT(Q), etc. level of theory. Finally, the LNO compression approximation is corrected at the MP2 level in each domain via a  $\Delta E_I^{\text{MP2}} = \delta E_I^{\text{domain MP2}} - \delta E_I^{\text{LNO-MP2}}$  correction.

As discussed before in detail [36, 91, 124, 125], this procedure involves a number of different approximations, for which we introduced a hierarchical set of composite thresholds: Loose, Normal, Tight, very Tight, etc. [91]. Note that the composite thresholds combine a set of settings which individually form a systematically convergent series for each approximation. Although the systematic convergence property is not necessarily inherited by the composite thresholds in all cases, in practise, they usually still provide a monotonically convergent energy series which is also suitable for extrapolation toward the local approximation free correlation energy [91]. To that end, we proposed to use a rather cautious extrapolation expression assuming only that monotonic convergence occurs in the threshold series [91]. Namely, an energy estimate is formed from the two tightest available results supposing that the subsequent step in the series will be smaller than the last. For instance, the extrapolation using Normal and Tight settings will give a Normal-Tight (N-T) result of

$$E^{\text{N-T}} = E^{\text{Tight}} + (E^{\text{Tight}} - E^{\text{Normal}})/2 \pm (E^{\text{Tight}} - E^{\text{Normal}})/2, \quad (5)$$

where an LNO error estimate is also defined via the  $\pm(E^{\text{Tight}} - E^{\text{Normal}})/2$  term [91]. We note that the expression in Equation (5) can be identically reformulated as  $E^{\text{N-T}} = E^{\text{Normal}} + 1.5(E^{\text{Tight}} - E^{\text{Normal}})$ , which formula was recently adopted by Bistoni et al. [101] to extrapolate DLPNO-CCSD(T) results in a very similar manner.

We have successfully employed this strategy to obtain systematically improved LNO-CCSD(T) interaction energies for a number of studies. For instance, complexes of alkali and alkaline earth metal ions representative of their non-covalent interaction in protein environments were investigated up to the very Tight convergence level in LNO-CCSD(T) combined with augmented quintuple- $\zeta$  basis sets [107–109]. Reaching this level of accuracy with a few tenths of a kcal/mol uncertainties compared to CCSD(T)/CBS in both the local correlation and the basis set error estimates has become relatively affordable for systems including 50–100 atoms. Our recent algorithmic developments also made well converged local CCSD(T)/CBS estimates accessible up to more than 100 atoms even for challenging complexes with extended, delocalised  $\pi$ -electron systems [27, 91]. Although with a considerable computational effort, even

very Tight and aug-cc-pV5Z level LNO-CCSD(T) computations were feasible for molecular complexes of up to 130 atoms, including the L7 set of complexes and an extensive fullerene complex [18, 27]. The challenging complexes of that study highlight the technical difficulties occurring with unusually extended, delocalised  $\pi$ -systems. Namely, their slowly decaying long-range interactions lead to a large number of important orbital pair interactions and thus to a considerable decrease in the efficiency of any local correlation approach [27, 89, 106]. Finally, in the cases free of large  $\pi$ -systems, such as protein-ligand complexes, our LNO-CCSD(T) implementation scales up to 1000 atoms [91].

This extensive exploration of intermolecular interactions revealed that the previous LNO algorithms did not provide completely balanced error measures for systems with and without ghost AOs used for the counterpoise (CP) correction of the basis set superposition error (BSSE). As expected, the BSSE is even more problematic for large molecular complexes with extensive diffuse AO basis sets than for the well-documented case of small-to medium sized complexes. The source of this unbalanced accuracy is in the construction of the LMO-specific orbital domains. Here, first those virtual MOs are gathered into the domains which are expected to be important for the correlation of the domain's LMOs. These virtual MOs reside on atoms whose AOs contribute considerably to the expansion of the domain's LMO. As pointed out in our previous work [36], a drawback of this initial virtual MO list surfaces if there are ghost atoms in the system holding AOs but not nuclei or electrons. Then, there could be some ghost atoms where the expansion coefficients of all LMOs are especially small and the virtual MOs residing on these ghost atoms are not added to the initial virtual MO list of the LMO domain.

To overcome this, we extended the initial virtual MO definition via a second virtual MO selection criterion [36]. In brief, we identify potentially important ghost atoms not only by the size of the LMO coefficients but also using a three-centre AO ERI estimate. First, a fixed threshold was determined for identifying the additional set of domain atoms and the corresponding orbitals, which was suitable for Normal threshold settings and up to medium-sized molecular complexes. In the present work, we update this selection by making it LNO threshold and target accuracy dependent as well as somewhat tighter than before. In this way, the accuracy is increased for large complexes with many ghost atoms and long-range interactions with a moderate additional computational effort. For the sake of reproducibility, the updated algorithm extending the one explained in Ref. [36] employs the  $(1 - T_{\text{EDO}})10^4 E_h$  threshold for the screening integrals. Thus, this threshold is now tied

to the LMO completeness threshold  $T_{\text{EDo}}$ . For instance, the new threshold amounts to 1, 0.5, and 0.1  $E_h$  with Normal, Tight, and very Tight settings, instead of the previous value of 2  $E_h$  used irrespective of the target accuracy. Let us note that this update is turned on by default in the latest, 2022 release of MRCC for the default `localcc=2021` set of options [127, 128]. Therefore, here we provide updated LNO-CCSD(T) benchmarks with these settings in Section 4.3.

### 3. Computational details

All new DF-based conventional [37], explicitly correlated [62], and LNO-based local [36, 91] CCSD(T) computations have been performed using the 2022 release of the MRCC quantum chemistry programme suite [127, 128]. For the explicitly correlated CCSD(T) computations the CCSD(F12\*) approach of Hättig and co-workers [55] is combined with our (T+) scaled-triples corrections [62]. The core electrons were frozen in all correlation calculations.

Correlation consistent basis sets augmented with diffuse functions only for the non-hydrogen, i.e. heavy atoms, that is, haug-cc-pVXZ [39, 129] ( $X = D, T, Q, 5$ ) were employed, which are abbreviated throughout the manuscript as haXZ. As the AO basis set of the CCSD(F12\*)(T+) computations, the correlation consistent  $X$ -tuple- $\zeta$  (haug-)cc-pVXZ-F12 ( $X = D, T, Q$ ), or haXZ-F12 in brief, basis sets developed for explicitly correlated calculations by Peterson *et al.* [130] were employed together with the corresponding cc-pVXZ-F12-OPTRI complementary auxiliary basis (CABS) bases of Yousaf and Peterson [131, 132]. The diffuse functions extending the original cc-pVXZ-F12 bases were adopted from the work of Martin and co-workers [119]. Following the recommendation of Shaw and Hill, their cc-pVDZ-F12-OPTRI+ basis set was employed instead of the original ‘OPTRI’ variant at the double- $\zeta$  level [133]. All Hartree–Fock (HF) IEs corresponding to the CCSD(F12\*)(T+) results contain the CABS correction (HF+CABS).

All of our computations employ the DF approximation with the corresponding DF auxiliary bases for the HF [134] and correlation [135] parts, that is, haug-cc-pVXZ-RI-JK and haug-cc-pVXZ-RI were employed with the haug-cc-pVXZ AO basis set. In the explicitly correlated calculations, the fitting bases were the aug-cc-pV( $X+1$ )Z-RI-JK sets of Weigend [134] and the aug-cc-pwCV( $X+1$ )Z-RI bases of Hättig [136], respectively. A Slater-type  $f_{12}$  correlation factor was used with exponents of 0.9, 1.0, and 1.0 with the (haug-)cc-pVDZ-F12, (haug-)cc-pVTZ-F12, and (haug-)cc-pVQZ-F12 basis sets, respectively [130].

Results utilising (ha)XZ(-F12) and (ha)( $X+1$ )Z(-F12) basis sets as well as complete basis set (CBS) extrapolation will be denoted by (ha)( $X,X+1$ )Z(-F12). For the extrapolation of HF energies, the two-point formula suggested by Karton and Martin is used with the recommended parameters [137, 138]. Conventional, LNO-, and DLPNO-CCSD(T) correlation energies were extrapolated employing the formula of Helgaker and co-workers with the exponent of 3 [40]. CCSD(F12\*) correlation energies using the (ha)XZ-F12 and (ha)( $X+1$ )Z-F12 basis sets are extrapolated via the formula of Ref. [40] with the exponents of 3.1458 for  $X = D$ , as well as 4.3448 and 4.6324 for  $X = T$  without and with the diffuse AOs, respectively, following the recommendation of Martin and co-workers [26].

The structures of the S66 compilation were taken from the original work of Řezáč, Riley, and Hobza [17]. DF approximation free CCSD(T)/haDZ, CCSD(T)/haTZ, HF+CABS/aQZ-F12, and MP2-F12/a(T,Q)Z-F12 energies were adopted from Ref. [26], while our previous DF-CCSD(T)/haTZ references were taken from Ref. [36]. DLPNO-CCSD(T0) and DLPNO-CCSD(T1) results computed with [100] and without [87] relying on the semi-canonical triples approximation (T0) were computed using the 4.1 version of the ORCA package [139].

To characterise the performance of the methods, the mean signed error (MSE), the mean absolute error (MAE), the root mean square deviation (RMSD), the standard deviation (STD), and the maximum absolute (MAX) error of the computed results will be applied.

Regarding the computational cost, obtaining all DF-CCSD(T)/haXZ ( $X = T, Q$ ) and the CCSD(F12\*)(T+)/haXZ-F12 ( $X = D, T$ ) references for the entire S66 set, including CP corrections, required the execution of about 30 exaFLOP that is,  $3 \cdot 10^{19}$  operations. Using a current many-core CPU of about 1 teraFLOP performance, this amounts to about 350,000 core hours or about a node year effort, which is rather affordable compared to the high-performance computing grants and resources available today.

## 4. Results and discussion

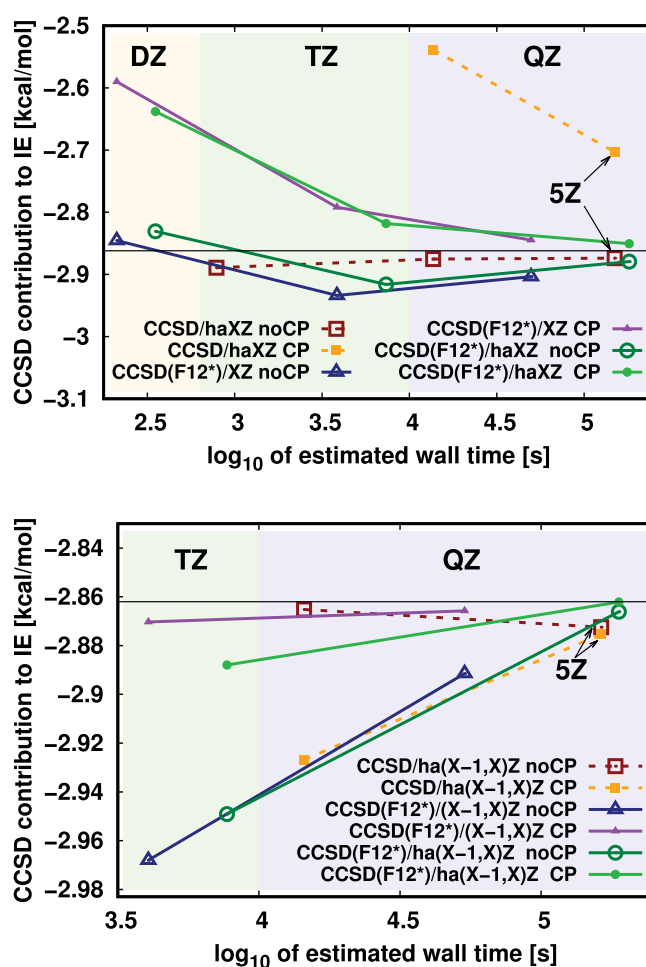
### 4.1. Conventional and F12-based CCSD(T): convergence to the CBS limit

Let us start with an analysis of one characteristic example, illustrating the convergence of the CCSD(T) correlation energies very close to the CBS limit, within cca. 1% or 0.01 kcal/mol remaining basis set uncertainty. Here, we select the dimer No. 20 of S66, the acetic acid dimer, which is ideal for this illustration for a number of reasons.

The dimer contains 16 atoms, which is around the size limit accessible with basis sets one cardinal number higher than our reference computations. That is, highly demanding but very well converged CCSD(T)/ha5Z and CCSD(F12\*)(T+)/haQZ-F12 computations are within the limit of feasibility (1456 and 1328 AOs, respectively). Additionally, the (AcOH)<sub>2</sub> IE exhibits one of the slowest convergence to the CBS limit of CCSD(T) and exhibits a double H-bond motif that may be challenging for certain DMC algorithms [11].

The most common approach to analyse the performance of CCSD(T)/XZ and CCSD(F12\*)(XZ)-F12 is to compare their convergence with respect to the cardinal number  $X$ . Here, we also take into account the computational cost in the analysis by considering the about one quarter to one third higher demand of CCSD(F12\*) with respect to CCSD with the same basis set, and more importantly, the marked size difference of the (ha)XZ and (ha)XZ-F12 basis sets even with the same  $X$  value. In other words, we are interested in the accuracy achieved compared to the invested number of operations or computer time taking into consideration that CCSD(F12\*)(T)/XZ-F12 computations are more demanding than CCSD(T)/XZ with the same  $X$ . To that end, Figures 1 and 2 plot the CCSD and (T) correlation energy contributions, respectively, to the IE of (AcOH)<sub>2</sub> as the function of the operation count requirement. For the sake of simplicity, we transform the operation counts to computational time estimates assuming a 2 teraFLOPS, that is  $10^{12}$  FLOP per second, (cca. 40 core) CPU and 50% peak performance utilisation, which efficiency is representative of our CCSD(T) code in MRCC [37]. The time estimates on the horizontal axes take into account the number of operations required for the rate determining fifth- and sixth-power-scaling terms for CCSD and the sixth- and seventh-power-scaling terms for (T). For instance, the large, ha5Z computations requiring about  $10^{5.2}$  and  $10^{5.1}$  teraFLOP for the CCSD and (T) parts, respectively, can be translated to about 1.9 and 1.4 days of wall time at a useful performance of 1 teraFLOPS, which is in accord with our measurements. Finally, in Figures 1 and 2, we also take into consideration the cost of CBS extrapolation, that is, the operation counts for the  $(X-1)Z$  and XZ computations, which are required to obtain the CBS( $X-1,X$ ) extrapolated values, are added together on the x axes. Consequently, the CBS( $X-1,X$ ) extrapolated results in the second panels of Figures 1 and 2 are slightly more demanding than the corresponding XZ results. The plotted data can be found in Tables S1 and S2 of the supplementary material (SM).

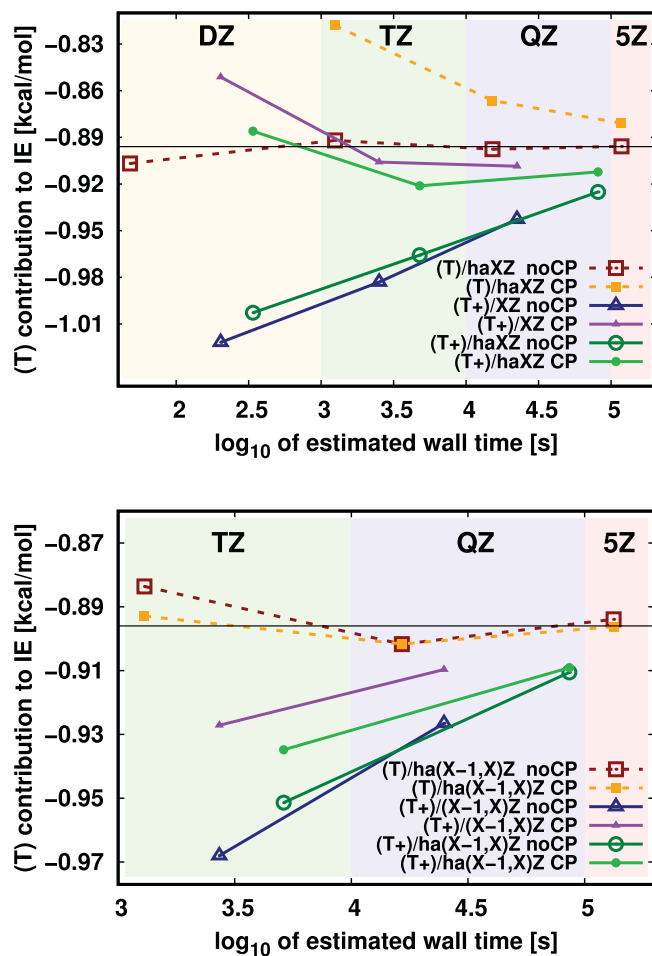
The CCSD and (T) IE contributions are plotted for CCSD(T)/haXZ, CCSD(F12\*)(T+)/XZ-F12, and CCSD(F12\*)(T+)/haXZ-F12 in Figures 1 and 2, where



**Figure 1.** CCSD correlation energy contribution to the (AcOH)<sub>2</sub> IE with respect to the required computational time. Top panel: CCSD/haXZ ( $X = D, T, Q, 5$ ) and CCSD(F12\*)(ha)XZ-F12 ( $X = D, T, Q$ ) results are collected without (label noCP) and with CP correction. Bottom panel: CBS( $X-1,X$ ) ( $X = T, Q, 5$ ) extrapolated results corresponding to the IEs of the top panel. The x axis values denote the  $\log_{10}$  of the estimated wall time [in seconds] corresponding to the required operation count and the assumption of a 2 teraFLOPS CPU with 50% peak performance utilisation. CCSD/haDZ noCP and CP, as well as the CCSD/haTZ CP results of  $-2.2$ ,  $-0.7$ , and  $-2.0$  kcal/mol, respectively, are not plotted for the sake of transparency.

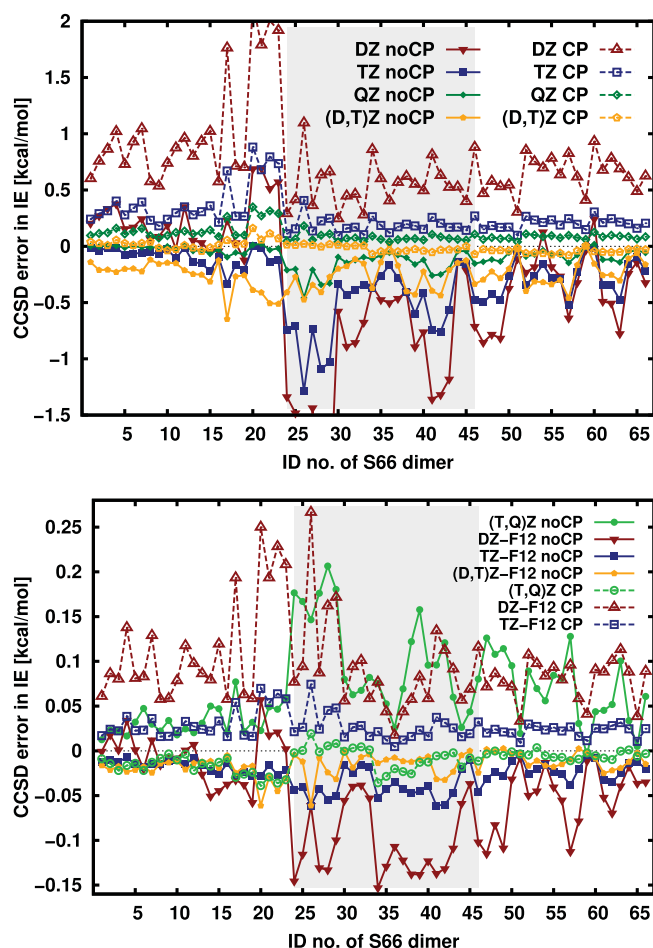
coloured vertical blocks collect results obtained with the same  $X$  cardinal number. Considering first the conventional CCSD results (red ‘CCSD noCP’ and orange ‘CCSD CP’ curves of Figure 1 with square symbols), we find a more rapid convergence for the raw IE contributions than with CP corrections. This pattern is not the most common case for IE computations as CP corrections usually improve the results, but this trend is characteristic of the first, hydrogen bonding dominated third of the S66 compilation (see Figure 3 below). While the CCSD/haDZ and CP corrected CCSD/haTZ results did not fit into Figure 1, the raw and the





**Figure 2.** (T) correlation energy contribution to the (AcOH)<sub>2</sub> IE with respect to the required computational time. Top panel: (T)/haXZ ( $X = D, T, Q, 5$ ) and (T+)/haXZ-F12 ( $X = D, T, Q$ ) results are collected without (label noCP) and with CP correction. Bottom panel: CBS( $X-1, X$ ) ( $X = T, Q, 5$ ) extrapolated results corresponding to the IEs of the top panel. The x axis values denote the log<sub>10</sub> of the estimated wall time [in seconds] corresponding to the required operation count and the assumption of a 2 teraFLOPS CPU with 50% peak performance utilisation. The (T)/haDZ result of  $-0.686$  kcal/mol value is omitted for transparency.

CP corrected CCSD results tend to about  $-2.872$  and  $-2.875$  kcal/mol, respectively, at the CCSD/ha(Q,5)Z level. Compared to that, the CCSD(F12\*) results, both with XZ-F12 (triangles) and haXZ-F12 (circles), and both without (blue and purple) and with CP (dark and light green) are already fairly accurate at the (ha)DZ-F12 level. However, considering the convergence pattern of the F12-based IEs, the good (ha)DZ-F12 performance is partly the result of error compensation. Nevertheless, (CP corrected) CCSD(F12\*)/ha(D,T)Z-F12 or better results can be considered converged and accurate within 1%. The agreement of raw and CP corrected CCSD(F12\*)/(T,Q)Z-F12 within 0.025 kcal/mol



**Figure 3.** CCSD IE deviations (correlation energy part) with respect to the CP corrected CCSD(F12\*)/ha(D,T)Z-F12 reference individually for the 66 dimers of the S66 set. Top panel: CCSD/haXZ ( $X = D, T, Q$ ) and CCSD/ha(D,T)Z results, bottom panel: CCSD/ha(T,Q)Z and the remaining CCSD/(F12\*)haXZ-F12 results, all without (noCP) and with CP correction. Dimers are bound mostly by dispersion (24–46) in the gray area, H-bonds (left) and mixed interactions (right) are collected separately.

and of raw and CP corrected CCSD(F12\*)/ha(T,Q)Z-F12 within 0.004 kcal/mol is again convincing. Additionally, the CCSD(F12\*)/ha(T,Q)Z-F12 results without and with CP of  $-2.866$  and  $-2.862$  kcal/mol agrees with the CCSD/ha(Q,5)Z results within 0.01 kcal/mol or 0.34% remaining uncertainty.

Considering the computational cost, it increases by about an order of magnitude with one step increase in the cardinal number. A much less emphasised aspect is that CCSD(F12\*)/XZ-F12 and CCSD(F12\*)/haXZ-F12 computations are about 3–5 and 10–12 times more demanding than CCSD/haXZ in this case. Since CCSD(F12\*)/(h)aXZ-F12 could be an order of magnitude more costly than CCSD/XZ, it is not fully informative when these kinds of cost analyses are simplified to only the  $X$  cardinal number value. Nevertheless, the

additional costs of CCSD(F12<sup>\*</sup>)/XZ-F12 computations certainly return, whereas it may depend on the particular application and target accuracy if the investment of the twice more operations with the additional diffuse functions in haXZ-F12 is worthwhile.

Continuing with the (T) contributions in Figure 2, the CP corrected (orange) and uncorrected (red) (T)/XZ results (square point type) behave similarly to the CCSD ones: the raw curve is fairly flat while the CP corrected results converge slower but in a satisfactory way. The two sets of results both approach  $-0.895$  kcal/mol within  $0.001$  kcal/mol at the (T)/ha(Q,5)Z level, but their agreement is already excellent at the ha(D,T)Z and ha(T,Q)Z levels. However, we note that the CP corrected (T)/haDZ result of  $-0.686$  kcal/mol (not plotted), exhibiting 23% error, does not appear to be highly reliable for use in extrapolation. In accord with numerical experience in the literature, the triples corrections using the CCSD(F12<sup>\*</sup>) amplitudes, even the scaled (T+) variant showed here, converge more slowly to the CBS limit. Since the basis set incompleteness error of the triples corrections is not decreased directly via the explicitly correlated treatment, the BSSE leads to overbinding, especially with the raw (T+) results. The CP correction uniformly improves the (T+) results, which are again within  $0.01$ – $0.02$  kcal/mol of the (T)/ha(Q,5)Z values starting from the (T+)/(ha)TZ-F12 level. Compared to the case of CCSD, the cost increase when using XZ-F12 and haXZ-F12 bases instead of XZ is milder as only about 1.5 to 3 times more operations are required for the (T) part with the (ha)XZ-F12 bases. Therefore, (ha)TZ-F12 level (T+) results, especially with CP corrections, are competitive, but (T)/ha(T,Q)Z appears to be somewhat more robust in this respect.

The comparison of the CCSD and (T) trends in Figures 1 and 2 is also instructive. Since the (T) term does not benefit directly from the F12 treatment, at least for this example, it is similarly hard to converge the CCSD and (T) IE contributions to close to  $0.01$  kcal/mol accuracy. The comparison of the CCSD and (T) expenses also show a trend that could be unexpected considering only the higher-order scaling of (T) and the corresponding common assumption of its higher cost. However, here, we find the opposite cost relation, partly caused by the 12 CCSD iterations, only 24 correlated electrons, and the better parallel scaling of (T). Namely, using the same basis set, the (T) term requires more operations than CCSD only with haDZ and haTZ; with the other basis sets, the CCSD or CCSD(F12<sup>\*</sup>) part is more costly than (T) by 10–220%. This adds an interesting perspective to the previous best reference of Ref. [26] labelled SILVER, where CCSD(F12<sup>\*</sup>)/aTZ-F12 was combined with (T)/ha(D,T)Z. Considering the required operations and

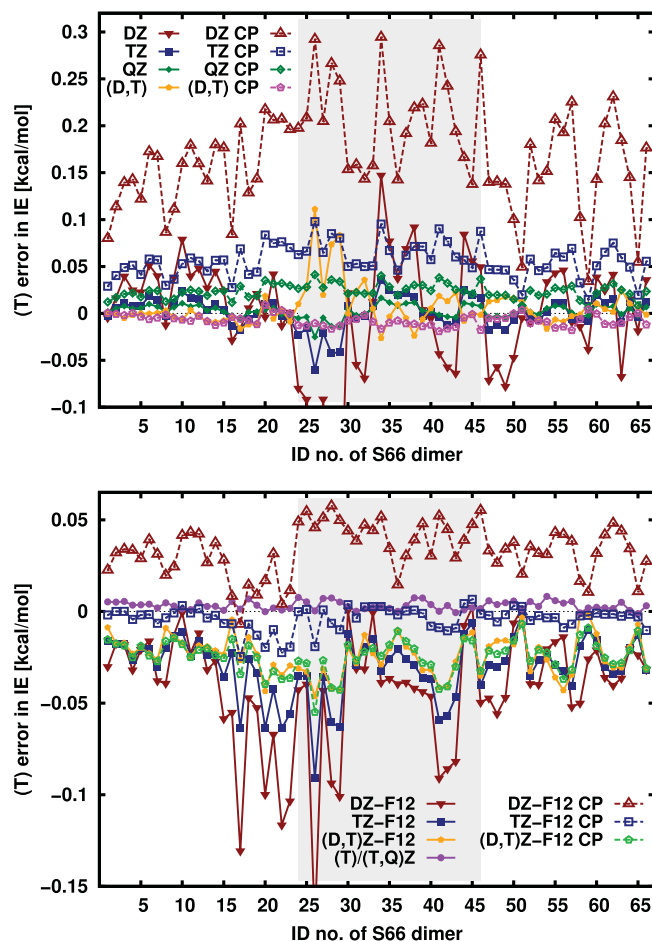
not just the basis set cardinal numbers, it is interesting to see that obtaining (T)/haQZ for (AcOH)<sub>2</sub> is only about 40% more demanding than CCSD(F12<sup>\*</sup>)/aTZ. While the extra cost of (T) is higher with more correlated electrons, still, this analysis uncovers an opportunity to improve upon the previous best S66 reference with an affordable computational effort. Naturally, the (T)/haQZ over CCSD(F12<sup>\*</sup>)/aTZ cost ratio is higher for the dimers with more electrons but remains lower than 3.4 even for the largest uracil dimers.

#### 4.2. Conventional and F12-based CCSD(T): new reference for the S66 set

Utilizing our recent, well-parallelised, and integral direct codes [37, 62] we obtained DF-based CCSD(T) results up to ha(T,Q)Z and CCSD(F12<sup>\*</sup>)(T+) results up to ha(D,T)Z-F12 for the entire S66 list. Considering the numerical experience of related studies for smaller or even S66-sized dimers [17, 25, 26, 42–44, 58–61], it is still not obvious which set of benchmarks is better from the following perspectives. For instance, the trends noted for (AcOH)<sub>2</sub> are not in accord with the common recommendations of using CP corrected or half-CP corrected results. The application of CBS extrapolation should also be investigated carefully, both for CCSD(F12<sup>\*</sup>) and especially for (T+). It is not obvious which set of results are more reliable: the CCSD/ha(T,Q)Z or the CCSD(F12<sup>\*</sup>)/ha(D,T)Z-F12 results and the (T)/ha(T,Q)Z or the scaled (T+)/(ha(D,T)Z-F12 results. Finally, it is unclear, whether the current standard of small/medium basis CCSD(T) and MP2-based CBS correction type composite approaches should be considered here, e.g. due to the well-documented overestimation of MP2-based  $\pi$ - $\pi$  stacking IEs.

First, one may look at the difference of the CP corrected and uncorrected results as a measure of the basis set convergence. This is a helpful indicator, but cannot determine alone whether the raw, CP, or half-CP results are the most accurate. Therefore, we also analysed the basis set convergence patterns of these possibilities. However, we could not find a clear connection between the absolute accuracy and measures of the level of convergence, such as the size of the last step taken between, e.g. the haDZ-F12 and haTZ-F12 results. The reason behind this is illustrated by the smaller steps of CCSD(F12<sup>\*</sup>)/XZ-F12 compared to CCSD(F12<sup>\*</sup>)/haXZ-F12 in Figure 1. Therefore, a somewhat more complicated route is chosen, that is, we looked at the convergence trends for all S66 dimers individually in order to find the more reliable reference.

The correlation energy contributions of the CCSD and (T) results are collected in Figures 3 and 4, respectively.



**Figure 4.** (T) IE deviations (correlation energy part) with respect to the CP corrected (T)/ha(T,Q)Z reference. Top panel: (T)/haXZ ( $X = D, T, Q$ ) and (T)/ha(D,T)Z results, bottom panel: raw (T)/ha(T,Q)Z and all (T+)/haXZ-F12 results, all without and with CP correction, the latter distinguished using a ‘CP’ label. Dimers are bound mostly by dispersion (24–46) in the gray area, H-bonds (left) and mixed interactions (right) are collected separately.

The dispersion dominated dimers are separated via gray shade in the middle from the hydrogen-bonded (left) and mixed (right) dimers. Conventional and mostly F12-based CCSD deviations from the CP corrected CCSD(F12<sup>\*</sup>)/ha(D,T)Z-F12 reference are gathered on the top and bottom panels of Figure 3, respectively. The first group of dimers with H-bonding character on the left, including (AcOH)<sub>2</sub> highlighted in Section 4.1, exhibits different trends compared to the dispersion dominated (middle) and mixed (right) dimers. Interestingly, the raw results (solid symbols) overbind only above the haDZ level, while the raw CCSD/haDZ IEs are underestimated, which do not appear to be dominated by BSSE. For the dispersion and mixed groups, CCSD/haXZ IEs with  $X = D, T, Q$  exhibit a satisfactory trend much more suitable for CBS extrapolation. In contrast to that, the CP results (hollow symbols) are overcorrected as usual but

exhibit a more reliable convergence pattern for the entire S66 set and, with a few exceptions, smaller errors than their raw counterparts.

Consequently, without CP, CCSD/ha(D,T)Z results (light orange pentagons) are often even worse than or comparable to raw CCSD/haTZ. However, with CP, the CCSD/ha(D,T)Z values are about 2–3 times better than even the CP corrected CCSD/haQZ (dark green diamonds). The raw and CP corrected CCSD/ha(T,Q)Z errors (green circles on the bottom panel) are more comparable to the performance of CCSD(F12<sup>\*</sup>). The ha(T,Q)Z extrapolation overshoots in both cases, but much less with CP than without it, thus a half-CP type averaging of raw and CP results is not helpful for CCSD/ha(T,Q)Z. The CCSD(F12<sup>\*</sup>)/haTZ-F12 calculations, both with and without CP (blue squares on the bottom panel), are more efficient and accurate than CCSD/ha(T,Q)Z, while CCSD(F12<sup>\*</sup>)/haDZ-F12 is comparable to CCSD/haQZ at a much smaller cost. Except for a few H-bonded dimers, the raw and CP results approach the reference consistently from below and above, respectively. The outliers again limit the accuracy of the raw and half-CP variants of CCSD(F12<sup>\*</sup>)/ha(D,T)Z-F12 due to their dependence on raw CCSD(F12<sup>\*</sup>)/haDZ-F12. In contrast, the raw and CP corrected CCSD(F12<sup>\*</sup>)/haTZ-F12 results are arranged fairly symmetrically on the two sides of our reference. Therefore, the half-CP CCSD(F12<sup>\*</sup>)/haTZ-F12 combination, which lacks only the diffuse AOs on H atoms with respect to the half-CP CCSD(F12<sup>\*</sup>)/aTZ-F12 reference choice of Ref. [26], is within 0.01 (0.02) kcal/mol RMSD (MAX error) of the CP CCSD(F12<sup>\*</sup>)/ha(D,T)Z-F12 reference selected in the present work. Our reference choice is supported by the much more homogeneous convergence trend of CCSD(F12<sup>\*</sup>)/haDZ-F12 and CCSD(F12<sup>\*</sup>)/haTZ-F12 with CP correction than without CP and thus ensures that the less well-behaved raw CCSD(F12<sup>\*</sup>)/haDZ-F12 results are not employed in the final reference.

The case of the (T) corrections, collected in Figure 4, is different since (i) the (T) contribution to the IE is 3–5 times smaller than that of CCSD, (ii) the (T) term does not directly benefit from the F12 treatment, but (iii) it is often assumed that (T) converges to its CBS limit faster than CCSD. Indeed, compared to the CP (T)/ha(T,Q)Z reference, the ratio of CCSD and (T) IE contributions and their deviations from the reference appear to be similar, leading to considerably smaller absolute errors in the (T) term. Similar to the case of CCSD, raw (T) errors with haDZ (red triangles) and haTZ (blue squares) have much less homogeneous behaviour than their CP corrected counterparts. However, for the (T) corrections, the dispersion dominated complexes are the outliers with larger negative errors at the haDZ and haTZ levels. Even though

the CP corrected haXZ ( $X = D, T, Q$ ) errors are, on the average, larger than the corresponding uncorrected ones, their homogeneous sign and convergence trend is much more suitable for extrapolation. Since mostly both the raw and the CP corrected results underestimate the reference (without CBS extrapolation), the use of half-CP is not advised. Considering the extrapolated values, the raw and CP corrected (T)/ha(D,T)Z results (light orange and magenta pentagons) agree relatively well, except for the most complicated  $\pi$ - $\pi$  stacking dimers containing benzene, pyridine, and uracil monomers (no. 24–29). Thus, especially for the dispersion dominated complexes in the gray shaded area, the use of the half-CP (T)/ha(D,T)Z reference of the previous best S66 benchmark [26] is not the best choice in the light of our new reference numbers.

While CP (T)/ha(D,T)Z is fairly convincing with only about twice as large errors as raw (T)/ha(T,Q)Z (purple circles of the bottom panel), the homogeneous convergence pattern of the CP corrected results suggests that CP (T)/ha(T,Q)Z is the best choice of these three options. Nevertheless, the average absolute deviation of raw (T)/ha(T,Q)Z from the CP (T)/ha(T,Q)Z reference is 0.003 kcal/mol, thus their average, that is, half-CP (T)/ha(T,Q)Z could be just as good for most purposes. Since the (T+)/haXZ-F12 results are found to be better than the unscaled (T)/haXZ-F12 ones [both obtained with CCSD(F12\*) amplitudes], here, we focus on the performance of (T+) and report the corresponding (T)/haXZ-F12 results in Figure S1 of the SM. Interestingly, the (T+) results, especially with CP, are fairly accurate even with the smallest haDZ-F12 basis set (red triangles of the bottom panel). The raw (T+)/haDZ-F12 and (T+)/haTZ-F12 results are surprisingly close to each other, while CP (T+)/haTZ-F12 offers substantial improvement over CP (T+)/haDZ-F12. In fact, at least compared to the CP (T)/ha(T,Q)Z reference, CP

(T+)/haTZ-F12 is the best choice if haQZ level results are not affordable. Let us also note the remarkable agreement of the raw and CP corrected (T+)/ha(D,T)Z-F12 results (light orange and green pentagons) with each other, but not as much with the CP (T)/ha(T,Q)Z reference. However, raw and CP (T)/ha(T,Q)Z agree with each other twice as well and appear to be more reliable. Moreover, we know much less about the behaviour of the very recent (T+) method [62]. For instance, it is unclear whether the (T+) correlation energies are subjectable to CBS extrapolation and if so, what would be the suitable exponent of  $X$ . This question will be a subject of a future study. Until then, we recommend to use CP (T+)/haTZ-F12 without extrapolation over (T+)/ha(D,T)Z-F12 IEs in the explicitly correlated context.

The statistical measures (RMSD and MAX errors) for CCSD and (T) IE contributions are collected in Tables 1 and 2. Half-CP (columns CP/2 of the tables) are also added besides the raw and CP corrected results. In addition to the trends observed for the individual IEs, we analyse half-CP corrected results and MAX errors via the statistical measures. The half-CP results are found to be the best performers without CBS extrapolation, while the CP corrected values prove to be more suitable in combination with extrapolation. While we are just outside of the 0.01 kcal/mol level agreement in the RMSD of CP CCSD/ha(T,Q)Z and CP CCSD(F12\*)/ha(D,T)Z-F12, at this stage, it is not certain that CP CCSD(F12\*)/ha(D,T)Z-F12 is converged within 0.01 kcal/mol. Additionally, the MAX errors being about 2–3 times larger than the RMSD, an about 0.05 kcal/mol MAX error can be assumed for CCSD(F12\*)/ha(D,T)Z-F12 with respect to CCSD/CBS. It is interesting to compare these MAX errors to the maximum absolute difference of the corresponding raw and CP corrected results (last columns of Tables 1 and 2), the latter of which serving as an upper bound measure of

**Table 1.** RMSD and MAX errors of raw, CP, and half-CP (CP/2) CCSD compared to CP CCSD(F12\*)/ha(D,T)Z-F12 IEs as reference [in kcal/mol].

Basis set	RMSD			MAX error			raw-CP
	raw	CP	CP/2	raw	CP	CP/2	
CCSD							
haDZ	0.788	0.846	<b>0.500</b>	2.708	2.164	<b>1.426</b>	3.804
haTZ	0.411	0.307	<b>0.176</b>	1.287	0.882	<b>0.440</b>	1.693
haQZ	0.128	0.122	<b>0.061</b>	0.458	0.350	<b>0.181</b>	0.641
ha(D,T)Z	0.288	<b>0.045</b>	0.149	0.646	<b>0.160</b>	0.296	0.701
ha(T,Q)Z	0.083	<b>0.015</b>	0.039	0.206	<b>0.039</b>	0.106	0.201
CCSD(F12*)							
haDZ-F12	0.075	0.108	<b>0.045</b>	0.153	0.266	<b>0.154</b>	0.328
haTZ-F12	0.030	0.030	<b>0.009</b>	0.061	0.074	<b>0.021</b>	0.136
ha(D,T)Z-F12	0.020	ref	0.010	0.061	ref	0.031	0.061
aTZ-F12 [26]	0.027	0.031	<b>0.009</b>	0.054	0.078	<b>0.024</b>	0.131

Note: The last column collects the maximum absolute difference of raw and CP results with the corresponding basis set. The best variant among the raw, CP, and CP/2 options for a given basis set is highlighted in bold.



**Table 2.** RMSD and MAX errors of raw, CP, and half-CP (CP/2) (T) compared to CP (T)/ha(T,Q)Z IEs as reference [in kcal/mol].

Basis set	RMSD			MAX error			raw-CP
	raw	CP	CP/2	raw	CP	CP/2	
(T) (with CCSD amplitudes)							
haDZ	<b>0.079</b>	0.179	0.093	0.360	0.294	<b>0.221</b>	0.653
haTZ	<b>0.017</b>	0.059	0.031	<b>0.060</b>	0.098	0.066	0.158
haQZ	<b>0.007</b>	0.025	0.014	<b>0.025</b>	0.041	0.028	0.066
ha(D,T)Z	0.023	<b>0.010</b>	0.011	0.111	<b>0.019</b>	0.050	0.122
ha(T,Q)Z	0.004	ref	0.002	0.008	ref	0.004	0.008
(T) [with CCSD(F12*) amplitudes]							
haDZ-F12	<b>0.036</b>	0.147	0.088	<b>0.069</b>	0.309	0.150	0.318
haTZ-F12	<b>0.013</b>	0.048	0.020	0.047	0.096	<b>0.032</b>	0.143
ha(D,T)Z-F12	0.030	<b>0.008</b>	0.019	0.068	<b>0.021</b>	0.045	0.048
(T+) [with CCSD(F12*) amplitudes]							
haDZ-F12	0.054	0.036	<b>0.020</b>	0.171	<b>0.058</b>	0.069	0.217
haTZ-F12	0.035	<b>0.008</b>	0.020	0.091	<b>0.025</b>	0.055	0.072
ha(D,T)Z-F12	0.026	0.026	0.026	<b>0.046</b>	0.055	0.051	0.011

Note: The last column collects the maximum absolute difference of raw and CP results with the corresponding basis set. The best variant among the raw, CP, and CP/2 options for a given basis set is highlighted in bold.

the basis set convergence. Indeed, in Table 1 the maximum ‘raw-CP’ values are considerably (2–6 times) higher than the MAX error of the best performing CCSD variant for a given basis set. This also suggests that CP CCSD(F12\*)/ha(D,T)Z-F12 is converged considerably better than its 0.061 kcal/mol maximum ‘raw-CP’ measure. The largest ‘raw-CP’ value corresponds to (AcOH)<sub>2</sub> (closely followed by the uracil dimer and the uracil-AcOH dimer). However, we know from the analysis of Section 4.1 that CP CCSD(F12\*)/ha(D,T)Z-F12 differs from the CBS limit by about 0.003 kcal/mol, and thus the major part of the 0.061 kcal/mol ‘raw-CP’ value comes from the error of the raw number. We also note that the CCSD(F12\*)/aTZ-F12 values used as reference in Ref. [26] (last row of Table 1) closely agree with our (DF-based) CCSD(F12\*)/haTZ-F12 results, indicating that the additional *p*- and *d*-type diffuse AOs on the hydrogen atoms of aTZ-F12 contribute negligibly to the IEs. Additionally, the ‘raw-CP’ value of CCSD(F12\*)/(h)aTZ-F12 is about twice as large as that of CCSD(F12\*)/ha(D,T)Z-F12, thus up to 0.02 kcal/mol improvement can be expected from updating the S66 reference to CP CCSD(F12\*)/ha(D,T)Z-F12.

Considering the corresponding statistical measures for the (T) contributions in Table 2, the hierarchy of the raw, CP, and half-CP results is much less clear cut compared to the case of CCSD. One commonality is that full CP appears to be best of all types of CBS extrapolated results, while raw is mostly but not always better than half-CP and CP without CBS extrapolation. The behaviour of the ‘raw-CP’ measure also differs for (T): here, it is often very close to or the same as the sum of the MAX raw and MAX CP errors, while in a few cases ‘raw-CP’ does not appear to be larger than the MAX

errors [cf. the (T+)/ha(D,T)Z-F12 case]. Nevertheless, the at most 0.008 kcal/mol MAX deviation of the raw and CP corrected (T)/ha(T,Q)Z results is very convincing and suggests that our new (T) reference is accurate within cca. 0.01 kcal/mol. This estimation holds for the case of (AcOH)<sub>2</sub>, where the (T)/ha(T,Q)Z IE components are within 0.008 and 0.005 kcal/mol of the (T)/ha(Q,5)Z values with and without CP, respectively.

The best converged result sets, that is, the CP (T)/ha(T,Q)Z reference, CP (T)/ha(D,T)Z-F12, and CP (T+)/haTZ-F12 are only about twice as close to each other as the analogous differences for CCSD in Table 1, indicating that converging the (T) contribution is similarly important, already for medium-sized complexes. The half-CP (T)/ha(D,T)Z variant employed in Ref. [26], with 0.011 and 0.050 kcal/mol RMSD and MAX errors, respectively, seems to be less reliable than its CP corrected analogue.

It is also worth noting the performance of the size-consistent (T+) scaling approach compared to the previously employed (T\*) and (T\*<sub>sc</sub>) scaling alternatives. In agreement with the observations of Ref. [62], the difference of the (T+) and (T\*) results is the largest for dimers with markedly different monomers, e.g. for the complexes of benzene or pyridine formed with water, MeOH, or AcOH. This size-consistency related difference can amount up to 0.1 kcal/mol with haDZ-F12 and remains too large for our purposes (up to 0.06 and 0.04 kcal/mol) even at the haTZ-F12 and ha(D,T)Z-F12 levels. Even if, for the dimers with markedly different monomers, the size-consistent (T\*<sub>sc</sub>) version is considerably better than (T\*), when all complexes are considered, the maximum deviations of (T+) and (T\*<sub>sc</sub>) are even higher than those of (T+) and (T\*), namely 0.2, 0.05, and 0.05 kcal/mol

at the haDZ-F12, haTZ-F12, and ha(D,T)Z-F12 levels, respectively. The worst performers for  $(T^*_{sc})$  are the  $\pi$ - $\pi$  stacking dimers containing benzene, pyridine, and uracil monomers (no. 24–29), where the MP2-based scaling factors obtained for the dimers are not adequate for the monomers. The performance of  $(T^*)$  and  $(T^*_{sc})$  found here is in accord with the findings of Ref. [25] for the S66x8 set, where such high size-consistency errors were also characterised as unacceptable.

In summary, we suggest combining CP corrected CCSD(F12<sup>\*</sup>)/ha(D,T)Z-F12 and  $(T)/ha(T,Q)Z$  to update the CCSD(T) reference IEs for the S66 set. This corresponds to the level of theory for the  $(T)$  term labelled as GOLD in the previous best converged S66 benchmark of Ref. [26]. The CCSD IEs employed here should be more converged than the SILVER level of Ref. [26], but we cannot test if their accuracy reaches that of the GOLD level of CCSD(F12<sup>\*</sup>)/cc-pVQZ-F12 suggested in Ref. [26]. Let us note in passing, that the largest CCSD(F12<sup>\*</sup>)/cc-pVQZ-F12 computation for S66 would involve 1664 AOs, which should be feasible in the near future upon some memory optimisation due in our CCSD(F12<sup>\*</sup>) code. Regarding the mean-field contribution, the half-CP HF+CABS/aQZ-F12 results of Ref. [26] are highly converged and suitable for combination with our new correlation energy contributions.

One remaining question is whether to add an MP2 based basis set correction ( $\Delta MP2$ ) to the CCSD results. Previous numerical experience shows that less converged CCSD results benefit from the MP2 correction. However, it is not possible to perform similar analysis comparing to better references at the size of the S66 set and at the convergence level of CCSD(F12<sup>\*</sup>)/ha(D,T)Z-F12. On the one hand, higher level of basis set convergence can be achieved for MP2 than for CCSD. For instance, the MP2-F12/a(T,Q)Z-F12 results of Ref. [26] exhibit for the ‘raw-CP’ convergence measure smaller RMSD and MAX values of 0.001 and 0.004 kcal/mol, respectively. This indicates higher level of basis set convergence for MP2-F12 at the a(T,Q)Z-F12 level than obtained for CCSD(F12<sup>\*</sup>) using only the ha(D,T)Z-F12 basis set level. For comparison, CP MP2-F12/ha(D,T)Z-F12 deviates from MP2-F12/a(T,Q)Z-F12 by 0.013 and 0.040 kcal/mol RMSD and MAX error values. One counterargument to using MP2-based basis set corrections is that the MP2/CBS results are quite far from the CCSD/CBS ones. Namely, the relation of MP2-F12 and CCSD(F12<sup>\*</sup>) IEs at the ha(D,T)Z-F12 level ranges from 10% underestimated [e.g. for (AcOH)<sub>2</sub>] to 30% overestimated (e.g. for  $\pi$ - $\pi$  stacking dimers), and it is little known about how these errors could influence the MP2-based CBS correction in combination with such well converged CCSD data.

For a more detailed understanding, the differences of CCSD and MP2 IEs contributions are collected in Table 3 compared to our best converged CP [CCSD(F12<sup>\*</sup>) – MP2-F12]/ha(D,T)Z-F12 reference. In brief, the results without explicit correlation do not exhibit strictly monotonic convergence with increasing basis set size, while the F12-based results show convincing trends. However, compared to the basis set completeness level of the CCSD results in Table 1, the [CCSD(F12<sup>\*</sup>) – MP2-F12]/ha(D,T)Z-F12 differences appear to be only about three times better converged. Therefore, the inaccuracies in MP2 can also affect the MP2-based CBS corrections, and we cannot expect that such MP2 corrected CCSD results match the convergence level of MP2-F12/a(T,Q)Z-F12. All in all, adding the MP2 correction to our final reference model seems to be beneficial, leading to: half-CP HF+CABS/aQZ-F12 + CP MP2/a(T,Q)Z-F12 + CP [CCSD(F12<sup>\*</sup>) – MP2-F12]/ha(D,T)Z-F12 + CP  $(T)/ha(T,Q)Z$ . This composite scheme involves higher level CCSD and  $(T)$  components than the one termed ‘14-carat GOLD’ in Ref. [26]. In order to simplify the discussion below we adopt the phrase of Ref. [26] and label our new reference as 14k-GOLD. The resulting 14k-GOLD reference IEs are given in Table 4.

Using the 14k-GOLD CCSD(T) reference, it is worthwhile analysing the performance of the lower-cost results collected in Table 5 both without and with MP2-based CBS corrections (left and right halves of Table 5, respectively). The RMSD measures of Table 5 show that the use of the  $\Delta MP2$  CBS corrections is beneficial with smaller basis sets up to haTZ, while the haQZ and ha(T,Q)Z results do not benefit from the correction. This trend appears to be reversed for the CCSD(F12<sup>\*</sup>)(T+) results, that is, MP2 corrected haTZ-F12 results are somewhat improved, while the uncorrected CCSD(F12<sup>\*</sup>)(T+)/haDZ-F12 IEs are preferred. The picture is much less clear when comparing the raw, CP, and half-CP as well as their MP2 (un)corrected versions with the same basis set. Depending on the AO basis set and the use of F12 approaches, basically any of the six CP and MP2 correction combinations can come up on top, and it is very hard to disentangle the compensation between the various sources of errors.

From the combinations of Table 5, CP CCSD(F12<sup>\*</sup>)(T+)/haTZ-F12 +  $\Delta MP2$  and CP CCSD(F12<sup>\*</sup>)(T+)/haDZ-F12 stand out as particularly accurate lower-cost alternatives. These only differ in a few aspects from the SILVER and BRONZE level composite approaches of Ref. [26] (cf. caption of Table 6). For a more detailed comparison, statistical measures for previous CCSD(T) benchmarks, including the SILVER, STERLING, and BRONZE composite schemes of Ref. [26], as well as

**Table 3.** RMSD and MAX errors with raw, CP, and half-CP (CP/2) for the difference of MP2 and CCSD compared to CP [CCSD(F12\*) – MP2-F12]/ha(D,T)Z-F12 IEs as reference [in kcal/mol].

Basis set	RMSD			MAX error			raw–CP
	raw	CP	CP/2	raw	CP	CP/2	
CCSD – MP2							
haDZ	0.096	0.187	0.137	0.186	0.441	0.313	0.273
haTZ	0.020	0.092	0.046	0.077	0.196	0.089	0.218
haQZ	0.043	0.059	0.012	0.127	0.118	0.023	0.244
ha(D,T)Z	0.038	0.057	0.029	0.104	0.135	0.078	0.170
ha(T,Q)Z	0.069	0.041	0.024	0.204	0.117	0.072	0.264
CCSD(F12*) – MP2-F12							
haDZ-F12	0.060	0.015	0.030	0.162	0.044	0.075	0.174
haTZ-F12	0.016	0.004	0.008	0.033	0.012	0.016	0.036
ha(D,T)Z-F12	0.007	ref	0.003	0.019	ref	0.010	0.019

Note: The last column collects the maximum absolute difference of raw and CP results with the corresponding basis set.

**Table 4.** The new, 14k-GOLD CCSD(T)/CBS estimates, in kcal/mol units, for the interaction energies of the S66 compilation using the following composite energy expression: half-CP HF+CABS/aQZ-F12 + CP MP2-F12/a(T,Q)Z-F12 + CP [CCSD(F12\*) – MP2-F12]/ha(D,T)Z-F12 + CP (T)/ha(T,Q)Z.

hydrogen bond		dispersion		mixed	
01 Water-Water	–4.991	24 Benzene-Benzene ( $\pi$ - $\pi$ )	–2.686	47 Benzene-Benzene	–2.806
02 Water-MeOH	–5.674	25 Pyridine-Pyridine ( $\pi$ - $\pi$ )	–3.760	48 Pyridine-Pyridine	–3.477
03 Water-MeNH <sub>2</sub>	–6.996	26 Uracil-Uracil ( $\pi$ - $\pi$ )	–9.716	49 Benzene-Pyridine	–3.267
04 Water-Peptide	–8.191	27 Benzene-Pyridine ( $\pi$ - $\pi$ )	–3.304	50 Benzene-Ethyne	–2.836
05 MeOH-MeOH	–5.827	28 Benzene-Uracil ( $\pi$ - $\pi$ )	–5.545	51 Ethyne-Ethyne	–1.528
06 MeOH-MeNH <sub>2</sub>	–7.625	29 Pyridine-Uracil ( $\pi$ - $\pi$ )	–6.661	52 Benzene-AcOH	–4.691
07 MeOH-Peptide	–8.313	30 Benzene-Ethene	–1.343	53 Benzene-AcNH <sub>2</sub>	–4.375
08 MeOH-Water	–5.071	31 Uracil-Ethene	–3.306	54 Benzene-Water	–3.263
09 MeNH <sub>2</sub> -MeOH	–3.091	32 Uracil-Ethyne	–3.675	55 Benzene-MeOH	–4.137
10 MeNH <sub>2</sub> -MeNH <sub>2</sub>	–4.189	33 Pyridine-Ethene	–1.784	56 Benzene-MeNH <sub>2</sub>	–3.171
11 MeNH <sub>2</sub> -Peptide	–5.442	34 Pentane-Pentane	–3.733	57 Benzene-Peptide	–5.219
12 MeNH <sub>2</sub> -Water	–7.360	35 Neopentane-Pentane	–2.583	58 Pyridine-Pyridine	–4.193
13 Peptide-MeOH	–6.250	36 Neopentane-Neopentane	–1.750	59 Ethyne-Water	–2.913
14 Peptide-MeNH <sub>2</sub>	–7.518	37 Cyclopentane-Neopentane	–2.375	60 Ethyne-AcOH	–4.933
15 Peptide-Peptide	–8.688	38 Cyclopentane-Cyclopentane	–2.957	61 Pentane-AcOH	–2.879
16 Peptide-Water	–5.178	39 Benzene-Cyclopentane	–3.482	62 Pentane-AcNH <sub>2</sub>	–3.495
17 Uracil-Uracil	–17.403	40 Benzene-Neopentane	–2.823	63 Benzene-AcOH	–3.716
18 Water-Pyridine	–6.929	41 Uracil-Pentane	–4.768	64 Peptide-Ethene	–2.977
19 MeOH-Pyridine	–7.465	42 Uracil-Cyclopentane	–4.055	65 Pyridine-Ethyne	–4.071
20 AcOH-AcOH	–19.383	43 Uracil-Neopentane	–3.660	66 MeNH <sub>2</sub> -Pyridine	–3.931
21 AcNH <sub>2</sub> -AcNH <sub>2</sub>	–16.485	44 Ethene-Pentane	–1.976		
22 AcOH-Uracil	–19.748	45 Ethyne-Pentane	–1.704		
23 AcNH <sub>2</sub> -Uracil	–19.424	46 Peptide-Pentane	–4.216		

the original (orig.) [17] and revised (rev.) [24] references of Řezáč and Hobza are collected in Table 6 with respect to the 14k-GOLD reference proposed here. Apparently, all but maybe the original set of IEs are found highly reliable, being within 0.075 kcal/mol MAX error and 0.033 kcal/mol RMSD of the new 14k-GOLD results. Considering the level of convergence for the individual terms in these composite schemes, such as the CCSD(T)/ha(D,T)Z part of “S66 revised” or the CCSD(F12\*)/(T<sub>sc</sub>)/DZ-F12 part of BRONZE, they benefit from some compensation of errors as also noted in Ref. [26]. These findings highlight the particularly successful design of previous composite schemes developed at the time when higher-level references were available only for the smallest few dimers in S66 [17, 24].

To assess the progress made via 14k-GOLD, let us also look at the difference of the raw and CP corrected

results corresponding to the energy components of the composite schemes as a measure of their basis set convergence level (see last row of Table 6). The respective values corresponding to 14k-GOLD are 0.008 and 0.024 kcal/mol RMSD and MAX difference, indicating that the 0.1 kcal/mol interaction energy gold-standard is indeed achieved. Compared to that, the same measures for SILVER are 0.023 and 0.056 kcal/mol RMSD and MAX difference, that is, they are about 2–3 times less converged according to this measure. The maximum of the |raw – CP| measure increases considerably up to 0.16, 0.47, 0.43, and 0.27 kcal/mol for the STERLING, BRONZE, S66 revisited, and S66 original benchmarks. Thus, the new 14k-GOLD reference exhibits a marked improvement regarding this basis set completeness measure, while it is proven to be very challenging to considerably improve upon the accuracy of previous composite

**Table 5.** RMSD with raw, CP, and half-CP (CP/2) and without and with the  $\Delta$ MP2 basis set correction for CCSD(T) and CCSD(F12\*)(T+) correlation energy contributions to the IEs compared to the here updated 14k-GOLD S66 reference [in kcal/mol].

Basis set	raw	CP	CP/2	raw	CP	CP/2
CCSD(T)			CCSD(T) + $\Delta$ MP2			
haDZ	0.853	0.998	0.548	0.143	<b>0.107</b>	0.117
haTZ	0.428	0.349	0.178	<b>0.021</b>	0.039	0.025
haQZ	0.141	0.134	<b>0.064</b>	0.142	0.117	0.065
ha(D,T)Z	0.294	0.055	0.161	0.055	0.063	<b>0.033</b>
ha(T,Q)Z	0.076	<b>0.024</b>	0.032	0.072	0.041	0.025
CCSD(F12*)(T+)			CCSD(F12*)(T+) + $\Delta$ MP2			
haDZ-F12	0.059	<b>0.032</b>	0.046	0.064	0.120	0.088
haTZ-F12	0.075	0.014	0.032	0.049	<b>0.006</b>	0.027
ha(D,T)Z-F12	0.057	0.038	0.047	0.026	0.026	0.026

**Table 6.** Statistical measures of previous composite CCSD(T) benchmarks for the S66 IEs compared to the reference updated here in kcal/mol units.

	SILVER [26]	STERLING [26]	BRONZE [26]	S66	rev. [24]	S66 orig. [17]
MSD	0.004	0.017	0.010	−0.031		0.003
MAD	0.007	0.022	0.028	0.031		0.078
MAX	0.038	0.075	0.073	0.048		0.290
RMSD	0.010	0.027	0.033	0.032		0.097
STD	0.009	0.021	0.032	0.009		0.098
MAX  raw−CP	0.056	0.156	0.465	0.428		0.266

Note: See text for explanation of the last row. SILVER [26]: half-CP MP2-F12/a(T,Q)Z-F12 + [CCSD(F12\*)−MP2-F12]/aTZ-F12 + (T)/ha(D,T)Z; STERLING [26]: half-CP MP2-F12/a(T,Q)Z-F12 + [CCSD(F12\*)−MP2-F12]/aTZ-F12 + raw (T)/sano(D,T)Z+; BRONZE [26]: half-CP CCSD(F12\*)(T<sub>CS</sub>)/DZ-F12; S66 rev. [24]: CP MP2/a(T,Q)Z + [CCSD(T) − MP2]/ha(D,T)Z; S66 orig. [17]: CP MP2/a(T,Q)Z + [CCSD(T) − MP2]/aDZ

energy expression based IEs. As expected, the largest difference between SILVER and our new reference is found for the dimers with the slowest basis set convergence, namely  $\pi$ - $\pi$  stacked complexes (e.g. uracil<sub>2</sub> and uracil-pyridine) and H-bonded complexes [e.g. (AcOH)<sub>2</sub>]. Having CCSD(F12\*)/haQZ-F12 and (T)/ha(Q,5)Z references at hand for (AcOH)<sub>2</sub>, we can be very confident that the new CCSD and (T) reference IE components are converged within 0.003 and 0.005 kcal/mol, respectively, while it would require an enormous effort to similarly assess the accuracy for the larger  $\pi$ - $\pi$  stacked dimers.

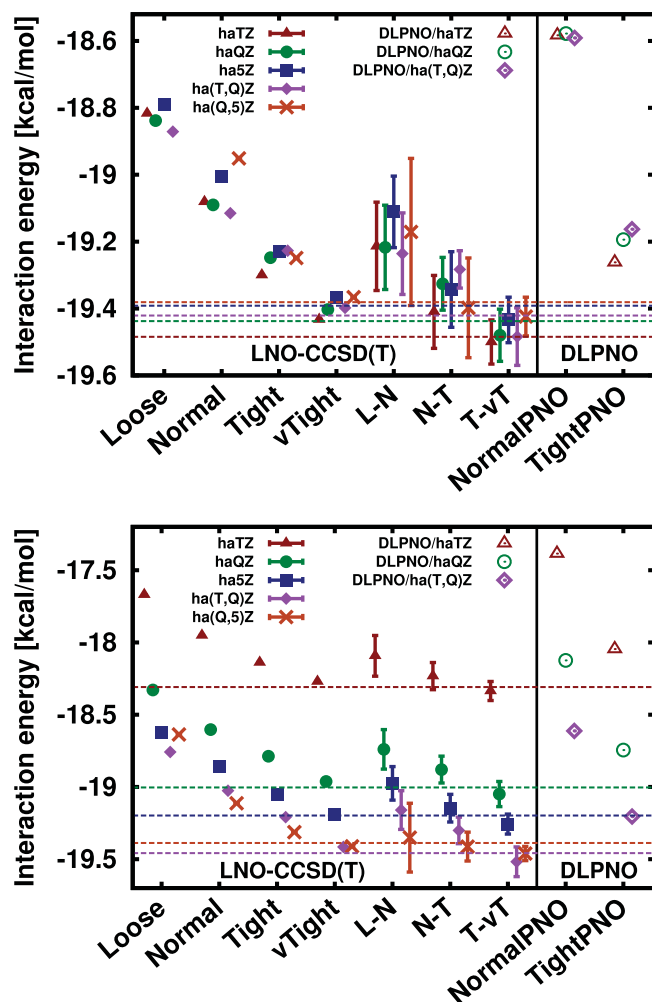
Besides the analysis of IE errors, it is also important to consider the relative errors as the S66 IEs spread across a wide range from −1.3 to −19.7 kcal/mol. In brief, we inspect the |raw−CP| RMSD and MAX error basis set convergence measures relative to the 14k-GOLD reference. These are found for the individual energy components to be as follows: 0.036% and 0.103% for MP2-F12/a(T,Q)Z-F12; 0.207% and 0.708% for [CCSD(F12\*)−MP2-F12]/a(D,T)Z-F12; 0.109% and 0.286% for (T)/ha(T,Q)Z. That is, the corresponding relative RMSD and MAX uncertainties are 0.266% and 0.784% for the total CCSD(T) IEs. Thus, this robust measure indicates that the IEs are converged with at most 0.8% relative accuracy, while the same measure is still accurate for the SILVER IEs being about twice as large (1.5%) as for 14k-GOLD.

### 4.3. Local natural orbital and other local CCSD(T) approaches

Local correlation approaches, such as the LNO-CCSD(T) as well as PNO-based schemes considered here, offer significant cost reduction for CCSD(T). For instance, they facilitate the increase of the basis set used for medium-sized molecules in the size range of S66 or allow for computations for much larger complexes than accessible via conventional CCSD(T). However, all of these local correlation methods introduce a set of approximations which have to be carefully assessed. Thus, first we exploit the possibility to test the accuracy of various local approximations against the new, large basis set references since so far local CCSD(T) benchmarks against exact CCSD(T) were not possible in the size range of S66 up to haQZ or ha(T,Q)Z [89, 91, 98, 100]. Second, we utilise well-converged LNO-CCSD(T) IEs up to ha(Q,5)Z to assess the basis set convergence level at the ha(T,Q)Z level used also for the (T) part of the new 14k-GOLD S66 reference IEs.

First, it is worthwhile utilising the CCSD(T)/ha5Z IEs of (AcOH)<sub>2</sub> presented in Sect. 4.1 to illustrate the local CCSD(T) benchmark process on an example, where all exact CCSD(T) references are available up to ha(Q,5)Z. The (AcOH)<sub>2</sub> IEs are collected on the top and bottom of Figure 5 without and with CP corrections, respectively. The left panels collect LNO-CCSD(T) IEs with





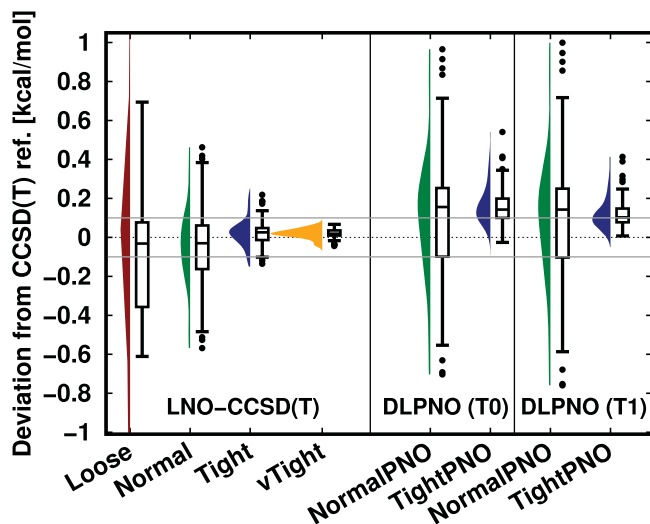
**Figure 5.** LNO-CCSD(T) (left panel, solid point type) and DLPNO-CCSD(T1) (right panel, hollow point type) IEs using various local correlation threshold settings (x axis labels) for the (AcOH)<sub>2</sub> dimer without (top) and with CP (bottom) corrections for various basis sets. Horizontal lines indicate the CP (un)corrected DF-CCSD(T) reference IEs using the same colour for the local CCSD(T) and the DF-CCSD(T) results with the same basis set. The DF-CCSD(T) reference IEs on the top panel without CP become less attractive with increasing the basis set, they appear in increasing order on the plot as haTZ, haQZ, ha(T,Q)Z, ha5Z, and ha(Q,5)Z. The CP corrected DF-CCSD(T) reference IEs on the bottom panel become more attractive with increasing the basis set, they appear in decreasing order on the plot as haTZ, haQZ, ha5Z, ha(Q,5)Z, and ha(T,Q)Z.

Loose to very Tight (vTight) threshold settings as well as the corresponding Loose–Normal (L–N), etc. extrapolated values. The right panels show DLPNO-CCSD(T1) IEs with the NormalPNO and TightPNO settings as recommended by the developers [87, 101]. Reference DF-CCSD(T) IEs, indicated via dashed horizontal lines in Figure 5, represent the local approximation free limit of both LNO-CCSD(T) and DLPNO-CCSD(T1) IEs due to the consistent use of the DF approximation in all three approaches.

Considering that (AcOH)<sub>2</sub> is one of the most complicated dimers of the S66 set in terms of the basis set requirement, the rate of convergence of the LNO-CCSD(T) results and their balanced accuracy across the basis sets is highly satisfactory. The Loose LNO-CCSD(T) results are only useful for quick but crude exploration, thus Normal, Tight, or vTight settings are required to reach, respectively, 0.3–0.4, 0.1–0.2, or even 0.02–0.05 kcal/mol deviations from the reference both with and without CP. In this case the LNO extrapolation consistently improves the IEs by about a whole step in the LNO setting hierarchy for all basis sets. For instance, the quality of the L–N and Normal–Tight (N–T) extrapolated LNO-CCSD(T) results match that obtained with the more demanding Tight and vTight settings, respectively. More importantly, the error bars assigned to the extrapolated IEs according to Equation (5) estimate the size of the next step in the hierarchy well, that is, Tight (vTight) results are indeed within the range defined by the error bars of L–N (N–T). As the L–N results are still about 0.2–0.3 kcal/mol away from the exact reference, their respective error bars slightly underestimate this distance toward the local approximation free limit. Therefore, for high accuracy production computations N–T or better settings are recommended, while Loose thresholds are useful to obtain an additional set of results cost-effectively to monitor the convergence pattern. It is interesting to note that the case of raw (AcOH)<sub>2</sub> IEs does not follow the trend observed in our previous LNO-CCSD(T)-based investigations of IEs [27, 91, 107–109], that is, the basis set error is usually larger than the error of the local approximations. Instead, the CP corrected (AcOH)<sub>2</sub> IEs on the bottom of Figure 5 exhibit the more typical case of larger basis set incompleteness than LNO errors.

For the DLPNO-CCSD(T1) results on the right panels of Figure 5, the homogeneity of the IE errors with various basis sets is similarly satisfactory. At least for the case of (AcOH)<sub>2</sub>, the NormalPNO settings are a bit less accurate than Loose LNO-CCSD(T). In general, NormalPNO DLPNO-CCSD(T1) performs better than Loose LNO-CCSD(T) for the entire S66 set as shown below, e.g. in Figure 6. Moreover, in a large step of improvement, TightPNO DLPNO-CCSD(T1) IEs for (AcOH)<sub>2</sub> become comparable to Tight or L–N LNO-CCSD(T) for all basis sets evaluated.

Let us continue with a similar convergence analysis for the entire S66 set. Tables 7 and 8 collect statistical measures illustrating the accuracy of LNO-CCSD(T) and DLPNO-CCSD(T) correlation energies and IEs, respectively, compared to the exact DF-CCSD(T) reference up to ha(T,Q)Z. Tables 7 and 8 show results only without CP corrections since the accuracy of the CP corrected



**Figure 6.** Signed IE errors with respect to the DF-CCSD(T) reference for LNO-CCSD(T) (left panel), DLPNO-CCSD(T0) (middle), and DLPNO-CCSD(T1) (right panel) using various threshold sets. Violin curves show the distribution of the signed IE errors, where the height of the curve (along the horizontal axis) indicates the frequency of the signed IE errors corresponding to an error value on the vertical axis. The horizontal lines of the boxes indicate the lower, median, and upper quartiles, respectively. Whiskers extend to the most distant data point whose IE error value lies within 1.5 times of the interquartile range (IQR), the latter of which is the difference of the lower and upper quartiles. Outliers beyond the whiskers, if any, are represented by dots. The gray horizontal lines highlight the  $\pm 0.1$  kcal/mol IE error mark.

correlation and interaction energies were found to be practically identical to those without CP corrections for both the LNO and the DLPNO approach.

Let us start with the MAE and MAX measures of the relative correlation energy errors collected in Table 7 with the haTZ, haQZ, and ha(T,Q)Z basis set levels. The top third of Table 7 shows the systematic convergence of the LNO-CCSD(T) correlation energies from Loose up to very very Tight (vvTight) settings. For instance, Loose LNO-CCSD(T)/haTZ correlation energies are already at the 0.045% relative error level, which decreases, in a highly regular manner, by about a factor of 2, in each step throughout the tightest settings. The MAX LNO-CCSD(T)/haTZ errors are similarly consistent being about 3 times higher than the corresponding MAE, that is, 99.9% accuracy is provided for all S66 dimers already with the Normal settings. Both error measures increase notably by about a factor of 2 when moving from haTZ to haQZ and then to ha(T,Q)Z, but the Tight results still maintain the at least 99.9% accuracy even at the ha(T,Q)Z level. Interestingly, LNO extrapolated results, collected in the middle third of Table 7 exhibit much more homogeneous relative correlation energy errors across the three basis set levels than those

without extrapolation. While the LNO-CCSD(T) correlation energies show minor improvements upon LNO extrapolation with the haTZ basis, the LNO extrapolation drastically improves the haQZ and ha(T,Q)Z correlation energy errors in Table 7 to a level comparable to the accuracy of LNO-CCSD(T)/haTZ. Since this is the first possibility to perform such an analysis on the LNO extrapolation for sufficiently large molecules at the haQZ basis set level, much more data are required to check the generality of the above trends.

Finally, let us compare the CP corrected LNO-CCSD(T) results using the updated domain construction algorithm of Section 2.2. As noted, the CP uncorrected and corrected interaction energies are almost identical. Namely, Loose MAE values for raw and CP differ at most by only 0.004%, which gradually disappears as the LNO thresholds are tightened. Interestingly, the CP corrected monomer correlation energies are barely affected by the updated domain construction algorithm, changes in the above relative correlation energy error statistics are negligible. The largest improvement is observed for the water monomer of the water-pyridine dimer (no. 18), where the change is 0.02% with Loose settings. The average absolute change for the entire set with Loose settings is only 0.002% and decreases even further when the LNO thresholds are tightened. This trend can be explained considering the relatively small size and the relatively short, equilibrium distance of the monomers. While the results with the increased domain sizes are encouraging from the perspective of the previous algorithm [36, 91], non-negligible changes may occur for larger complexes (e.g. for the L7 set [18, 27]) or for more distant monomers, such as those in the S66x8 compilation [17].

The analogous DLPNO-CCSD(T) relative correlation energy errors are also collected in the bottom third of Table 7 with both NormalPNO and TightPNO settings, as well as with both the DLPNO-(T0) and the DLPNO-(T1) treatment of the triples corrections [87, 100]. It is important to highlight that the DLPNO-approximation free limit of the DLPNO-CCSD(T0) correlation energies would be a DF-CCSD(T0) set of reference values, which is not identical to the DF-CCSD(T) reference used here. Nevertheless, looking at the DLPNO-CCSD(T0) results too reveals that the differences of the (T0) and (T1) correlation energies are about 0.2% across all basis sets and DLPNO thresholds settings considered. Interestingly, the DLPNO-CCSD(T1) results improve with increasing basis set size. While TightPNO DLPNO-CCSD(T1)/haTZ is comparable to Loose LNO-CCSD(T)/haTZ, NormalPNO DLPNO-CCSD(T1)/haQZ is closer to Normal LNO-CCSD(T)/haQZ, which trend may be related to the compensation between basis set and DLPNO errors noted

**Table 7.** MAE and MAX relative correlation energy errors (in %) with respect to the DF-CCSD(T) reference for various LNO- and DLPNO-CCSD(T) settings without CP corrections. Top and middle thirds: LNO-CCSD(T) using various composite threshold sets without and with LNO extrapolation, respectively. Bottom third: DLPNO-CCSD(T) using NormalPNO and TightPNO settings as well as the (T0) and (T1) triples correction algorithms.

Local threshold	haTZ		haQZ		ha(T,Q)Z	
	MAE	MAX	MAE	MAX	MAE	MAX
Loose	0.045	0.134	0.225	0.378	0.401	0.573
Normal	0.029	0.094	0.081	0.167	0.148	0.244
Tight	0.016	0.046	0.032	0.076	0.057	0.108
vTight	0.007	0.018	0.019	0.043	0.032	0.060
vvTight	0.003	0.010	0.011	0.022	0.017	0.031
Loose-Normal	0.025	0.092	0.028	0.065	0.035	0.080
Normal-Tight	0.011	0.026	0.013	0.035	0.015	0.045
Tight-vTight	0.005	0.013	0.014	0.027	0.021	0.037
vTight-vvTight	0.002	0.007	0.006	0.013	0.010	0.019
NormalPNO (T0)	0.321	0.534	0.261	0.471	0.221	0.429
NormalPNO (T1)	0.141	0.312	0.091	0.259	0.064	0.223
TightPNO (T0)	0.238	0.367	0.211	0.329	0.193	0.306
TightPNO (T1)	0.049	0.113	0.029	0.089	0.020	0.074

**Table 8.** MAE and MAX interaction energy errors [in kcal/mol] with respect to the DF-CCSD(T) reference for various LNO- and DLPNO-CCSD(T) settings without CP corrections. Further details are given in the caption of Table 7.

Local threshold	haTZ		haQZ		ha(T,Q)Z	
	MAE	MAX	MAE	MAX	MAE	MAX
Loose	0.313	1.319	0.340	1.387	0.364	1.454
Normal	0.159	0.568	0.169	0.602	0.178	0.668
Tight	0.053	0.219	0.054	0.199	0.058	0.194
vTight	0.027	0.067	0.020	0.063	0.021	0.077
vvTight	0.013	0.048	0.013	0.055	0.017	0.069
Loose-Normal	0.099	0.346	0.093	0.308	0.095	0.367
Normal-Tight	0.059	0.187	0.057	0.165	0.057	0.150
Tight-vTight	0.024	0.109	0.020	0.077	0.022	0.065
vTight-vvTight	0.012	0.059	0.015	0.068	0.019	0.076
NormalPNO (T0)	0.269	0.965	0.265	0.883	0.265	0.957
NormalPNO (T1)	0.275	0.999	0.271	0.907	0.272	1.020
TightPNO (T0)	0.167	0.541	0.155	0.520	0.146	0.503
TightPNO (T1)	0.125	0.413	0.113	0.397	0.107	0.382

previously [140]. Considering both the MAE and MAX relative correlation energy measures and all basis sets, TightPNO DLPNO-CCSD(T1) practically reaches 99.9% accuracy for all species in S66. This is, however, slightly worse than the convergence level of the more cost-efficient Loose-Normal LNO-CCSD(T) correlation energies, indicating that a smaller compensation of errors is sufficient in LNO correlation energies to provide comparable quality IEs. Finally, since not many benchmarks are available in the literature dedicated to assess the performance of the DLPNO method with CP corrections, it is valuable to note on the near identical performance of all DLPNO variants with and without CP corrections for all considered basis set choices.

To continue with the analogous analysis for IEs, the corresponding MAE and MAX IE error measures with respect to the DF-CCSD(T) reference are collected in

Table 8 for the same LNO and DLPNO settings and basis sets. For the sake of brevity, we again show only the CP uncorrected statistics because the CP corrected ones are practically identical for both local correlation approaches and for all three levels of basis set.

Regarding the LNO-CCSD(T) IEs, we again find systematic convergence and a factor of 2, sometimes even a factor of 3, improvement in the MAE values for all three levels of basis set. The MAX errors, being consistently about 3–4 times larger than the MAE values, thus follow a similar trend. The Normal (Tight) LNO settings provide 0.1–0.2 (sub 0.1) kcal/mol IE accuracy on the average, but Tight or vTight settings are needed for similar accuracy regarding the MAX errors. Comparing the haTZ, haQZ, and the ha(T,Q)Z results, we find a slight decrease in the quality of the IEs with increasing the basis set size. However, this effect is much milder than what is observed for

the basis set dependence of the LNO-CCSD(T) correlation energies above. In other words, the larger correlation energy errors compensate very well when forming the interaction energies, indicating that the IEs may converge somewhat faster with the basis set size than the correlation energies. Such an effect also contributes to the trend that, except for the L–N variant, the LNO extrapolation does not considerably improve the IE error statistics. An additional factor to consider is that so close to the converged IEs, the natural orbital approximation may not dominate the overall LNO error. Thus, other sources of errors with possibly different signs could become comparable, which, in turn, can affect the performance of the LNO extrapolation. Nevertheless, such cases can be identified by observing the convergence pattern of the IEs of individual dimers.

The analogous data for the DLPNO-CCSD(T) variants are collected in the bottom third of Table 8. The performance of all four DLPNO-CCSD(T) variants, just as for the correlation energies, is consistent for all three basis set levels. Based on the MAE and MAX IE error measures alone, both the (T0) and the (T1) variants with NormalPNO settings are about 15–45% more accurate than Loose LNO-CCSD(T), while the TightPNO variants are comparable to or similarly outperform Normal LNO-CCSD(T) with the haTZ and the larger basis sets, respectively.

Besides considering only the MAE and MAX error measures, it is useful to look at the complete error distributions too. Since the IE error trends are fairly consistent for all three basis sets, and with and without CP for both the LNO and DLPNO methods, the IE error distributions are only displayed for the case of haTZ in Figure 6. Violin curves of Figure 6 show the distribution of the signed IE errors, where the height of the curve indicates the frequency of the signed IE errors corresponding to an error value on the vertical axis. On the right side of the error distribution curves, the horizontal lines of the boxes indicate the lower, median, and upper quartiles, respectively, encompassing the middle 50% of the data points. Whiskers extend to the most distant data point whose IE error value lies within 1.5 times of the interquartile range (IQR), the latter of which is the difference of the lower and upper quartiles. Outliers beyond the whiskers, if any, are represented by dots.

One apparent feature of Figure 6 is the high symmetry of the LNO-CCSD(T) error curves around the 0 kcal/mol error line and the correspondingly small mean of the signed errors. Clearly, most of the Loose and about half of the Normal LNO-CCSD(T) IEs show errors above 0.1 kcal/mol, but the two middle quartiles of Normal LNO-CCSD(T) are almost completely within the gray lines indicating the  $\pm 0.1$  kcal/mol error mark. Moreover,

most of the Tight and all of the vTight results fall in the  $\pm 0.1$  kcal/mol error range. It is interesting to have a closer look at the dots representing the outliers. For both Normal and Tight LNO-CCSD(T), the IE of the most challenging H-bonded ( $\pi$ - $\pi$  stacking) dimers are consistently underestimated (overestimated) [e.g. H-bonded complexes of AcOH, AcNH<sub>2</sub>, and uracil, as well as the  $\pi$ - $\pi$  stacking dimers of uracil, benzene, and pyridine]. The latter trend was also observed for more extended  $\pi$ - $\pi$  dimers investigated in Ref. [27].

Considering the DLPNO-CCSD(T) results, the middle and right panels of Figure 6 reveal error distributions shifted toward underestimated IEs with MSE measures being consistently in the 0.1–0.2 kcal/mol range. The middle quartile and 1.5 IQR performance of NormalPNO DLPNO-CCSD(T) is comparable to Loose LNO-CCSD(T) but the outliers with NormalPNO DLPNO-CCSD(T) are markedly better. The TightPNO DLPNO-CCSD(T) results, especially with the (T1) triples correction, have middle quartile and 1.5 IQR widths comparable to those of Tight LNO-CCSD(T), but the symmetry, MSE, and outlier performance of Tight LNO-CCSD(T) are apparently better.

Having the expected accuracy of the local approximations established at least up to the ha(T,Q)Z level of basis set, let us exploit their efficiency to approach the basis set limit of the IEs even closer, at the ha(Q,5)Z level. One complication is that the AO basis set for some of the dimers exhibit considerable linear dependency in the ha5Z basis. To our knowledge, only the LNO approach has features implemented to treat the quasi-redundancy of such large basis sets [91], thus the ha5Z results are only computed with the LNO-CCSD(T) method. For comparison, we also gather recently reported and well-converged PNO-based explicitly correlated CCSD(T) results with up to haQZ-F12 basis sets [98, 99].

First, we consider the difference of the raw and CP corrected LNO-CCSD(T) results as a measure of the basis set convergence. While the average  $|\text{raw}-\text{CP}|$  value converges convincingly as 0.74, 0.26, and 0.09 kcal/mol with the haTZ, haQZ, and ha5Z basis sets, respectively, the ha(T,Q)Z and ha(Q,5)Z  $|\text{raw}-\text{CP}|$  results tend to similar values in the 0.04–0.05 kcal/mol range using even vvTight LNO settings. Compared to that, the average LNO-CCSD(T) error in the  $|\text{raw}-\text{CP}|$  measure is about 0.01 kcal/mol, at least at the ha(T,Q)Z level where the reference is known. This is in accord with the 0.017 kcal/mol MAE value in Table 8 obtained for the vvTight LNO-CCSD(T)/ha(T,Q)Z IEs. The corresponding MAX error being 0.07 kcal/mol indicates that the LNO error is probably somewhat larger than the basis set uncertainty for the CBS extrapolated results, but some compensation of LNO errors can be expected in the  $|\text{raw}-\text{CP}|$  measures.



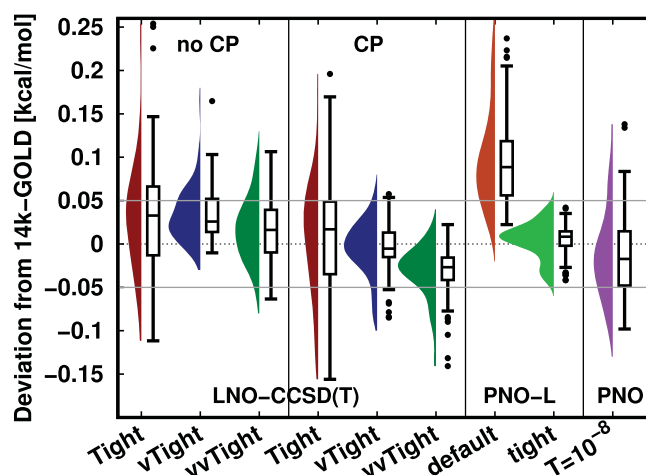
**Table 9.** MAE and MAX interaction energy errors [in kcal/mol] with respect to the 14k-GOLD reference for various LNO- and PNO-based CCSD(T) schemes without and with CP corrections.

Local threshold	raw		CP	
	MAE	MAX	MAE	MAX
Loose	0.544	2.203	0.532	2.218
Normal	0.160	0.664	0.157	0.560
Tight	0.057	0.254	0.053	0.196
vTight	0.035	0.165	0.023	0.085
vvTight	0.031	0.106	0.035	0.141
default PNO-L [99]			0.100	0.276
tight PNO-L [99]			0.017	0.042
PNO/aTZ [98]	0.042	0.138		
PNO/a(D,T)Z [98]	0.103	0.435		

Note: Top third: LNO-CCSD(T)/ha(Q,5)Z with Loose to vvTight settings. Middle: CP corrected PNO-LCCSD(T\*)-F12b/haQZ-F12 with default and tight settings [99]. Bottom: not CP corrected PNO-CCSD(F12\*)(T) with  $10^{-8}$  PNO thresholds and aTZ or a(D,T)Z level of basis set [98].

We can also compare the LNO-CCSD(T)/ha(Q,5)Z results to the new 14k-GOLD reference, keeping in mind that the latter has still up to a few tens of cal/mol uncertainty in its basis set convergence. The corresponding MAE and MAX error values are collected in Table 9 both without and with CP corrections. Again, the convergence toward the LNO approximation free limit is very convincing as the Tight MAE values already drop below 0.06 kcal/mol, while vTight and vvTight are very close to each other in the 0.02–0.04 kcal/mol MAE range. Interestingly, vTight LNO-CCSD(T)/ha(Q,5)Z results with CP are somewhat closer to the 14k-GOLD reference than with vvTight settings, indicating that CP corrected LNO-CCSD(T)/ha(Q,5)Z could tend to a slightly different CBS limit than the 14k-GOLD values obtained with smaller AO basis sets.

To inspect this in more detail, distributions of the signed deviation of LNO-CCSD(T)/ha(Q,5)Z with respect to the 14k-GOLD reference are plotted on the violin curves of Figure 7. The corresponding boxes and whiskers on the right side of the distributions again indicate the middle two quartiles and the 1.5 IQR. A notable difference can be observed if we compare the convergence pattern of the LNO-CCSD(T)/ha(Q,5)Z deviations with respect to the 14k-GOLD reference to the case of LNO-CCSD(T) errors against the exact DF-CCSD(T) reference in Figure 6. For instance, the widths of the distributions do not drop as significantly when stepping from Tight to vTight as against the exact reference, which indicates that the remaining basis set incompleteness error of 14k-GOLD may interfere with the systematically decreasing LNO errors when tending from Tight toward vTight and vvTight settings. This observation holds for both the raw and the CP corrected LNO-CCSD(T)/ha(Q,5)Z results on the two left panels

**Figure 7.** Signed IE deviations with respect to the 14k-GOLD reference for raw and CP corrected LNO-CCSD(T)/ha(Q,5)Z (left two panels), CP corrected PNO-LCCSD(T\*)-F12b/haQZ-F12 [99] (third panel), and raw PNO-CCSD(F12\*)(T)/aTZ [98] (rightmost panel). See the caption of Figure 6 for details of the violin and the box-and-whisker plot notation.

of Figure 7. However, raw LNO-CCSD(T)/ha(Q,5)Z IEs tend to an LNO approximation free limit that underestimates 14k-GOLD, while their CP corrected counterparts settle at the opposite side overestimating 14k-GOLD. Since 14k-GOLD is also constructed using CP corrected energy components, CP corrected vvTight LNO-CCSD(T)/ha(Q,5)Z may reveal here that 14k-GOLD is not attractive enough, at least on the scale of few tens of cal/mol. On the other hand, the maximum LNO uncertainty is probably somewhat larger than a few tens of cal/mol, especially for the  $\pi$ - $\pi$  stacking dimers, which are the outliers among the CP corrected vvTight results. Therefore, it remains very challenging to separate the comparable sized local and basis set incompleteness errors in the difference of LNO-CCSD(T)/ha(Q,5)Z and 14k-GOLD.

Additional information may be obtained in this regard by inspecting alternative local correlation schemes, such as the PNO-LCCSD(T\*)-F12b/haQZ-F12 results of Ma and Werner [99] as well as the PNO-CCSD(F12\*)(T)/aTZ IEs reported by Schmitz and Hättig [98]. Both sets of results utilise the explicitly correlated treatment at the CCSD level and fairly tight PNO truncation thresholds ( $T = 10^{-8}$ ). The PNO-LCCSD(T\*)-F12b results benefit from the large, haQZ-F12 basis set and CP correction, while the non CP corrected PNO-CCSD(F12\*)(T)/aTZ IEs take advantage of the excellent performance of CCSD(F12\*)(T) noted above in Sect. 4.2. Inspecting Table 9 and Figure 7, we find the PNO-LCCSD(T\*)-F12b results with its default domain construction options about halfway between our CP corrected Normal

and Tight LNO-CCSD(T)/ha(Q,5)Z results. The performance of PNO-CCSD(F12\*)(T)/aTZ is close to Tight LNO-CCSD(T)/ha(Q,5)Z with an even smaller number of outliers. The PNO-LCCSD(T\*)-F12b results clearly benefit from the tight domain settings being even a bit closer to 14k-GOLD than the tighter LNO variants. Interestingly, the tight PNO-LCCSD(T\*)-F12b/haQZ-F12 results, on the average, slightly underestimate the IEs of 14k-GOLD despite both having CP corrections.

All in all, the investigated local CCSD(T) approaches all have excellent qualities as well as some remaining drawbacks, which prevent us from drawing decisive conclusions. For instance, the LNO results are very tightly converged and reach the ha(Q,5)Z level, but do not benefit from the advantages of the F12 treatment. While the CCSD part of tight PNO-LCCSD(T\*)-F12b/haQZ-F12 could be very well converged, the scaled (T\*) treatment still has uncertainties up to few tens of cal/mol, as pointed out here as well as in Ref. [99]. At this point, because of the agreement of several different sources, we can be confident that the uncertainty of our CCSD(T)/CBS estimates does not exceed 0.05 kcal/mol, but it will be quite challenging to shrink this uncertainty window further in a more decisive manner.

Finally, it is important to analyse the same data from the perspective of relative IE errors. This is a challenging property, especially for the dispersion dominated dimers. For these dimers, a large portion of the repulsive HF and attractive correlation energy contributions to the IEs partly cancel each other, and thus the total IE is often considerably smaller than the separate HF and CC components. This amplifies the errors caused in the correlation energy contributions, when they are compared to the smaller total IEs. Such relative IE errors, evaluated against the exact DF-CCSD(T) reference, are collected in Table 10.

Here, we find the MAE of Normal LNO-CCSD(T) at the 3–4% range, while Tight and vTight settings improve this metric to about 1% and 0.5%, respectively. Compared

to that, NormalPNO and TightPNO DLPNO-CCSD(T1) MAEs are at the 5–6% and 2% range, respectively. Thus, the comparison of LNO-CCSD(T) and DLPNO-CCSD(T) relative errors shows trends similar to the case of the absolute IE errors. However, the MAX errors being about 4–5 times larger than the MAE values reveal an important aspect. Namely, large relative errors can occur also for cases with fairly acceptable absolute errors if the total IE is also small. For instance, the largest relative error of 2.2% for vTight LNO-CCSD(T)/haTZ corresponds to the 0.04 kcal/mol IE error obtained for the neopentane dimer due to its about  $-1.94$  kcal/mol total IE. Nevertheless, the MAX LNO-CCSD(T) errors reaching 2–3% and about 1.5% with vTight and vvTight settings, respectively, are fairly promising as it is feasible to obtain LNO-CCSD(T) results with such well-converged settings even above 100 atoms [27, 91].

## 5. Conclusions and outlook

In this report, we utilised our state-of-the-art DF-based conventional [37], explicitly correlated [62], and local natural orbital (LNO) [36, 91] CCSD(T) approaches to push forward the understanding of the basis set convergence of CCSD(T) interaction energies in medium-sized molecular complexes. These recent methodological and algorithmic developments allowed us for the first time to compute augmented quadruple- and quintuple- $\zeta$  level DF-CCSD(T) and LNO-CCSD(T) interaction energies (IEs), respectively, for the entire S66 molecular complex compilation. A comprehensive analysis of previous references and our new benchmarks suggest updating the S66 reference using the CCSD(F12\*)(T)/ha(D,T)Z-F12-based CCSD contributions combined with (T)/ha(T,Q)Z triples corrections. Following the terminology of Martin and co-workers [26], the new CCSD(T)/CBS estimate is labelled by (14-karat) 14k-GOLD. The new reference exhibits about 0.01 (0.05) kcal/mol average (maximum) basis set convergence uncertainty, while the relative accuracy of the correlation energies appear to fall below the maximum 1% mark for the first time. The CCSD(F12\*)(T+)/haTZ-F12 interaction energies, relying on our recent size-consistent (T+) method [62], offer a close second best option with only at most 0.016 kcal/mol deviation from 14k-GOLD without the need for quadruple- $\zeta$  level computations.

While shrinking these uncertainties even further at the moment poses considerable challenges for conventional CCSD(T) codes, there is a reason for optimism. For instance, the largest CCSD(T) computation five years ago in the previous S66 benchmark work [26] contained 1200 atomic orbitals for the uracil dimer, which required 8 days runtime on 96 cores, 1.5 TB memory, and 18 TB

**Table 10.** MAE and MAX relative errors of total IEs (in %) with respect to the DF-CCSD(T) reference for various LNO- and DLPNO-CCSD(T) settings without CP corrections.

Basis set:	haTZ		haQZ		ha(T,Q)Z	
Local threshold	MAE	MAX	MAE	MAX	MAE	MAX
Loose	6.6	32.2	8.2	42.4	9.8	52.3
Normal	3.1	12.8	3.9	17.9	4.4	22.9
Tight	0.9	3.2	1.1	5.0	1.4	6.8
vTight	0.6	2.2	0.5	2.3	0.5	3.0
vvTight	0.3	1.4	0.3	1.5	0.4	1.8
NormalPNO	5.2	21.6	5.8	28.5	6.4	34.8
TightPNO	2.3	4.9	2.2	4.7	2.2	4.6

Note: Further details are given in the caption of Table 7.

SSD disk space. Compared to that, our largest CCSD(T) computation so far with 31% more orbitals required only 3 days on 224 cores and 0.33 TB memory using our recent implementation in the MRCC package [37]. Employing the approximate but much more efficient LNO-CCSD(T) approach, we were able to estimate the S66 IEs at the LNO-CCSD(T)/ha(Q,5)Z level. Considering also the local approximation error in the latter, the uncertainty level of about 0.05 kcal/mol of the best CCSD(T) estimates has been corroborated using multiple independent approaches including previous benchmarks [26, 99], 14k-GOLD, and tightly converged LNO-CCSD(T)/ha(Q,5)Z. Such highly converged results show that the deviation of DMC and CCSD(T) IEs at the  $0.3 \pm 0.2$  kcal/mol scale, e.g. for the benzene or uracil dimers [27], cannot solely be explained by the basis set incompleteness error of CCSD(T) IEs.

The new DF-CCSD(T)/haQZ references also allowed us to benchmark our LNO and the commonly employed DLPNO local correlation approaches against exact references for the first time with such a large basis set at the system size of the S66 dimers. Importantly, both the LNO and the DLPNO approaches were found to maintain their accuracy levels established up to the haTZ level also with the larger haQZ basis set, upon basis set extrapolation, and in combination with counterpoise corrections. For instance, LNO-CCSD(T) IEs with Tight settings exhibit about 0.05 (0.2) kcal/mol average (maximum) deviations with respect to the exact reference, while the TightPNO DLPNO-CCSD(T1) IE errors are about twice as high. If needed, more tightly converged LNO-CCSD(T) can reliably approach the local approximation free CCSD(T) results matching the about 0.01 (0.05) kcal/mol average (maximum) uncertainty of the new basis set incompleteness estimate. Due to its efficient implementation with small memory and disk I/O requirements, LNO-CCSD(T) is thus an excellent choice to generate large data-sets with highly-accurate references used routinely for the parametrization or assessment of DFT methods [141, 142] and for the training step in machine learning approaches [21–23, 143–151].

It is important to note that there is a large gap between the size and complexity of the S66 dimers and those of the systems that are in the applicability range of local CCSD(T) approaches as, for example, our LNO-CCSD(T) code was already shown to scale up to a 1000-atom protein-ligand complex with a quadruple- $\zeta$  basis set [91]. The presented excellent performance of the local correlation approaches is partly due to the moderate system size and IE strength of the S66 dimers. However, approaching the rigorous few percent relative accuracy in the CCSD(T) IEs is already important for the most complicated but medium-sized S66 dimers. Such accuracy

enables rigorous and informative comparisons with alternative methods capable of high-accuracy, including local correlation approaches, DMC, SAPT, RPA, and maybe even the leading DFT methods [25]. Moreover, as the system size grows to a very large scale, the IE has a faster than linear scaling component in terms of the size of the interacting surface and has been documented to reach up to 100 kcal/mol CCSD and 10–20 kcal/mol (T) contributions for a 132-atom fullerene-in-a-nanoring [27] and a 1023-atom protein-ligand complex [91] studied recently. In such cases, a relative error of even better than a few percent is required for well-converged IEs. Therefore, when more challenging and more extended molecular complexes are studied, we recommend cautious convergence tests and extrapolation toward the local approximation free CCSD(T) limit to estimate [91] and to avoid potentially high local correlation errors.

## Acknowledgments

This paper is dedicated to Professor Péter G. Szalay on the happy occasion of his 60th birthday. The authors wish to express their gratitude to Péter for the fruitful scientific collaborations and support.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## Funding

The authors are grateful for the financial support from the National Research, Development, and Innovation Office (NKFIH, Grant No. KKP126451). The research reported in this paper is part of project BME-EGA-02, implemented with the support provided by the Ministry of Innovation and Technology of Hungary from the National Research, Development and Innovation Fund, financed under the TKP2021 funding scheme. The computing time granted on the Hungarian HPC Infrastructure at NIIF Institute, Hungary, and DECI resources Kay with support from the PRACE aisbl are gratefully acknowledged.

## Supplementary material

Tabulated data plotted in Figures 1 and 2 for the CCSD and (T) interaction energies of (AcOH)<sub>2</sub>. The (T) IE contributions of CCSD(F12\*)(T) for all S66 dimers analogously to the bottom panel of Figure 4. Moreover, s66\_14kGOLD.xls contains the new 14k-GOLD reference values and all the new benchmark data computed at the CCSD(T)/haXZ (X = T, Q) and CCSD(F12\*)(T)/haXZ-F12 (X = D, T) level of theory.

## ORCID

Péter R. Nagy  <http://orcid.org/0000-0001-6692-0879>  
 László Gyevi-Nagy  <http://orcid.org/0000-0002-4757-3098>  
 Balázs D. Lőrincz  <http://orcid.org/0000-0002-7573-8099>  
 Mihály Kállay  <http://orcid.org/0000-0003-1080-6625>

## References

- [1] J. Řezáč and P. Hobza, *Chem. Rev.* **116**, 5038 (2016). doi:10.1021/acs.chemrev.5b00526
- [2] M. Kodrycka and K. Patkowski, *J. Chem. Phys.* **151**, 070901 (2019). doi:10.1063/1.5116151
- [3] E.G. Hohenstein and C.D. Sherrill, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2**, 304 (2012). doi:10.1002/wcms.84
- [4] Y.S. Al-Hamdani and A. Tkatchenko, *J. Chem. Phys.* **150**, 010901 (2019). doi:10.1063/1.5075487
- [5] K. Szalewicz, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2**, 254 (2012). doi:10.1002/wcms.86
- [6] K. Patkowski, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **10**, e1452 (2020). doi:10.1002/wcms.1452
- [7] K. Carter-Fenk, K.U. Lao and J.M. Herbert, *Acc. Chem. Res.* **54**, 3679 (2021). doi:10.1021/acs.accounts.1c00387
- [8] M. Dubecký, L. Mitás and P. Jurečka, *Chem. Rev.* **116**, 5188 (2016). doi:10.1021/acs.chemrev.5b00577
- [9] M. Dubecký, P. Jurečka, R. Derian, P. Hobza, M. Otyepka and L. Mitás, *J. Chem. Theory Comput.* **9**, 4287 (2013). doi:10.1021/ct4006739
- [10] A. Zen, J.G. Brandenburg, A. Michaelides and D. Alfè, *J. Chem. Phys.* **151**, 134105 (2019). doi:10.1063/1.5119729
- [11] M. Dubecký, P. Jurečka, L. Mitás, M. Ditte and R. Fanta, *J. Chem. Theory Comput.* **15**, 3552 (2019). doi:10.1021/acs.jctc.9b00096
- [12] A. Heßelmann, in *Non-Covalent Interactions in Quantum Chemistry and Physics*, edited by A.O. de la Roza and G.A. DiLabio (Elsevier, Amsterdam, 2017), p. 65.
- [13] B.D. Nguyen, G.P. Chen, M.M. Agee, A.M. Burow, M.P. Tang and F. Furche, *J. Chem. Theory Comput.* **16**, 2258 (2020). doi:10.1021/acs.jctc.9b01176
- [14] R.J. Bartlett and M. Musiał, *Rev. Mod. Phys.* **79**, 291 (2007). doi:10.1103/RevModPhys.79.291
- [15] K. Raghavachari, G.W. Trucks, J.A. Pople and M. Head-Gordon, *Chem. Phys. Lett.* **157**, 479 (1989). doi:10.1016/S0009-2614(89)87395-6
- [16] J. Řezáč and P. Hobza, *J. Chem. Theory Comput.* **9**, 2151 (2013). doi:10.1021/ct400057w
- [17] J. Řezáč, K.E. Riley and P. Hobza, *J. Chem. Theory Comput.* **7**, 2427 (2011). doi:10.1021/ct2002946
- [18] R. Sedlak, T. Janowski, M. Pitoňák, J. Řezáč, P. Pulay and P. Hobza, *J. Chem. Theory Comput.* **9**, 3364 (2013). doi:10.1021/ct400036b
- [19] S. Grimme, *Chem. Eur. J.* **18**, 9955 (2012). doi:10.1002/chem.v18.32
- [20] K. Kříž and J. Řezáč, *J. Chem. Inf. Model.* **60**, 1453 (2020). doi:10.1021/acs.jcim.9b01171
- [21] A.G. Donchev, A.G. Taube, E. Decolvenaere, C. Hargus, R.T. McGibbon, K.H. Law, B.A. Gregersen, J.L. Li, K. Palmo, K. Siva and M. Bergdorf, *Sci. Data* **8**, 1 (2021). doi:10.1038/s41597-021-00833-x
- [22] Z.M. Sparrow, B.G. Ernst, P.T. Joo, K.U. Lao and R.A. DiStasio Jr., *J. Chem. Phys.* **155**, 184303 (2021). doi:10.1063/5.0068862
- [23] J. Řezáč, *J. Chem. Theory Comput.* **16**, 2355 (2020). doi:10.1021/acs.jctc.9b01265
- [24] J. Řezáč, K.E. Riley and P. Hobza, *J. Chem. Theory Comput.* **7**, 3466 (2011). doi:10.1021/ct200523a
- [25] B. Brauer, M.K. Kesharwani, S. Kozuch and J.M.L. Martin, *Phys. Chem. Chem. Phys.* **18**, 20905 (2016). doi:10.1039/C6CP00688D
- [26] M.K. Kesharwani, A. Karton, N. Sylvetsky and J.M.L. Martin, *Aust. J. Chem.* **71**, 238 (2018). doi:10.1071/CH17588
- [27] Y.S. Al-Hamdani, P.R. Nagy, D. Barton, M. Kállay, J.G. Brandenburg and A. Tkatchenko, *Nat. Commun.* **12**, 3927 (2021). doi:10.1038/s41467-021-24119-3
- [28] A. Asadchev and M.S. Gordon, *J. Chem. Theory Comput.* **9**, 3385 (2013). doi:10.1021/ct400054m
- [29] V.M. Anisimov, G.H. Bauer, K. Chadalavada, R.M. Olson, J.W. Glenski, W.T.C. Kramer, E. Aprà and K. Kowalski, *J. Chem. Theory Comput.* **10**, 4307 (2014). doi:10.1021/ct500404c
- [30] M. Pitoňák, F. Aquilante, P. Hobza, P. Neogrády, J. Noga and M. Urban, *Collect. Czech. Chem. Commun.* **76**, 713 (2011). doi:10.1135/cccc2011048
- [31] T. Janowski and P. Pulay, *J. Chem. Theory Comput.* **4**, 1585 (2008). doi:10.1021/ct800142f
- [32] E. Deumens, V.F. Lotrich, A. Perera, M.J. Ponton, B.A. Sanders and R.J. Bartlett, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **1**, 895 (2011). doi:10.1002/wcms.77
- [33] C. Peng, J.A. Calvin and E.F. Valeev, *Int. J. Quantum Chem.* **119**, e25894 (2019). doi:10.1002/qua.v119.12
- [34] T. Shen, Z. Zhu, I.Y. Zhang and M. Scheffler, *J. Chem. Theory Comput.* **15**, 4721 (2019). doi:10.1021/acs.jctc.8b01294
- [35] F. Hummel, T. Tsatsoulis and A. Grüneis, *J. Chem. Phys.* **146**, 124105 (2017). doi:10.1063/1.4977994
- [36] P.R. Nagy, G. Samu and M. Kállay, *J. Chem. Theory Comput.* **14**, 4193 (2018). doi:10.1021/acs.jctc.8b00442
- [37] L. Gyevi-Nagy, M. Kállay and P.R. Nagy, *J. Chem. Theory Comput.* **16**, 366 (2020). doi:10.1021/acs.jctc.9b00957
- [38] L. Gyevi-Nagy, M. Kállay and P.R. Nagy, *J. Chem. Theory Comput.* **17**, 860 (2021). doi:10.1021/acs.jctc.0c01077
- [39] T.H. Dunning Jr., *J. Chem. Phys.* **90**, 1007 (1989). doi:10.1063/1.456153
- [40] T. Helgaker, W. Klopper, H. Koch and J. Noga, *J. Chem. Phys.* **106**, 9639 (1997). doi:10.1063/1.473863
- [41] S.F. Boys and F. Bernardi, *Mol. Phys.* **19**, 553 (1970). doi:10.1080/00268977000101561
- [42] L.A. Burns, M.S. Marshall and C.D. Sherrill, *J. Chem. Theory Comput.* **10**, 49 (2014). doi:10.1021/ct400149j
- [43] M.S. Marshall, L.A. Burns and C.D. Sherrill, *J. Chem. Phys.* **135**, 194102 (2011). doi:10.1063/1.3659142
- [44] S. Alessandrini, V. Barone and C. Puzzarini, *J. Chem. Theory Comput.* **16**, 988 (2019). doi:10.1021/acs.jctc.9b01037
- [45] A.E. DePrince and C.D. Sherrill, *J. Chem. Theory Comput.* **9**, 2687 (2013). doi:10.1021/ct400250u
- [46] W. Klopper, F.R. Manby, S. Ten-no and E.F. Valeev, *Int. Rev. Phys. Chem.* **25**, 427 (2006). doi:10.1080/01442350600799921
- [47] C. Hättig, W. Klopper, A. Köhn and D.P. Tew, *Chem. Rev.* **112**, 4 (2012). doi:10.1021/cr200168z
- [48] H.J. Werner, T.B. Adler, G. Knizia and F.R. Manby, in *Recent Progress in Coupled Cluster Methods*, edited by P.



- Čársky, J. Paldus, and J. Pittner (Springer-Verlag, Heidelberg, 2010). pp. 573.
- [49] L. Kong, F.A. Bischoff and E.F. Valeev, *Chem. Rev.* **112**, 75 (2012). doi:10.1021/cr200204r
- [50] A. Grüneis, S. Hirata, Y.y. Ohnishi and S. Ten-no, *J. Chem. Phys.* **146**, 080901 (2017). doi:10.1063/1.4976974
- [51] T.B. Adler, G. Knizia and H.J. Werner, *J. Chem. Phys.* **127**, 221106 (2007). doi:10.1063/1.2817618
- [52] G. Knizia, T.B. Adler and H.J. Werner, *J. Chem. Phys.* **130**, 054104 (2009). doi:10.1063/1.3054300
- [53] M. Torheyden and E.F. Valeev, *Phys. Chem. Chem. Phys.* **10**, 3410 (2008). doi:10.1039/b803620a
- [54] E.F. Valeev and T.D. Crawford, *J. Chem. Phys.* **128**, 244113 (2008). doi:10.1063/1.2939577
- [55] C. Hättig, D.P. Tew and A. Köhn, *J. Chem. Phys.* **132**, 231102 (2010). doi:10.1063/1.3442368
- [56] R.A. Bachorz, F.A. Bischoff, A. Glöss, C. Hättig, S. Höfener, W. Klopper and D.P. Tew, *J. Comput. Chem.* **32**, 2492 (2011). doi:10.1002/jcc.v32.11
- [57] M.K. Kesharwani, N. Sylvetsky, A. Köhn, D.P. Tew and J.M.L. Martin, *J. Chem. Phys.* **149**, 154109 (2018). doi:10.1063/1.5048665
- [58] B. Brauer, M.K. Kesharwani and J.M.L. Martin, *J. Chem. Theory Comput.* **10**, 3791 (2014). doi:10.1021/ct500513b
- [59] D.A. Sirianni, L.A. Burns and C.D. Sherrill, *J. Chem. Theory Comput.* **13**, 86 (2017). doi:10.1021/acs.jctc.6b00797
- [60] L.A. Burns, M.S. Marshall and C.D. Sherrill, *J. Chem. Phys.* **141**, 234111 (2014). doi:10.1063/1.4903765
- [61] O. Marchetti and H.J. Werner, *J. Phys. Chem. A* **113**, 11580 (2009). doi:10.1021/jp9059467
- [62] M. Kállay, R.A. Horváth, L. Gyevi-Nagy and P.R. Nagy, *J. Chem. Phys.* **155**, 034107 (2021). doi:10.1063/5.0057426
- [63] R. Zaleśny, M. Papadopoulos, P. Mezey and J. Leszczynski, *Linear-Scaling Techniques in Computational Chemistry and Physics: Methods and Applications*, Challenges and Advances in Computational Chemistry and Physics, Springer, Netherlands, 2011.
- [64] R.M. Richard, K.U. Lao and J.M. Herbert, *Acc. Chem. Res.* **47**, 2828 (2014). doi:10.1021/ar500119q
- [65] P. Pulay, *Chem. Phys. Lett.* **100**, 151 (1983). doi:10.1016/0009-2614(83)80703-9
- [66] G.E. Scuseria and P.Y. Ayala, *J. Chem. Phys.* **111**, 8330 (1999). doi:10.1063/1.480174
- [67] Y. Jung, Y. Shao and M. Head-Gordon, *J. Comput. Chem.* **28**, 1953 (2007). doi:10.1002/(ISSN)1096-987X
- [68] F.H. Bangerter, M. Glasbrenner and C. Ochsenfeld, *J. Chem. Theory Comput.* **17**, 211 (2021). doi:10.1021/acs.jctc.0c00934
- [69] M. Del Ben, J. Hutter and J. VandeVondele, *J. Chem. Theory Comput.* **8**, 4177 (2012). doi:10.1021/ct300531w
- [70] G. Schmitz, C. Hättig and D.P. Tew, *Phys. Chem. Chem. Phys.* **16**, 22167 (2014). doi:10.1039/C4CP03502J
- [71] J.J. Eriksen, P. Baudin, P. Ettenhuber, K. Kristensen, T. Kjærgaard and P. Jørgensen, *J. Chem. Theory Comput.* **11**, 2984 (2015). doi:10.1021/acs.jctc.5b00086
- [72] A.D. Findlater, F. Zahariev and M.S. Gordon, *J. Phys. Chem. A* **119**, 3587 (2015). doi:10.1021/jp509266g
- [73] T. Ishikawa, K. Sakakura and Y. Mochizuki, *J. Comput. Chem.* **39**, 1970 (2018). doi:10.1002/jcc.v39.24
- [74] Y. Jin and R.J. Bartlett, *J. Phys. Chem. A* **123**, 371 (2018). doi:10.1021/acs.jpca.8b07947
- [75] W. Li, H. Dong, J. Ma and S. Li, *Acc. Chem. Res.* **54**, 169 (2020). doi:10.1021/acs.accounts.0c00580
- [76] D.B. Krisiloff, C.M. Krauter, F.J. Ricci and E.A. Carter, *J. Chem. Theory Comput.* **11**, 5242 (2015). doi:10.1021/acs.jctc.5b00762
- [77] W. Li and P. Piecuch, *J. Phys. Chem. A* **114**, 6721 (2010). doi:10.1021/jp1038738
- [78] Y. Mochizuki, K. Yamashita, T. Nakano, Y. Okiyama, K. Fukuzawa, N. Taguchi and S. Tanaka, *Theor. Chem. Acc.* **130**, 515 (2011). doi:10.1007/s00214-011-1036-3
- [79] W. Li and S. Li, *J. Chem. Phys.* **121**, 6649 (2004). doi:10.1063/1.1792051
- [80] M. Kobayashi and H. Nakai, *J. Chem. Phys.* **131**, 114108 (2009). doi:10.1063/1.3211119
- [81] W. Li, Z. Ni and S. Li, *Mol. Phys.* **114**, 1447 (2016). doi:10.1080/00268976.2016.1139755
- [82] J. Friedrich and K. Walczak, *J. Chem. Theory Comput.* **9**, 408 (2013). doi:10.1021/ct300938w
- [83] O. Masur, M. Schütz, L. Maschio and D. Usvyat, *J. Chem. Theory Comput.* **12**, 5145 (2016). doi:10.1021/acs.jctc.6b00651
- [84] M. Feldt and R.A. Mata, *J. Chem. Theory Comput.* **14**, 5192 (2018). doi:10.1021/acs.jctc.8b00727
- [85] M. Schütz, J. Yang, G.K.L. Chan, F.R. Manby and H.J. Werner, *J. Chem. Phys.* **138**, 054109 (2013). doi:10.1063/1.4789415
- [86] D. Usvyat, L. Maschio and M. Schütz, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **8**, e1357 (2018).
- [87] Y. Guo, C. Riplinger, U. Becker, D.G. Liakos, Y. Minenkov, L. Cavallo and F. Neese, *J. Chem. Phys.* **148**, 011101 (2018). doi:10.1063/1.5011798
- [88] M. Bensberger and J. Neugebauer, *J. Chem. Phys.* **155**, 224102 (2021). doi:10.1063/5.0071347
- [89] Q. Ma and H.J. Werner, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **8**, e1371 (2018). doi:10.1002/wcms.1371
- [90] Z. Rolik, L. Szegedy, I. Ladjánszki, B. Ládóczy and M. Kállay, *J. Chem. Phys.* **139**, 094105 (2013). doi:10.1063/1.4819401
- [91] P.R. Nagy and M. Kállay, *J. Chem. Theory Comput.* **15**, 5275 (2019). doi:10.1021/acs.jctc.9b00511
- [92] P.B. Szabó, J. Csóka, M. Kállay and P.R. Nagy, *J. Chem. Theory Comput.* **17**, 2886 (2021). doi:10.1021/acs.jctc.1c00093
- [93] P.O. Löwdin, *Phys. Rev.* **97**, 1474 (1955). doi:10.1103/PhysRev.97.1474
- [94] A.G. Taube and R.J. Bartlett, *Collect. Czech. Chem. Commun.* **70**, 837 (2005). doi:10.1135/cccc20050837
- [95] A.E. DePrince and C.D. Sherrill, *J. Chem. Theory Comput.* **9**, 293 (2013). doi:10.1021/ct300780u
- [96] P.R. Nagy, L. Gyevi-Nagy and M. Kállay, *Mol. Phys.* **119**, e1963495 (2021). doi:10.1080/00268976.2021.1963495
- [97] G. Schmitz and C. Hättig, *J. Chem. Phys.* **145**, 234107 (2016). doi:10.1063/1.4972001
- [98] G. Schmitz and C. Hättig, *J. Chem. Theory Comput.* **13**, 2623 (2017). doi:10.1021/acs.jctc.7b00180
- [99] Q. Ma and H.J. Werner, *J. Chem. Theory Comput.* **15**, 1044 (2019). doi:10.1021/acs.jctc.8b01098

- [100] C. Riplinger, P. Pinski, U. Becker, E.F. Valeev and F. Neese, *J. Chem. Phys.* **144**, 024109 (2016). doi:10.1063/1.4939030
- [101] A. Altun, F. Neese and G. Bistoni, *J. Chem. Theory Comput.* **16**, 6142 (2020). doi:10.1021/acs.jctc.0c00344
- [102] A. Benali, H. Shin and O. Heinonen, *J. Chem. Phys.* **153**, 194113 (2020). doi:10.1063/5.0026275
- [103] J.L. Chen, T. Sun, Y.B. Wang and W. Wang, *J. Comput. Chem.* **41**, 1252 (2020). doi:10.1002/jcc.v41.13
- [104] A.S. Christensen, M. Elstner and Q. Cui, *J. Chem. Phys.* **143**, 084123 (2015). doi:10.1063/1.4929335
- [105] J. Calbo, E. Orti, J.C. Sancho-García and J. Aragó, *J. Chem. Theory Comput.* **11**, 932 (2015). doi:10.1021/acs.jctc.5b00002
- [106] F. Ballesteros, S. Dunivan and K.U. Lao, *J. Chem. Phys.* **154**, 154104 (2021). doi:10.1063/5.0042906
- [107] V. Wineman-Fisher, Y. Al-Hamdani, P.R. Nagy, A. Tkatchenko and S. Varma, *J. Chem. Phys.* **153**, 094115 (2020). doi:10.1063/5.0022058
- [108] V. Wineman-Fisher, J.M. Delgado, P.R. Nagy, E. Jakobsson, S.A. Pandit and S. Varma, *J. Chem. Phys.* **153**, 104113 (2020). doi:10.1063/5.0022060
- [109] S. Rahman, V. Wineman-Fisher, P.R. Nagy, Y. Al-Hamdani, A. Tkatchenko and S. Varma, *Chem. Eur. J.* **27**, 11005 (2021). doi:10.1002/chem.v27.42
- [110] I. Sandler, S. Sharma, B. Chan and J. Ho, *J. Phys. Chem. A* **125**, 9838 (2021). doi:10.1021/acs.jpca.1c06648
- [111] B. Chan, W. Dawson and T. Nakajima, *J. Phys. Chem. A* **126**, 2119 (2022). doi:10.1021/acs.jpca.2c00308
- [112] N. Sylvestsky, A. Banerjee, M. Alonso and J.M.L. Martin, *J. Chem. Theory Comput.* **16**, 3641 (2020). doi:10.1021/acs.jctc.0c00297
- [113] I. Efremenko and J.M.L. Martin, *J. Phys. Chem. A* **125**, 8987 (2021). doi:10.1021/acs.jpca.1c05124
- [114] E. Paulechka and A. Kazakov, *J. Chem. Theory Comput.* **14**, 5920 (2018). doi:10.1021/acs.jctc.8b00593
- [115] E. Semidalas and J.M.L. Martin, *J. Chem. Theory Comput.* **18**, 883 (2022). doi:10.1021/acs.jctc.1c01126
- [116] M.J.O. Deegan and P.J. Knowles, *Chem. Phys. Lett.* **227**, 321 (1994). doi:10.1016/0009-2614(94)00815-9
- [117] D.P. Tew, W. Klopper, C. Neiss and C. Hättig, *Phys. Chem. Chem. Phys.* **9**, 1921 (2007). doi:10.1039/B617230J
- [118] C. Peng, J.A. Calvin, F. Pavošević, J. Zhang and E.F. Valeev, *J. Phys. Chem. A* **120**, 10231 (2016). doi:10.1021/acs.jpca.6b10150
- [119] N. Sylvestsky, M.K. Kesharwani and J.M.L. Martin, *J. Chem. Phys.* **147**, 134106 (2017). doi:10.1063/1.4998332
- [120] W. Li, P. Piecuch, J.R. Gour and S. Li, *J. Chem. Phys.* **131**, 114109 (2009). doi:10.1063/1.3218842
- [121] J. Friedrich and M. Dolg, *J. Chem. Theory Comput.* **5**, 287 (2009). doi:10.1021/ct800355e
- [122] Z. Rolik and M. Kállay, *J. Chem. Phys.* **135**, 104111 (2011). doi:10.1063/1.3632085
- [123] M. Kállay, *J. Chem. Phys.* **142**, 204105 (2015). doi:10.1063/1.4921542
- [124] P.R. Nagy, G. Samu and M. Kállay, *J. Chem. Theory Comput.* **12**, 4897 (2016). doi:10.1021/acs.jctc.6b00732
- [125] P.R. Nagy and M. Kállay, *J. Chem. Phys.* **146**, 214106 (2017). doi:10.1063/1.4984322
- [126] D. Mester, P.R. Nagy and M. Kállay, *J. Chem. Theory Comput.* **15**, 6111 (2019). doi:10.1021/acs.jctc.9b00735
- [127] M. Kállay, P.R. Nagy, D. Mester, Z. Rolik, G. Samu, J. Csontos, J. Csóka, P.B. Szabó, L. Gyevi-Nagy, B. Hégyel, I. Ladjánszki, L. Szegedy, B. Ladóczki, K. Petrov, M. Farkas, P.D. Mezei and Á. Ganyecz, *J. Chem. Phys.* **152**, 074107 (2020). doi:10.1063/1.5142048
- [128] M. Kállay, P.R. Nagy, D. Mester, L. Gyevi-Nagy, J. Csóka, P.B. Szabó, Z. Rolik, G. Samu, J. Csontos, B. Hégyel, Á. Ganyecz, I. Ladjánszki, L. Szegedy, B. Ladóczki, K. Petrov, M. Farkas, P.D. Mezei and R.A. Horváth, *Mrcc, a Quantum Chemical Program Suite*. See <https://www.mrcc.hu/> (Accessed March 18, 2022).
- [129] R.A. Kendall, T.H. Dunning Jr. and R.J. Harrison, *J. Chem. Phys.* **96**, 6796 (1992). doi:10.1063/1.462569
- [130] K.A. Peterson, T.B. Adler and H.J. Werner, *J. Chem. Phys.* **128**, 084102 (2008). doi:10.1063/1.2831537
- [131] K.E. Yousaf and K.A. Peterson, *J. Chem. Phys.* **129**, 184108 (2008). doi:10.1063/1.3009271
- [132] K.E. Yousaf and K.A. Peterson, *Chem. Phys. Lett.* **476**, 303 (2009). doi:10.1016/j.cplett.2009.06.003
- [133] R.A. Shaw and J.G. Hill, *J. Chem. Theory Comput.* **13**, 1691 (2017). doi:10.1021/acs.jctc.7b00140
- [134] F. Weigend, *J. Comput. Chem.* **29**, 167 (2008). doi:10.1002/(ISSN)1096-987X
- [135] F. Weigend, A. Köhn and C. Hättig, *J. Chem. Phys.* **116**, 3175 (2002). doi:10.1063/1.1445115
- [136] C. Hättig, *Phys. Chem. Chem. Phys.* **7**, 59 (2005). doi:10.1039/B415208E
- [137] A. Karton and J.M.L. Martin, *Theor. Chem. Acc.* **115**, 330 (2006). doi:10.1007/s00214-005-0028-6
- [138] F. Neese and E.F. Valeev, *J. Chem. Theory Comput.* **7**, 33 (2011). doi:10.1021/ct100396y
- [139] F. Neese, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **8**, e1327 (2018). doi:10.1002/wcms.1327
- [140] D.G. Liakos, M. Sparta, M.K. Kesharwani, J.M.L. Martin and F. Neese, *J. Chem. Theory Comput.* **11**, 1525 (2015). doi:10.1021/ct501129s
- [141] N. Mardirossian and M. Head-Gordon, *Mol. Phys.* **115**, 2315 (2017). doi:10.1080/00268976.2017.1333644
- [142] L. Goerigk, A. Hansen, C. Bauer, S. Ehrlich, A. Najibi and S. Grimme, *Phys. Chem. Chem. Phys.* **19**, 32184 (2017). doi:10.1039/C7CP04913G
- [143] L. Cheng, M. Welborn, A.S. Christensen and T.F. Miller, *J. Chem. Phys.* **150**, 131103 (2019). doi:10.1063/1.5088393
- [144] R.T. McGibbon, A.G. Taube, A.G. Donchev, K. Siva, F. Hernández, C. Hargus, K.H. Law, J.L. Klepeis and D.E. Shaw, *J. Chem. Phys.* **147**, 161725 (2017). doi:10.1063/1.4986081
- [145] A.P. Bartók, S. De, C. Poelking, N. Bernstein, J.R. Kermode, G. Csányi and M. Ceriotti, *Sci. Adv.* **3**, e1701816 (2017). doi:10.1126/sciadv.1701816
- [146] T. Nuddejima, Y. Ikabata, J. Seino, T. Yoshikawa and H. Nakai, *J. Chem. Phys.* **151**, 024104 (2019). doi:10.1063/1.5100165
- [147] S. Chmiela, H.E. Sauceda, K.R. Müller and A. Tkatchenko, *Nat. Commun.* **9**, 3887 (2018). doi:10.1038/s41467-018-06169-2
- [148] P.D. Mezei and A.O. von Lilienfeld, *J. Chem. Theory Comput.* **16**, 2647 (2020). doi:10.1021/acs.jctc.0c00181
- [149] G. Montavon, M. Rupp, V. Gobre, A. Vazquez-Maya goitia, K. Hansen, A. Tkatchenko, M. Klaus-Robert and

- O.A. von Lilienfeld, *New J. Phys.* **15**, 095003 (2013). doi:[10.1088/1367-2630/15/9/095003](https://doi.org/10.1088/1367-2630/15/9/095003)
- [150] R. Ramakrishnan, P.O. Dral, M. Rupp and O.A. von Lilienfeld, *J. Chem. Theory Comput.* **11**, 2087 (2015). doi:[10.1021/acs.jctc.5b00099](https://doi.org/10.1021/acs.jctc.5b00099)
- [151] J.S. Smith, B.T. Nebgen, R. Zubatyuk, N. Lubbers, C. Devereux, K. Barros, S. Tretiak, O. Isayev and A.E. Roitberg, *Nat. Commun.* **10**, 2903 (2019). doi:[10.1038/s41467-019-10827-4](https://doi.org/10.1038/s41467-019-10827-4)