

Minerva Access is the Institutional Repository of The University of Melbourne

**Author/s:**

Chan, B;Goerigk, L;Radom, L

**Title:**

On the inclusion of post-MP2 contributions to double-Hybrid density functionals

**Date:**

2016-01-15

**Citation:**

Chan, B., Goerigk, L. & Radom, L. (2016). On the inclusion of post-MP2 contributions to double-Hybrid density functionals. *Journal of Computational Chemistry*, 37 (2), pp.183-193. <https://doi.org/10.1002/jcc.23972>.

**Persistent Link:**

<https://hdl.handle.net/11343/209008>

# On the Inclusion of Post-MP2 Contributions to Double-Hybrid Density Functionals

Bun Chan,<sup>‡[a]\*</sup> Lars Goerigk,<sup>‡[b]\*</sup> Leo Radom<sup>[a]\*</sup>

<sup>[a]</sup> *School of Chemistry, The University of Sydney and Centre of Excellence for Free Radical Chemistry and Biotechnology, Sydney, NSW 2006, Australia*

<sup>[b]</sup> *School of Chemistry, The University of Melbourne, Parkville, VIC 3010, Australia*

<sup>‡</sup> *Bun Chan and Lars Goerigk are equal contributors.*

In the present study, we explore the effect of supplementing the DuT spin-component-scaled double-hybrid density functional method with post-MP2 correlation terms. We find that the inclusion of additional MP3 correlation energies has almost no effect on the performance. Further addition of correlation effects from MP4 generally leads to a small improvement in the performance. However, we find that the inclusion of the higher-order perturbative correlation effects does not rectify some major shortcomings of DuT for more challenging systems, and the use of MP4 in fact leads to a significant deterioration in the performance in some cases. We also find that the use of correlation energies from CCSD(T) instead of those from MP3 and MP4 does not lead to a substantial improvement over the MP4-based method, both in general and in some difficult cases that we have examined. An additional observation is that, for large systems that are dominated by non-covalent interactions, DuT and the two MP $n$ -based post-MP2 double-hybrid density functional theory procedures all benefit from the inclusion of dispersion corrections. Overall, our investigation suggests that the current generation of MP2-based double-hybrid density functional theory methods may already be providing close to the optimal performance that can be achieved with the double-hybrid methodology paired with spin-component-scaling. Development of even better double hybrids is an active research field and we hope that our study provides valuable insights. We recommend the continuing use of existing MP2-based double-hybrid methods as a bridging level between hybrid density functional procedures and high-level wave-function-based procedures.

## 1. Introduction

Kohn-Sham (KS) density functional theory (DFT)<sup>[1]</sup> has without doubt become one of the most important methodologies in modern quantum chemistry. It is in principle an exact theory, but because the form of the ‘true’ functional remains unknown, we have to rely on approximations to it, and over the past decades a ‘zoo’ of density functionals has emerged (see for example References [2] and [3] for overviews). While systematic improvements in the spirit of the well-known hierarchy of wave-function electron-correlation methods<sup>[4]</sup> are difficult to achieve in DFT development, Perdew’s metaphoric Jacob’s Ladder scheme at least allows the classification of functionals according to their expected accuracy and computational effort, with those on a higher rung of the ladder expected to be closer to the exact solution for a given system.<sup>[5]</sup> The currently two highest rungs (rungs 4 and 5) have become particularly important in the field of computational thermochemistry. Rung 4 includes hybrid-functionals, in which parts of the DFT exchange functional are replaced by a nonlocal Fock-exchange component, well known from Hartree-Fock (HF) theory.<sup>[6]</sup> Rung-5 functionals include nonlocal correlation effects and therefore rely on information from virtual KS orbitals. Numerous ways of including such nonlocal effects are possible, with two of the most prominent representatives being random-phase-approximation (RPA)<sup>[7-10]</sup> methods and double-hybrid density functionals (DHDFs).<sup>[11]</sup>

In the present study, we limit our focus to DHDFs of the form introduced by Grimme with the B2-PLYP functional in 2006.<sup>[12]</sup> From a theoretical point of view, modern DHDFs have brought wave-function-type and DFT methods a step closer to one another, with the exchange and correlation terms including, respectively, a Fock-exchange component –as in hybrid functionals– and a nonlocal correlation component obtained via second-order Møller–Plesset-type theory (MP2),<sup>[13]</sup> as shown in the general form for a DHDF in eq. (1):

$$E^{DHDF} = (1 - a_x)E_X^{DFT} + a_x E_X^{HF} + (1 - a_c)E_C^{DFT} + a_c E_C^{MP2} \quad (1)$$

The  $E_{X/C}^{DFT}$  terms in eq. (1) are conventional exchange and correlation functionals, respectively,  $E_X^{HF}$  is the Fock-exchange component,  $E_C^{MP2}$  the MP2-type correlation contribution, and  $a_x$  and  $a_c$  are the mixing parameters for these nonlocal exchange and

correlation portions. A typical DHDF calculation consists of two steps. First, the self-consistent-field (SCF) equations are solved for the hybrid portion of the functional comprising the first three terms in eq. (1). Subsequently, the resulting KS orbitals are used for the MP2 calculation.

Many DHDFs, such as B2-PLYP, employ Becke-88 (B88)<sup>[14]</sup> exchange and Lee-Yang-Parr (LYP)<sup>[15]</sup> correlation. However, other combinations of  $E_X^{DFT}$  and  $E_C^{DFT}$  have also been explored; see References [11,16,17] for detailed reviews of these methods. The mixing parameters  $a_x$  and  $a_c$  can be derived on theoretical grounds (see for example, Reference [18]), but DHDFs based on empirical parameter fits have been shown to be generally more accurate and robust.<sup>[11,19]</sup> The latest DHDF procedures provide computational chemists with powerful tools that come close to the ‘chemical accuracy’ target of 4 kJ mol<sup>-1</sup> for reaction energies and barrier heights, and 0.4 kJ mol<sup>-1</sup> for noncovalent interaction energies, while still being applicable to systems of reasonable size.<sup>[11,17,20,21]</sup> Moreover, they have been shown to be more generally accurate than MP2-type wave-function procedures.<sup>[21]</sup>

The use of MP2, which can be considered the most economical wave-function electron-correlation method, undoubtedly has the advantage of being relatively computationally efficient. In fact, for medium-sized systems of up to  $\sim 50$  atoms, efficient implementations of the resolution-of-the-identity (RI-MP2) technique<sup>[22]</sup> into most quantum-chemistry codes make DHDFs a valuable alternative to conventional hybrid functionals because of only a marginal increase in computational cost. In wave-function theory, the accuracy of MP2 has been improved by either introducing separate, spin-dependent scaling factors for the electron-pair correlation energies (e.g. SCS-MP2<sup>[23,24]</sup> and SOS-MP2<sup>[25]</sup>), by allowing optimization of MP2 amplitudes and orbital rotations (OO-MP2)<sup>[26,27]</sup>, or by mixing in portions of higher-order MP terms (e.g. SCS-MP3,<sup>[28]</sup> MP2.5,<sup>[29]</sup> or MP3.5<sup>[30]</sup>).

In the context of DHDFs, only the SCS-,<sup>[17-34]</sup> SOS-<sup>[20,35]</sup> and OO-MP2<sup>[36]</sup> ideas have been tested, while the inclusion of higher-order terms has not yet been assessed. Before the advent of DHDFs as defined in eq. (1), Truhlar and co-workers had introduced multi-coefficient methods –sometimes termed “doubly hybrid” approaches– that mix hybrid density functional results with wave-function contributions obtained at the MP $n$  ( $n \geq 2$ )

and coupled-cluster (CC) levels.<sup>[37,38]</sup> However, in these cases the wave-function contributions were obtained conventionally with HF orbitals and energies. Higher-order terms in a DFT framework have also been discussed in the context of “*ab initio* DFT”.<sup>[39]</sup> However, to the best of our knowledge, no assessments have yet been made of DHDFs that follow a definition similar to that of eq. (1) but that additionally contain higher-order correlation terms that are obtained with the help of KS orbitals.

In the present study, we aim to close this gap by systematically examining a series of wave-function-type correlation procedures for their application in a DHDF framework. In particular, we try to answer the question of whether the inclusion of such higher-order terms improves current DHDFs –in a manner similar to the way in which higher-order terms have significantly improved wave-function-based MP2– or if current DHDFs are already close to the upper limit of accuracy for thermochemical applications. To this end, we first describe the theoretical and computational details for the methods employed in the present study. This is followed by analyzing the effect of the inclusion of MP3 and MP4 terms. In Section 3.3, we discuss basis-set effects for these additional post-MP2 components. In Section 3.4, the alternative of using higher-level CCSD(T)-type corrections is investigated, while in Section 3.5 a comparison of our methods with analogous, composite wave-function methods is presented. In Section 3.6, we discuss the inclusion of dispersion effects in the post-MP2 procedures, and we end the discussion with an analysis of larger systems and difficult cases.

We hope that our study provides useful insights regarding DHDFs, that it contributes to our understanding of these procedures, and that it helps with the development of future DFT approximations. In particular, we wish to address the question: how far should one go beyond MP2 in DHDF procedures and still be able to justify the increased cost in terms of improved performance.

## **2. Theoretical and Computational Details**

The number of DHDF procedures has rapidly grown in recent years, and we carefully choose the procedures on which to base our investigation. To establish possible benefits from including higher-order terms in a double-hybrid methodology, it seems logical to start

our investigation with an appropriate and robust DHDF in general thermochemical applications. Two important factors have to be considered in making this choice. Firstly, the mixing parameter for Fock-exchange and nonlocal correlation can either be determined empirically or from theoretical observations. While the latter approach can be insightful from an academic point of view, it has generally been found that an empirical procedure is overall more accurate and robust.<sup>[11,19]</sup> Second, the inclusion of spin-component- or spin-opposite-scaling has been shown to provide an improvement over the standard MP2 contribution.<sup>[17,20,32,34]</sup> In particular, Kozuch and Martin’s DSD-functionals have been found to represent a significant improvement over “first-generation” DHDFs that follow eq. (1). DSD-functionals are described by eq. (2):<sup>[17,32,34]</sup>

$$E^{DSD-DF} = (1 - \alpha_x)E_X^{DFT} + \alpha_x E_X^{HF} + c_C E_C^{DFT} + c_O E_C^{OS-MP2} + c_S E_C^{SS-MP2} + E^{disp}, \quad (2)$$

where  $c_O$  and  $c_S$  are new mixing parameters for the opposite-spin-scaled and same-spin-scaled MP2 nonlocal correlation contributions,  $c_C$  is the scale parameter for the DFT correlation contribution, and  $E^{disp}$  is a London-dispersion contribution. Note that unlike DHDFs obeying eq. (1), the parameter  $c_C$  in eq. (2) is fitted independently of the scaling factors for the perturbative correlation contribution. However, the sum of  $c_C$  and  $c_O$  is usually close to unity,<sup>[17,34]</sup> which indicates that the two parameters remain indirectly coupled.<sup>17</sup> The DSD functionals take into account the fact that DHDFs, like other common DFT approximations, do not fully describe London-dispersion effects and therefore need dispersion corrections.<sup>[11,20,21,40,41]</sup> DSD functionals have employed Grimme’s additive dispersion corrections of the DFT-D type (DFT-D2,<sup>[42]</sup> DFT-D3,<sup>[43]</sup> DFT-D3(BJ))<sup>[44]</sup> during the parameter optimization.

The first functional following eq. (2) employed B88 exchange and LYP correlation (DSD-BLYP).<sup>[32]</sup> In subsequent studies, Kozuch and Martin have tested various combinations of DFT exchange and correlation, and they proposed the use of Perdew–Burke–Ernzerhof (PBE) exchange<sup>[45]</sup> with Perdew-86 (P86) correlation<sup>[46,47]</sup> (DSD-PBEP86),<sup>[34]</sup> or PBE-hole (PBEh) exchange<sup>[48]</sup> with Becke-95 (B95) correlation<sup>[49]</sup> (DSD-PBEhB95).<sup>[17]</sup> In particular, these three functionals were validated against 11 benchmark sets and the anthracene-dimerization reaction.<sup>[17]</sup> Subsequently, Goerigk and Grimme analyzed these functionals

with their extensive GMTKN30 database<sup>[20,50]</sup> of 30 benchmark sets for main-group thermochemistry, kinetics and noncovalent interactions, and they concluded that the average errors over the entire database are very similar for all three DSD-type methods.<sup>[11]</sup> It is noteworthy that Goerigk and Grimme’s study dealt with main-group compounds, while Kozuch and Martin’s study also tested their methods for transition-metal systems. Importantly, both studies agree that DSD-type functionals are superior to lower-rung density functional approximations. In our current study, we base our investigation on a single DSD-type method. However, it is likely that the effects of post-MP2 terms discussed herein will apply equally to the other methods.

Another important consideration is the computational efficiency in relation to the basis-set dependence, which is particularly important because of the use of the computationally more demanding post-MP2 correlation procedures in the present study. To this end, we note that the procedures in the DSD-family are typically parameterized with quadruple- $\zeta$  (QZ)-type basis sets, which can be computationally very demanding in applications to larger systems. Detailed studies on the basis-set dependence of DHDFs have concluded that using a double- $\zeta$  basis set does not lead to a better performance when compared with that of hybrid functionals.<sup>[21]</sup> However, for applications to larger systems use of a triple- $\zeta$  (TZ)-type basis set for DHDFs seems to represent a reasonable compromise.<sup>[21,33,51]</sup> In this context, it is worthwhile to note that DHDFs have a smaller basis-set dependence than the related MP2 wave-function procedure. To try to remedy shortcomings associated with the use of a TZ basis, Chan and Radom have proposed the DuT-D3 procedure,<sup>[33]</sup> which is a refit of DSD-BLYP designed to be used with the moderate-sized A'VTZ+d basis set<sup>[52]</sup> and combined with Grimme’s DFT-D3 dispersion correction with “zero damping”.<sup>[43]</sup> It has been shown to perform as well as DSD-BLYP with a QZ-type basis set for over 800 thermochemical data points covering a diverse range of systems and properties,<sup>[33]</sup> and its good performance has also been demonstrated in a number of other benchmark studies.<sup>[53-55]</sup> We thus deem it to be the best compromise between accuracy and computational efficiency, and we choose this method as our baseline method. The fitted parameters for DuT are  $a_X = 0.70$ ,  $c_C = 0.59$ ,  $c_O = 0.47$  and  $c_S = 0.36$ .

For the present study, we additionally include post-MP2 perturbative terms for our “Higher-order Double-Hybrid Density Functional” (HDH-DF-MP $n$ ) in the following way:

$$\begin{aligned}
 E^{HDH-DF-MPn} &= (1 - a_X)E_X^{DFT} + a_X E_X^{HF} + c_C E_C^{DFT} + c_O E_C^{OS-MP2} + c_S E_C^{SS-MP2} \\
 &+ c_3 \Delta E_C^{MP3} + c_4 \Delta E_C^{MP4},
 \end{aligned} \tag{3}$$

where  $\Delta E_C^{MP3}$  and  $\Delta E_C^{MP4}$  are the third- and fourth-order MP $n$  contributions, with their respective scale factors  $c_3$  and  $c_4$ . Alternatively, we also define such a method with coupled-cluster terms instead of MP $n$  terms:

$$\begin{aligned}
 E^{HDH-DF-CC} &= (1 - a_X)E_X^{DFT} + a_X E_X^{HF} + c_C E_C^{DFT} + c_O E_C^{OS-MP2} + c_S E_C^{SS-MP2} \\
 &+ c_{CC} \Delta E_C^{CCSD} + c_T \Delta E_C^{(T)},
 \end{aligned} \tag{4}$$

where  $\Delta E_C^{CCSD}$  is the difference between the coupled-cluster singles-doubles contribution and the MP2 correlation energy, and  $\Delta E_C^{(T)}$  is the perturbative triples correction used in CCSD(T),<sup>[56]</sup> with the scale parameters  $c_{CC}$  and  $c_T$ , respectively. The total energy according to eqs. (3) and (4) is obtained first by a standard hybrid-DFT calculation, involving the first three terms. The nonlocal correlation contributions are subsequently obtained by using the KS orbitals and amplitudes, according to Grimme’s DHDF idea.<sup>[12]</sup> As already outlined in the Introduction, Truhlar and co-workers previously investigated methods that combined hybrid-DFT contributions with higher-order wave-function terms.<sup>[38]</sup> However, the latter contributions were obtained with HF orbitals, and different basis sets were employed for the various DFT and wave-function contributions. Thus, although higher-order terms, such as coupled cluster, were for the first time connected with DFT energies in this approach, it differs from the subsequent DHDF idea primarily because KS orbitals are not used; differences and parallels between these concepts are discussed thoroughly in Reference [11].

In our discussion, we include those higher-order terms in eqs. (3) and (4) and examine their effects. As the addition of higher-order excitation effects may significantly reduce the need for the inclusion of a dispersion correction, we start our discussion with the

dispersion-uncorrected variant, namely DuT, as our baseline method. The effect of dispersion corrections is discussed later in Section 3.6.

The DHDF calculations were carried out with Gaussian 09,<sup>[57]</sup> while all dispersion corrections were obtained with Grimme’s standalone program.<sup>[58]</sup> Geometries, zero-point vibrational energies, thermal corrections for 298 K enthalpies, and benchmark total energies that are used in the present investigation were taken from previous studies.<sup>[20,33,50,59,60]</sup> Specifically, those for species in the subsets of the E3 test set were taken from References 33, 59 and 60, while those for the subsets of GMTKN30 were taken from References 20 and 50. In the present study, the parameters for the correlation energy components in eq. (3) and (4) were fitted to the G2/97 training set.<sup>[61]</sup> A preliminary test on larger systems did not change the overall trends in the values of the parameters, and we decided to use the G2/97 set, which includes moderately-sized molecules and is well-balanced in its composition. Following previous practice,<sup>[62-66]</sup> the optimization procedure for the correlation energy terms involved minimizing the average of the mean absolute deviation (MAD) from benchmark values and the standard deviation (SD) of these deviations. Parameters for the dispersion correction were determined using a modified version of the DFT-D3 training set (see Section 3.6 for more details).<sup>[43]</sup> All relative energies are reported in  $\text{kJ mol}^{-1}$ .

### 3. Results and Discussion

*3.1 The Use of An Additional MP3 Term.* We begin our discussion with an analysis of the effect of adding post-MP2 components to DuT. We will first focus on the use of the MP3 higher-order perturbative treatments. For the full G2/97 test set, the inclusion of MP3 correlation with an optimized coefficient  $c_3$  statistically yields a performance almost identical to that of DuT, with an MAD from experimental values of  $6.2 \text{ kJ mol}^{-1}$  (Table 1, column B), which may be compared with  $6.3 \text{ kJ mol}^{-1}$  for DuT alone (column A). We note that the coefficient for the MP3 component is very small ( $c_3 = 0.0076$ ). Therefore, the very similar results obtained with and without the MP3 component are not surprising.

**Table 1.** Comparison of scale parameters and performance (mean absolute deviations from experimental values,  $\text{kJ mol}^{-1}$ ) for wave-function-type correlation components in DuT and HDH-DF-MP $n$  (eq. 3) procedures for the G2/97 test set and its subsets.

	A	B	C	D	E
Method <sup>a</sup>	DuT	DuT + $\Delta$ MP3	DuT(reopt) <sup>b</sup> + $\Delta$ MP3	DuT + $\Delta$ MP3 + $\Delta$ MP4	DuT(reopt) <sup>b</sup> + $\Delta$ MP3 + $\Delta$ MP4
$c_0$	0.4737	0.4737	0.4690	0.4737	0.4273
$c_s$	0.3551	0.3551	0.3672	0.3551	0.5334
$c_3$	-	0.0076	0.0091	0.0991	0.2658
$c_4$	-	-	-	0.1386	0.2471
G2/97 <sup>c</sup>	6.3	6.2	6.2	6.0	5.5
$\Delta H_f$	5.0	4.9	4.8	5.3	4.7
IE	8.2	8.2	8.2	7.5	6.9
EA	6.7	6.7	6.7	6.0	5.4
PA	4.7	4.7	4.8	4.6	5.7
<sup>a</sup> See eq. (3). <sup>b</sup> DuT(reopt) signifies DuT with $c_0$ and $c_s$ and post-MP2 coefficients all reoptimized. <sup>c</sup> $\Delta H_f$ = heat of formation, IE = ionization energy, EA = electron affinity and PA = proton affinity.					

As the MP2 parameters in DuT are fitted to accurate experimental benchmark values, the scaled MP2 components may already be approximating higher-order effects implicitly, which may account for the small value for the MP3 component. To investigate such a possibility, we have reoptimized the two MP2 coefficients, namely those for the opposite-spin (OS) and same-spin (SS) components, together with the simultaneous fitting of the MP3 coefficient. We find that the values for all three parameters do not change substantially, with the MP3 coefficient remaining very small (column C). Furthermore, the resulting MAD ( $6.2 \text{ kJ mol}^{-1}$ ) is also very similar to that of DuT itself. It thus appears that the inclusion of MP3 correlation alone is unlikely to lead to a substantial improvement over DuT.

*3.2 The Inclusion of An MP4-Type Correlation Effect.* When additional, higher-level correlation effects at the MP4 level are also included, without reoptimization of the OS and SS MP2 contributions, we again do not notice a substantial improvement in the performance (MAD = 6.0 kJ mol<sup>-1</sup>, Table 1, column D) over DuT itself (6.3 kJ mol<sup>-1</sup>, column A). We observe, however, that the contribution of the MP3 term becomes more significant, with a scale parameter of  $c_3 = 0.0991$ . The factor obtained for the MP4 component is even larger with  $c_4 = 0.1386$ . With the higher-order correlation effects now contributing considerably to the total energy, there is a larger potential for interdependence between the four wave-function-type correlation components. We therefore optimized the scale factors for all of the MP2OS, MP2SS, MP3 and MP4 terms (column E). The resulting method has an MAD of 5.5 kJ mol<sup>-1</sup>, which represents a more notable improvement over DuT and the other three new methods discussed so far.

A closer inspection of the four newly-fitted coefficients of the optimized MP4-type method reveals a curious result. In DuT, DSD-BLYP, SCS-MP2 and the SCS-MP2-related composite method G4(MP2)-6X, the MP2OS parameters are larger than the corresponding MP2SS counterparts. Such a qualitative ordering has been attributed<sup>[23]</sup> to the systematic underestimation of OS-type correlation effects by unscaled MP2, and a corresponding overestimation of SS correlations. However, in the optimized MP4-based DHDF procedure (column E), the MP2OS coefficient is smaller than the MP2SS coefficient.

Such a contrasting observation may seem unphysical for the new MP4-type procedure, and one might suspect that this could be a result of "overfitting" of the four parameters. However, we note that the G4(MP2)-6X procedure contains a total of twelve parameters, with six being scale parameters for the correlation components. As mentioned before, the relative size of the SCS-MP2-type parameters in G4(MP2)-6X is similar to that of SCS-MP2 itself. We therefore deem it not necessarily be the case that the optimized parameters in the MP4-type DHDF procedure are unphysical, and feel that they simply reflect the nature of the combination of the MP2, MP3 and MP4 components. Presumably, the higher-order MP4 excitations have an excessive effect in the opposite direction to that of MP2SS correlations, and as a result the MP2SS parameter is adjusted to become larger than that

for MP2OS. The elucidation of the exact nature of their behavior, however, is beyond the scope of the present study.

*3.3 Basis-Set Effect for the MP3 and MP4 Components.* So far, the results discussed are all based on calculations using a single triple- $\zeta$  basis set, namely A'VTZ+d. The use of such a basis set for MP3 and (especially) MP4 represents a significant increase in the demand for computational resources over that for the MP2-based DuT. It is noteworthy that many useful wave-function-type composite procedures, such as the G4(MP2)-6X method mentioned before, employ progressively smaller basis sets for higher-order terms. In order to minimize the impact of including post-MP2 effects on the computational efficiency, we have examined the use of the smaller A'VDZ+d basis set for the calculation of these terms (Table 2).

To obtain the “composite” DHDF energy, we have carried out two computations for each species. These include the DuT calculation using the default A'VTZ+d basis set, and an analogous MP3- and MP4-type DuT calculation in which the post-MP2 terms were calculated with the A'VDZ+d basis set. In these cases, the coefficients for the MP3 and MP4 components were reoptimized together with simultaneous refitting of the MP2OS and MP2SS parameters. For the optimized MP3-type (columns B and C) and MP4-type (columns D and E) methods, it can be seen that the use of the smaller basis set for the MP3 and MP4 components does not lead to a significant change in the performance. Thus, in our subsequent investigation into the effect of adding even more costly coupled-cluster-type correlation terms to DuT, we calculate the higher-order terms with the A'VDZ+d basis set.

**Table 2.** Comparison of scale parameters and performance (mean absolute deviations from experimental values, kJ mol<sup>-1</sup>) for DuT and HDH-DF-MP<sub>n</sub> procedures [eq. (3)] using A'VDZ+d or A'VTZ+d basis sets for the post-MP2 terms for the G2/97 test set and its subsets.

	A	B	C	D	E
Method <sup>a</sup>	DuT	DuT(reopt) <sup>b</sup> + ΔMP3	DuT(reopt) <sup>b</sup> + ΔMP3	DuT(reopt) <sup>b</sup> + ΔMP3 + ΔMP4	DuT(reopt) <sup>b</sup> + ΔMP3 + ΔMP4
post-MP2 basis set		A'VTZ+d	A'VDZ+d	A'VTZ+d	A'VDZ+d
$c_0$	0.4737	0.4690	0.4706	0.4273	0.4307
$c_5$	0.3551	0.3672	0.3643	0.5334	0.5302
$c_3$	–	0.0091	0.0101	0.2658	0.2464
$c_4$	–	–	–	0.2471	0.2356
G2/97 <sup>c</sup>	6.3	6.2	6.2	5.5	5.5
ΔH <sub>f</sub>	5.0	4.8	4.8	4.7	4.6
IE	8.2	8.2	8.2	6.9	6.9
EA	6.7	6.7	6.7	5.4	5.6
PA	4.7	4.8	4.7	5.7	5.4

<sup>a</sup> See eq. (3). <sup>b</sup> DuT(reopt) signifies DuT with  $c_0$  and  $c_5$  and post-MP2 coefficients all reoptimized. <sup>c</sup> ΔH<sub>f</sub> = heat of formation, IE = ionization energy, EA = electron affinity and PA = proton affinity.

*3.4 The Use of CCSD(T)-Type Corrections.* We now explore the addition of even higher-order excitation terms to DuT, namely the difference between the MP2 contributions and CCSD (ΔCCSD) and the additional perturbative triples correction Δ(T) [see eq. (4)]. Table 3 contains the results for DuT (column A), the related composite MP3- and MP4-type methods (B and C) as defined in the previous section, and the new composite CCSD- and CCSD(T)-based DHDF protocols (D and E). In each case, the coefficients of the MP2 terms, as well as the post-MP2 terms, are optimized. We can see that, in terms of the performance

for the G2/97 test set, the use of coupled cluster does not lead to an improved performance compared with MP4. This comes about as a result of the better performance of the CCSD-type procedures for  $\Delta H_f$  and PA but worse performance for IE and PA. Interestingly, the inclusion of the perturbative triples corrections does not appear to have a noticeable effect, despite the optimized coefficient for the  $\Delta(T)$  term being reasonably significant and comparable in magnitude to that for the  $\Delta\text{CCSD}$  term.

**Table 3.** Comparison of scale parameters and performance (mean absolute deviations from experimental values, kJ mol<sup>-1</sup>) for DuT and post-MP2 DuT-type composite protocols using MP $n$  [eq. (3)] and coupled-cluster procedures [eq. (4)] for the G2/97 test set and its subsets<sup>a,b</sup>

	A	B	C	D	E
Method <sup>a</sup>	DuT	DuT(reopt) <sup>b</sup> + $\Delta\text{MP3}$	DuT(reopt) <sup>b</sup> + $\Delta\text{MP3}$ + $\Delta\text{MP4}$	DuT(reopt) <sup>b</sup> + $\Delta\text{CCSD}$	DuT(reopt) <sup>b</sup> + $\Delta\text{CCSD}$ + $\Delta(T)$
$c_0$	0.4737	0.4706	0.4307	0.4585	0.4625
$c_5$	0.3551	0.3643	0.5302	0.4436	0.4066
$c_3$	–	0.0101	0.2464	–	–
$c_4$	–	–	0.2356	–	–
$c_{CC}$	–	–	–	0.0714	0.0564
$c_T$	–	–	–	–	0.0722
G2/97 <sup>c</sup>	6.3	6.2	5.5	5.6	5.6
$\Delta H_f$	5.0	4.8	4.6	4.2	4.4
IE	8.2	8.2	6.9	7.7	7.7
EA	6.7	6.7	5.6	5.9	5.8
PA	4.7	4.7	5.4	4.9	4.8

<sup>a</sup> See eqs. (3) and (4). <sup>b</sup> DuT(reopt) signifies DuT with  $c_0$  and  $c_5$  and post-MP2 coefficients all reoptimized. <sup>c</sup>  $\Delta H_f$  = heat of formation, IE = ionization energy, EA = electron affinity and PA = proton affinity.

In relation to the comparable performances for the CCSD- and CCSD(T)-type protocols, it is of interest to more closely examine the optimized scale parameters. We can see that the coefficients for the MP2OS components for the two methods are quite similar. However, the MP2SS and the  $\Delta$ CCSD parameters for the CCSD(T)-based method (E) are notably smaller than the corresponding ones for the CCSD-type method (D). Thus, it appears that within this test set, the effect of perturbative triples is approximated reasonably in the CCSD-type protocol by a combination of scaled-up MP2SS and  $\Delta$ CCSD terms. It is noteworthy that the MP2OS coefficients for the CCSD- and CCSD(T)-type methods are larger than the respective MP2SS parameters. This is in contrast to the corresponding relative magnitudes for the MP4-based procedure (column C), but is consistent with typical MP2-based DHDF methods (e.g., column A) as well as the MP3-type protocol (column B). Thus, while  $\Delta$ MP4 has an effect that appears to be opposite to that for MP2SS as discussed earlier,  $\Delta$ CCSD and  $\Delta$ (T) seem to augment MP2 in a more balanced manner.

We also note that, when compared with the standard DuT calculation for the relatively modest-sized  $C_6H_5Me^{•+}$  radical cation (relative CPU time = 1.0 unit), the composite approach with the inclusion of up to fourth-order perturbative contributions using the small A'VDZ+d basis set already more than doubles the relative computer time (2.4 units). For this molecule, the composite approach with the inclusion of CCSD(T) correlation is approximately three times more expensive than DuT itself (3.2 units). As we will discuss in Section 3.6 below, for the larger systems that we examine in the present study, the severe scaling of CCSD(T) has rendered their computations unfeasible in some instances.

*3.5 Comparison with the G4(MP2)-6X Wave-Function-Only Composite Procedure.* The CCSD(T) composite DuT-type procedure discussed in the previous section is, in spirit, similar to wave-function-type composite procedures, which often contain a CCSD(T) component calculated using a small basis set to augment lower-level methods computed in conjunction with larger basis sets. In these composite methods, the MP2 procedure is a fairly popular choice as the low-level method, or one of the low-level methods. For instance, the critical part of the total composite energy for the G4(MP2)-6X procedure is defined as:

$$\begin{aligned}
E[\text{G4}(\text{MP2})\text{-6X}] &= \text{HF/CBS} \\
&+ E_{\text{corr}}[\text{SCS-MP2/G3MP2LargeXP}] \\
&+ \Delta E_{\text{corr}}[\text{scaled-CCSD/6-31G(d)}] \\
&+ \Delta E_{\text{corr}}[\text{scaled-(T)/6-31G(d)}] \\
&+ \text{HLC}
\end{aligned}$$

The sum of the terms HF/CBS and  $E_{\text{corr}}[\text{SCS-MP2/G3MP2LargeXP}]$  can be considered to be an approximation of an SCS-MP2 energy at the complete-basis-set (CBS) limit. The additional scaled CCSD and (T) components are added in order to include higher-order post-MP2 correlation effects but with the smaller 6-31G(d) basis set. The HLC (higher-level correction) is an empirical term intended to compensate for remaining deficiencies.

The CCSD(T) DuT-type procedure is thus similar to G4(MP2)-6X in two principal ways: (1) it has an SCS-MP2 component obtained with a triple- $\zeta$  basis set, and (2) the SCS-MP2 energy is supplemented with a CCSD(T) energy obtained with a double- $\zeta$  basis set, with the CCSD and (T) components scaled separately. Some major differences between the two methods are: (3) the use of a Hartree-Fock reference wave function in G4(MP2)-6X versus the use of hybrid-DFT in the post-MP2 DHDF, (4) the lack of a CBS component in the DHDF, (5) a somewhat different approach for the scaling of the CCSD(T) components, and (6) the lack of an HLC in the composite double-hybrid scheme. Of these similarities and differences, it has been shown already that the HF/CBS component in the G4 procedure, which is related to G4(MP2)-6X, has an effect of approximately 0.5 kJ mol<sup>-1</sup> on the performance. The effects of (5) and (6) on G4(MP2)-6X can be examined straightforwardly, and these will be considered in this section in order to make a more clear-cut comparison with the CCSD(T) DuT-type method.

When the raw G4(MP2)-6X energies without the scaling of the components and the HLC (Table 4, column A) are used, there are large discrepancies for the G2/97 set of heat of formation ( $\Delta H_f$ ) values, with an MAD of 63.7 kJ mol<sup>-1</sup>. The MADs for the ionization energies (IEs, 12.7 kJ mol<sup>-1</sup>) and electron affinities (EAs, 12.8 kJ mol<sup>-1</sup>) are also quite substantial. The standard G4(MP2) procedure (B) includes the HLC, and this largely remedies the deficiency in the underlying composite energies, with an MAD of 4.4 kJ mol<sup>-1</sup> for the

complete test set. Columns C and D show the performance of alternative composite approaches when the HLC is not included, but instead with scaled correlation components. Both of these approaches lead to fairly good performance, with the one having more flexibility in its parameterization (D, 5.1 kJ mol<sup>-1</sup>) slightly outperforming the other one (column C, 6.6 kJ mol<sup>-1</sup>). The G4(MP2)-6X procedure, which utilizes both the HLC and the scaling of all six correlation energy components, performs the best (column E, 4.1 kJ mol<sup>-1</sup>).

**Table 4.** Comparison of the mean absolute deviations (kJ mol<sup>-1</sup>) from experimental values for the G2/97 test set for G4(MP2)-6X (column E) and the related protocols (columns A–D), and for the CCSD(T)-type DHDF (DuT) procedures (columns F and G).

	A	B <sup>a</sup>	C	D	E <sup>b</sup>	F <sup>c</sup>	G
Method	G4(MP2), no HLC	G4(MP2)	G4(MP2), no HLC, 4 scale coeffs	G4(MP2), no HLC, 6 scale coeffs	G4(MP2) -6X, i.e., G4(MP2), 6 scale coeffs	DuT + ΔCCSD + Δ(T), 4 scale coeffs	DuT + ΔCCSD + Δ(T), 6 scale coeffs
Reference	HF	HF	HF	HF	HF	DFT	DFT
HLC	×	✓	×	×	✓	×	✓
Scaling of correlation energy components							
SCS-MP2/sm	×	×	×	✓	✓	×	✓
SCS-MP2/lg	×	×	✓	✓	✓	✓	✓
ΔCCSD, Δ(T)	×	×	✓	✓	✓	✓	✓
Mean absolute deviations							
G2/97	31.7	4.4	6.6	5.1	4.1	5.6	5.0
ΔH <sub>f</sub>	63.7	3.6	5.5	3.4	3.0	4.4	4.0
IE	12.7	4.8	7.5	6.1	4.7	7.7	6.9
EA	12.8	5.3	8.6	6.6	5.3	5.8	5.0
PA	2.4	3.1	2.8	3.3	2.8	4.8	4.6

For the five methods considered in Table 4 so far, the one in column C is closest in its formulation to the CCSD(T) DuT-type procedure (column F). We can see that the DuT-type

protocol (G2/97 MAD = 5.6 kJ mol<sup>-1</sup>) somewhat outperforms the pure-wave-function-type counterpart (column C, 6.6 kJ mol<sup>-1</sup>). One might thus reasonably expect that, if we include additional flexibility and a fully-optimized HLC in the formulation of this composite DHDF, we could arrive at a method that outperforms G4(MP2)-6X. However, we find this not to be the case in general (column G, 5.0 kJ mol<sup>-1</sup>). We have also attempted to further improve the performance of the composite DHDF method by incorporating an additional term for the effect of using a larger quadruple- $\zeta$  basis set for the underlying DFT component. However, this only leads to a marginal improvement of 0.1 kJ mol<sup>-1</sup> in the overall MAD. Presumably, the deviations for the wave-function-type method are more systematic in relation to the specific form of the HLC, and this may be the reason for the larger improvement in the performance in comparison with the DuT-type composite protocol.

*3.6 Inclusion of a Dispersion Correction.* So far we have only discussed systems for which London-dispersion interactions are small and we primarily focused on the effects of adding higher-order terms to the underlying DuT method. However, London-dispersion effects do play a crucial role for larger systems, not only in the description of noncovalent interaction energies but also for general thermochemical problems. In 2007, Schwabe and Grimme pointed out that DHDFs still require dispersion corrections, albeit to a smaller extent than conventional DFT approximations, due to their nonlocal correlation component.<sup>[40]</sup> Various dispersion corrections for DFT methods have been published, see e.g. References [67–69] for reviews. Here we use Grimme’s established and accurate DFT-D3 approach with Becke-Johnson<sup>[70–72]</sup> damping [DFT-D3(BJ)],<sup>[44]</sup> which has been shown to provide accurate noncovalent-interaction energies, thermochemistry and geometries.<sup>[21,41,44,69,73–75]</sup> It is an atom-pairwise approach of the form:

$$E^{DFT-D3(BJ)} = -\frac{1}{2} \sum_{A \neq B} \sum_{n=6,8} s_n \frac{C_n^{AB}}{R_{AB}^n + \left( a_1 \sqrt{C_8^{AB} / C_6^{AB}} + a_2 \right)^n}, \quad (5)$$

where  $R_{AB}$  is the distance between two atoms  $A$  and  $B$ , with  $n = 6$  taking into account the long-range and  $n = 8$  the medium-range behavior of the dispersion energy scaled by the factors  $s_n$ . The dispersion coefficients  $C_n^{AB}$  are based on first principles, while  $a_1$  and  $a_2$  are

parameters influencing the dispersion contribution at short distances at which the energy contribution from the specific density functional approximation also plays a role.

For conventional DFT approximations,  $s_6$  adopts a value of unity to ensure the correct asymptotic long-range behavior. A method that includes a proper description of London-dispersion would have  $s_6 = 0$ . A double-hybrid includes a scaled-down perturbative correlation contribution and its  $s_6$  is typically between 0.4 and about 0.8.<sup>[11]</sup> Goerigk and Grimme suggested a procedure for estimating this contribution from accurate rare-gas dimer interaction energies<sup>20</sup> and we follow the same scheme here for the two composite procedures containing either the  $\Delta$ MP3 correction or the  $\Delta$ MP3 +  $\Delta$ MP4 correction. Because of computational limitations, we were not able to carry out CCSD(T) calculations that are necessary for the determination of the D3(BJ) parameters for the CCSD(T)-based approach at this stage. However, we expect that our findings for the two related methods can be transferred to this approach. The resulting  $s_6$  values are shown in Table 5 and they are 0.45 (for  $\Delta$ MP3) and 0.42 (for  $\Delta$ MP3 +  $\Delta$ MP4). Table 5 also shows the  $s_6$  for Dut-D3(BJ), which was determined for our current study. It is slightly higher, with  $s_6 = 0.50$ . Thus, higher-order contributions, when added in the way we do it here, do not remove the dispersion problem of the DFT approximation. It is therefore expected that these methods still need a dispersion correction in applications to larger systems.

**Table 5.** Optimized values for the DFT-D3(BJ) dispersion coefficients and mean absolute deviations (MADs in kJ mol<sup>-1</sup>) from benchmark values for some noncovalent-interaction benchmark sets for DuT-D3(BJ)/A'VTZ+d and the two composite HDF-DH-MP $n$  procedures.<sup>a</sup>

	A	B	C
Method	DuT	DuT(reopt) <sup>b</sup> + $\Delta$ MP3	DuT(reopt) <sup>b</sup> + $\Delta$ MP3 + $\Delta$ MP4
$c_0$	0.4737	0.4706	0.4307
$c_s$	0.3551	0.3643	0.5302
$c_3$	-	0.0101	0.2464
$c_4$	-	-	0.2356
$s_6$	0.50	0.45	0.42
$a_1$	0.0	0.0	0.0
$s_8$	0.4753	0.4517	0.6999
$a_2$	7.3518	7.0805	8.1296
	Mean absolute deviations		
ACONF	0.8 (1.5)	0.7 (1.5)	0.8 (1.2)
SCONF	1.4 (1.9)	1.3 (1.9)	0.5 (0.7)
CYCONF	0.3 (0.6)	0.2 (0.6)	0.3 (0.4)
RG6	0.1 (1.4)	0.0 (1.4)	0.1 (1.3)
ADIM6	2.1 (7.2)	1.9 (7.1)	1.8 (5.4)
S22 <sup>c</sup>	1.0 (2.8)	1.1 (2.8)	1.2 (1.7)
S22(H-bonds) <sup>d</sup>	0.8 (1.4)	0.9 (1.4)	1.3 (0.8)
S22(dispersion) <sup>e</sup>	1.6 (4.4)	1.7 (4.5)	1.7 (2.7)
S22(mixed) <sup>f</sup>	0.7 (2.5)	0.7 (2.5)	0.7 (1.5)
S22+	0.7 (1.9)	0.7 (1.9)	0.8 (1.2)

<sup>a</sup> Values in parentheses are MADs for the dispersion-uncorrected approaches. <sup>b</sup> DuT(reopt) signifies DuT with  $c_0$  and  $c_s$  and post-MP2 coefficients all reoptimized. <sup>c</sup> Two systems were not included because of computational limitations (see SI). <sup>d</sup> Hydrogen-bonded dimers. <sup>e</sup>Dispersion-bound dimers. <sup>f</sup>Dimers that bind due to hydrogen bonds and dispersion.

The parameters  $a_1$ ,  $s_8$ , and  $a_2$  are the only empirical parameters in eq. (5) and usually they are fitted to a set of around 140 relative energies. This set is made up of the following benchmark sets: ACONF<sup>[76]</sup> for alkane conformations, SCONF<sup>[50,77]</sup> for sugar conformations, CYCONF<sup>[78]</sup> for cysteine conformations, PCONF<sup>[79]</sup> for tripeptide conformations, RG6<sup>[43]</sup> for rare-gas dimers, ADIM6<sup>[43]</sup> for alkane dimers, S22<sup>[80,81]</sup> for noncovalently-bound dimers of various types, and S22+<sup>[43]</sup> for the same systems as S22, only at longer, non-equilibrium distances between the monomers. Because of computational limitations, we had to put to one side the PCONF set, and we carried out the least-squares fitting procedure for the remaining systems. The values for the resulting parameters are shown in Table 5 for the underlying DuT procedure and the two composite HDH-DF-MP $n$  approaches. It turns out that in all three cases, the parameter  $a_1$  is zero, which parallels findings for other DHDFs that include an SCS-MP2 contribution.<sup>[17]</sup>

Table 5 also shows the MADs from benchmark values for the various sets that were part of the training set; the values in parentheses are the results for the dispersion-uncorrected methods. For all three approaches and for all the benchmark sets, we see the clear benefit from adding the dispersion correction, as the deviations are reduced by 50% or more; for CYCONF and RG6 the final deviations are even within the uncertainty range of the *ab initio* reference values. The S22 set is divided into three subsets with dimers that are either stabilized by hydrogen bonds, dispersion interactions or a mixture of both. Table 5 shows the MADs for the three separate groups. For the hydrogen-bonded dimers, adding a dispersion correction improves the MADs of the DuT and the DuT+ $\Delta$ MP3 approaches from 1.4 to 0.8 and 0.9 kJ mol<sup>-1</sup>, respectively, while adding a dispersion correction to DuT+ $\Delta$ MP3+ $\Delta$ MP4 leads to an increase in the MAD (from 0.8 to 1.3 kJ mol<sup>-1</sup>). To better understand this behavior, it is instructive to also analyze the mean deviations (MDs). The first two methods strongly underbind the dimers, with MDs of -1.4 kJ mol<sup>-1</sup> when not dispersion corrected. Adding the fourth-order correction already leads to an overbinding tendency (MD = 0.7 kJ mol<sup>-1</sup>), which is further enhanced by adding the dispersion correction. The same behavior is not observed for the purely dispersion-bound and mixed complexes. In all cases, the MDs are negative and the MADs are larger than for the dispersion-corrected systems.

However, when comparing the three different approaches with one another, no substantial improvement is generally observed from adding the post-MP2 corrections. The only exception is the SCONF data set, for which the method that includes a  $\Delta$ MP4 contribution reduces the DuT deviation by more than 50%. In fact, the MAD of 0.5 kJ mol<sup>-1</sup> is the best value yet reported for this benchmark set. Thus, we can conclude that the post-MP2 procedures benefit from dispersion corrections, but that, with one exception, they do not generally lead to better performance than that of standard DHDFs for relative energies in noncovalently-bound dimers or conformers. In the next section, we will investigate these methods for some additional thermochemical problems.

*3.7 Evaluation for larger systems/difficult cases.* So far, we have established that adding higher-order correlation terms to the DuT method according to either eq. (3) or eq. (4) only has a marginal effect on thermochemical properties of small molecules and on noncovalent interaction energies in small to medium-sized systems. Two more aspects need to be addressed before we can draw our final conclusions. First, we need to test how these methods behave for thermochemical problems involving larger systems, and second, we need to analyze if outliers in a benchmark set are treated better, i.e. if the deviation range is lowered by adding higher-order terms.

For this purpose, we take some examples from the large GMTKN30 database for general main-group thermochemistry, kinetics and noncovalent interactions.<sup>[20,50]</sup> In fact, the benchmark sets for noncovalent interaction energies discussed in the previous section are part of this compilation. However, GMTKN30 also contains a range of benchmark sets for reaction energies and we have investigated three of those that we regard as challenging for a computational method. The first benchmark set we discuss is the DC9 set,<sup>[12,50,82-87]</sup> which contains nine “difficult cases” (see SI or Reference [50] for more details). Due to the size of some of the systems, we had to reduce the set to five systems (DC5) for the coupled-cluster-based approaches. MADs from benchmark values for DC9 and DC5 are shown in Table 6 for DuT/A'VTZ+d (column A), the composite approach containing a  $\Delta$ MP3 correction obtained at the double- $\zeta$  level (column B), the composite approach with an additional  $\Delta$ MP4 correction (column C), the approach containing a  $\Delta$ CCSD correction

(column D) and finally the approach that additionally includes the  $\Delta(T)$  correction (column E). The results in columns A–C are dispersion-corrected (see previous section).

While adding just the  $\Delta MP3$  correction has a minimal effect on the results for both DC9 and DC5 when compared with DuT-D3(BJ), adding the fourth-order  $\Delta MP4$  correction actually increases the MADs by up to 5 kJ mol<sup>-1</sup>. The two main reasons for this are an increase in the absolute deviation for the cage-to-bowl isomerization of C<sub>20</sub> from 3 kJ mol<sup>-1</sup> for DuT-D3(BJ) (column A) to 34.4 kJ mol<sup>-1</sup> (column C), and to a smaller extent for the decomposition of S<sub>8</sub> into four S<sub>2</sub> molecules, with an increase from 7.1 to 15.7 kJ mol<sup>-1</sup>. As the C<sub>20</sub> isomerization is not included in DC5, the increase in the MADs (2.5 kJ mol<sup>-1</sup>) is not as high as for DC9. The coupled-cluster-based approaches show similar results to DuT-D3(BJ) for DC5.

The second test set is a set of 43 radical stabilization energies (RSE43).<sup>[27]</sup> As can be seen from Table 6, none of the post-MP2 methods improve on the original DuT-D3(BJ) result; all MADs are between 3.6 and 4.0 kJ mol<sup>-1</sup>. The final benchmark set taken from GMTKN30 is the DARC set containing 14 Diels-Alder reactions.<sup>88</sup> MADs for DuT-D3(BJ) and the dispersion-corrected, composite HDH-DF-MP $n$  methods are also included in Table 6. Only a small improvement is observed when adding the  $\Delta MP4$  correction, with the MAD decreasing from 7.5 to 7.1 kJ mol<sup>-1</sup>.

**Table 6.** Mean absolute deviations (MADs in kJ mol<sup>-1</sup>) from benchmark values for four benchmark sets, and signed deviations for total atomization energies of two systems for DuT-D3(BJ)/A'VTZ+d, the two composite HDF-DH-MP<sub>n</sub> procedures and the two composite HDH-DF-CC procedures.<sup>a,b</sup>

	A	B	C	D	E
Method	DuT	DuT(reopt) <sup>c</sup> + ΔMP3	DuT(reopt) <sup>c</sup> + ΔMP3 +ΔMP4	DuT(reopt) <sup>c</sup> + ΔCCSD	DuT(reopt) <sup>c</sup> + ΔCCSD + Δ(T)
<i>c<sub>0</sub></i>	0.4737	0.4706	0.4307	0.4585	0.4625
<i>c<sub>S</sub></i>	0.3551	0.3643	0.5302	0.4436	0.4066
<i>c<sub>3</sub></i>	–	0.0101	0.2464	–	–
<i>c<sub>4</sub></i>	–	–	0.2356	–	–
<i>CC</i>	–	–	–	0.0714	0.0564
<i>CT</i>	–	–	–	–	0.0722
Mean absolute deviations					
DC9	11.1	10.9	16.2	–	–
DC5	6.7	6.5	9.2	6.8	7.0
RSE43	3.6	3.6	3.8	3.7	4.0
DARC	7.5	7.0	7.1	–	–
Signed deviations for total atomization energies					
anthracence	35.6	34.7	2.1	–3.1	2.7
adamantane	–7.5	–6.7	–21.3	–31.1	–32.4
<sup>a</sup> DC9 and DC5 represent nine and five difficult cases for DFT, RSE43 are 43 radical stabilization energies, DARC is a database of Diels–Alder reaction energies. <sup>b</sup> The values in columns A–C are dispersion corrected, the values in columns D–E are not. <sup>c</sup> DuT(reopt) signifies DuT with <i>c<sub>0</sub></i> and <i>c<sub>S</sub></i> and post-MP2 coefficients all reoptimized.					

We conclude our analysis with the discussion of the atomization energies of anthracene and adamantane. Goerigk and Grimme included these two systems as part of their parameter training set for the optimized  $\sigma$ TPSS<sup>[50]</sup> meta-GGA and the PWPB95<sup>[20]</sup> double-hybrid functionals and they argued that these two systems pose difficult cases for DFT approximations. The total atomization energies are of the same order of magnitude (between 11,200 and 12,400 kJ mol<sup>-1</sup>). However, none of the tested DFT approximations were found to perform well for both systems at the same time.<sup>[50]</sup> Also, even the best DHDFs were found to describe only one of the systems adequately, while they have a larger error for the other.<sup>[20]</sup> Therefore, it is worthwhile to examine here whether post-MP2 corrections can provide a more balanced description of both systems.

Inspection of Table 6 shows that the signed deviation from the benchmark atomization energies for the DuT-D3(BJ) method is moderate for adamantane (-7.5 kJ mol<sup>-1</sup>) but is large for anthracene (35.6 kJ mol<sup>-1</sup>). The  $\Delta$ MP3 correction (column B) does not significantly change the DuT-D3(BJ) values. However, when the  $\Delta$ MP4 correction (column C) is included, the deviation for anthracene is reduced considerably from 35.6 kJ mol<sup>-1</sup> to 2.1 kJ mol<sup>-1</sup>. Unfortunately, the opposite effect is observed for adamantane: inclusion of the  $\Delta$ MP4 correction leads to a larger underestimation (21.3 kJ mol<sup>-1</sup>) than DuT-D3(BJ) (-7.5 kJ mol<sup>-1</sup>). The use of CCSD(T) correlation energies (columns D and E) leads to still larger deviations for adamantane. Thus, we conclude that this challenging problem of simultaneously obtaining reasonable atomization energies for anthracene and adamantane still persists, and that for these systems adding higher-order correlation terms to the DHDF does not provide any additional robustness.

Overall, the results in Tables 3–6 suggest that the inclusion of post-MP2 correlation components that go beyond MP4 may not justify the associated cost. This contrasts with the case of wave-function-type procedures based on a Hartree–Fock reference, where the performance of CCSD(T) is notably and almost-universally better than that of MP4. We note in this regard that DHDF procedures are affected by the deficiencies in the underlying exchange functional. Although the contribution of this factor to the results of the present study is not entirely clear, it undoubtedly hinders the potential for systematic

improvement of DHDFs in their current form, when one focuses only on the correlation component.

#### 4. Conclusions

In the present study, we have explored the effect of supplementing the DuT double-hybrid density functional (DHDF) method with post-MP2 correlation terms. Several key findings emerge from our investigation.

1. The inclusion of additional correlation energy from MP3 has almost no effect on the performance of the DuT method. Consistent with this observation, we also find that the optimized coefficient for the  $\Delta$ MP3 term is very small, regardless of whether the MP2 parameters are refitted or not.
2. Further addition of MP4 correlation effects to the MP3-based method, with both the  $\Delta$ MP3 and  $\Delta$ MP4 coefficients fitted in conjunction with the default MP2 parameters, leads to a small improvement in the performance for the G2/97 set. A more notable improvement can be seen when all four  $MP_n$  parameters (for MP2OS, MP2SS,  $\Delta$ MP3 and  $\Delta$ MP4) are optimized. The magnitudes of the fitted parameters appear to suggest that the correlation effects captured by the  $\Delta$ MP4 term in this scheme oppose those represented by the MP2SS component.
3. The post-MP2 terms can be included using the aug'-cc-pVDZ+d basis set, which is less computationally demanding than the aug'-cc-pVTZ+d basis set used for the underlying density functional and MP2 components, without a significant deterioration in the performance. This requires refitting of the coefficients for the  $MP_n$  terms.
4. Using the  $\Delta$ CCSD and  $\Delta$ (T) correlation energies from a CCSD(T) calculation with the Kohn–Sham orbitals instead of the  $\Delta$ MP3 and  $\Delta$ MP4 terms does not lead to any notable improvement over the MP4-based method for the G2/97 set. We note that the performance of this CCSD(T)-based DHDF is somewhat better than an analogous wave-function composite procedure, namely G4(MP2)-6X without its empirical correction terms. However, when empirical higher-level corrections are included in both the

density-functional-based and wave-function-based protocols, the wave-function method shows a better performance.

5. We find that, for large systems that are dominated by non-covalent interactions, DuT and the two  $MP_n$ -based post-MP2 DHDF procedures all benefit from the inclusion of D3(BJ) dispersion corrections. The performance of the non-dispersion-corrected methods, and progression of the optimized  $s_6$  parameter for the MP2-, MP3- and MP4-based methods, indicate that the inclusion of higher-order correlations do lead to a (minor) progressive improvement in the description of dispersion interactions, but not sufficient for a sole-reliance on the electronic structure method alone. We also note that the inclusion of the MP4 term leads to an overestimation of binding in hydrogen-bonded dimers.
6. An examination of the performance of the post-MP2-type DHDF methods shows that the inclusion of higher-order correlation effects, in general, does not rectify some of the major shortcomings of DuT for more challenging systems. The use of MP4 in fact leads to a significant deterioration in performance in some cases.

Overall, our investigation suggests that the current generation of MP2-based DHDF methods may already yield close to the optimal performance that can be achieved by mixing density-functional and wave-function components using the general methodology introduced by Grimme in B2-PLYP. The inclusion of post-MP2 perturbative correlation effects may lead to a slight overall improvement, but it can also result in significant deterioration in the performance in some cases. The use of coupled-cluster terms is more robust than stopping at the MP4 level. However, its higher-cost and, more importantly, its seemingly less good performance than analogous wave-function-type composite methods such as G4(MP2)-6X, does not make it an attractive choice. DHDF development is still an active field of research and we hope that our study provides important insights into the addition of higher-order terms. We recommend the continuing use of existing MP2-based DHDF methods (including SCS- or SOS-MP2 contributions) as a bridging level between hybrid-DFT procedures and wave-function-based composite procedures.

## Acknowledgements

LG is the recipient of an Australian Research Council (ARC) Discovery Early Career Researcher Award (project number DE140100550) and of the 2014 Selby Research Award from the Selby Scientific Foundation. We gratefully acknowledge the generous allocation of computing time from the National Computational Infrastructure (NCI) National Facility within the National Computational Merit Allocation Scheme and from Intersect Australia Ltd.

## References

- [1] W. Kohn, L. J. Sham, *Phys. Rev.* **1965**, *140*, A1133.
- [2] R. Peverati, D. G. Truhlar, *Phil. Trans. R. Soc. A* **2013**, *372*, 20120476.
- [3] A. D. Becke, *J. Chem. Phys.* **2014**, *140*, 18A301.
- [4] W. J. Hehre, L. Radom, P. v. R. Schleyer, and J. A. Pople, *Ab Initio Molecular Orbital Theory*; Wiley: New York, 1986.
- [5] J. P. Perdew, K. Schmidt, in *Density Functional Theory and Its Applications to Materials*, AIP Conference Proceedings; V. E. Van Doren, C. Van Alsenoy, P. Geerlings, Eds.; American Institute of Physics: Melville, **2001**; Vol. 577, pp 1–20.
- [6] A. D. Becke, *J. Chem. Phys.* **1993**, *98*, 1372–1377.
- [7] D. C. Langreth, J. P. Perdew *Phys. Rev. B* **1977**, *15*, 2884–2901.
- [8] S. Kurth, J. P. Perdew *Phys. Rev. B* **1999**, *59*, 10461.
- [9] J. Paier, B. G. Janesko, T. M. Henderson, G. E. Scuseria, A. Grüneis, G. Kresse *J. Chem. Phys.* **2010**, *132*, 094103.
- [10] H. Eshuis, J. E. Bates, F. Furche *Theor. Chem. Acc.* **2012**, *131*, 1084.
- [11] L. Goerigk, S. Grimme *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2014**, *4*, 576–600.
- [12] S. Grimme *J. Chem. Phys.* **2006**, *124*, 034108.

- [13] C. Møller, M. S. Plesset *Phys. Rev.* **1934**, *46*, 618–622.
- [14] A. D. Becke *Phys. Rev. A* **1988**, *38*, 3098–3100.
- [15] C. Lee, W. Yang, R. G. Parr *Phys. Rev. B* **1988**, *37*, 785–789.
- [16] J. C. Sancho-García, C. Adamo *Phys. Chem. Chem. Phys.* **2013**, *15*, 14581–14594.
- [17] S. Kozuch, J. M. L. Martin *J. Comput. Chem.* **2013**, *34*, 2327–2344.
- [18] E. Bremond, C. Adamo *J. Chem. Phys.* **2011**, *135*, 024106.
- [19] D. Bousquet, E. Bremond, J. C. Sancho-García, C. Adamo *J. Chem. Theory Comput.* **2013**, *9*, 3444–3452.
- [20] L. Goerigk, S. Grimme *J. Chem. Theory Comput.* **2011**, *7*, 291–309.
- [21] L. Goerigk, S. Grimme *Phys. Chem. Chem. Phys.* **2011**, *13*, 6670–6688.
- [22] F. Weigend, M. Häser *Theor. Chem. Acc.* **1997**, *97*, 331–340.
- [23] S. Grimme *J. Chem. Phys.* **2003**, *118*, 9095–9102.
- [24] S. Grimme, L. Goerigk, R. F. Fink *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 886–906.
- [25] Y. Jung, R. C. Lochan, A. D. Dutoi, M. Head-Gordon *J. Chem. Phys.* **2004**, *121*, 9793–9802.
- [26] R. C. Lochan, M. Head-Gordon *J. Chem. Phys.* **2007**, *126*, 164101.
- [27] F. Neese, T. Schwabe, S. Kossmann, B. Schirmer, S. Grimme *J. Chem. Theory Comput.* **2009**, *5*, 3060–3073.
- [28] S. Grimme *J. Comput. Chem.* **2003**, *24*, 1529–1537.
- [29] M. Pitoňák, P. Neogrady, J. Černý, S. Grimme, P. Hobza *ChemPhysChem* **2009**, *10*, 282–289.

- [30] A. Karton, L. Goerigk *J. Comput. Chem.* **2015**, *36*, 622–632.
- [31] J. D. Chai, M. Head-Gordon *J. Chem. Phys.* **2009**, *131*, 174105.
- [32] S. Kozuch, D. Gruzman, J. M. L. Martin *J. Phys. Chem. C* **2010**, *114*, 20801–20808.
- [33] B. Chan, L. Radom *J. Chem. Theory Comput.* **2011**, *7*, 2852–2863.
- [34] S. Kozuch, J. M. L. Martin *Phys. Chem. Chem. Phys.* **2011**, *13*, 20104–20107.
- [35] T. Benighaus, R. A. DiStasio Jr., R. C. Lochan, R. D. Chai, M. Head-Gordon *J. Phys. Chem. A* **2008**, *112*, 2702–2712.
- [36] R. Peverati, M. Head-Gordon *J. Chem. Phys.* **2013**, *139*, 024110.
- [37] Y. Zhao, B. J. Lynch, D. G. Truhlar *J. Phys. Chem. A* **2004**, *108*, 4786–4791.
- [38] Y. Zhao, B. J. Lynch, D. G. Truhlar *Phys. Chem. Chem. Phys.* **2005**, *7*, 43–52.
- [39] R. J. Bartlett, I. Grabowski, S. Hirata, S. Ivanov *J. Chem. Phys.* **2005**, *122*, 034104.
- [40] T. Schwabe, S. Grimme *Phys. Chem. Chem. Phys.* **2007**, *9*, 3397–3406.
- [41] L. Goerigk, H. Kruse, S. Grimme, *Chem. Phys. Chem.* **2011**, *12*, 3421–3433.
- [42] S. Grimme *J. Comput. Chem.* **2006**, *27*, 1787–1799.
- [43] S. Grimme, J. Antony, S. Ehrlich, H. Krieg *J. Chem. Phys.* **2010**, *132*, 154104.
- [44] S. Grimme, S. Ehrlich, L. Goerigk *J. Comput. Chem.* **2011**, *32*, 1456–1465.
- [45] J. P. Perdew, K. Burke, M. Ernzerhof *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
- [46] J. P. Perdew *Phys. Rev. B* **1986**, *33*, 8822–8824.
- [47] J. P. Perdew *Phys. Rev. B* **1986**, *34*, 7406.
- [48] M. Ernzerhof, J. P. Perdew *J. Chem. Phys.* **1998**, *109*, 3313–3320.
- [49] A. D. Becke *J. Chem. Phys.* **1996**, *104*, 1040–1046.

- [50] L. Goerigk, S. Grimme *J. Chem. Theory Comput.* **2010**, *6*, 107–126.
- [51] A. Karton, J. M. L. Martin *J. Chem. Phys.* **2011**, *135*, 144119.
- [52] Martin, J. M. L.; de Oliveira, G. *J. Chem. Phys.* **1999**, *111*, 1843–1856.
- [53] B. Chan, L. Radom, *J. Phys. Chem. A* **2012**, *116*, 4975–4986.
- [54] B. Chan, L. Radom *J. Phys. Chem. A* **2013**, *117*, 3666–3675.
- [55] B. Chan, G. E. Ball, *J. Chem. Theory Comput.* **2013**, *9*, 2199–2208.
- [56] K. Raghavachari, G. W. Trucks, J. A. Pople, M. Head-Gordon *Chem. Phys. Lett.* **1989**, *157*, 479–483.
- [57] Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R. E.; Stratmann, O.; Yazyev, A. J.; Austin, R.; Cammi, C.; Pomelli, J. W.; Ochterski, R.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, Revision C.01, Gaussian, Inc.: Wallingford CT, 2009.
- [58] S. Grimme, DFT-D3 program, Grimme group website: <http://www.thch.uni-bonn.de/tc/index.php?section=downloads&subsection=DFT-D3&lang=english>.
- [59] B. Chan, P. M. W. Gill, L. Radom *J. Chem. Theory Comput.* **2012**, *8*, 4899–4906.

- [60] B. Chan, A. T. B. Gilbert, P. M. W. Gill, L. Radom *J. Chem. Theory Comput.* **2014**, *10*, 3777–3783.
- [61] L. A. Curtiss, R. C. Redfern, K. Raghavachari, J. A. Pople *J. Chem. Phys.* **1997**, *106*, 1063–1079.
- [62] B. Chan, M. L. Coote, L. Radom *J. Chem. Theory Comput.* **2010**, *6*, 2647–2653.
- [63] B. Chan, J. Deng, L. Radom *J. Chem. Theory Comput.* **2011**, *7*, 112–120.
- [64] B. Chan, L. Radom *J. Chem. Theory Comput.* **2012**, *8*, 4259–4269.
- [65] B. Chan, L. Radom *J. Chem. Theory Comput.* **2013**, *9*, 4769–4778.
- [66] B. Chan, L. Radom *Theor. Chem. Acc.* **2013**, *133*, 1426.
- [67] S. Grimme *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2011**, *1*, 211–228.
- [68] J. Klimes, A. Michaelides *J. Chem. Phys.* **2012**, *137*, 120901.
- [69] L. Goerigk *J. Chem. Theory Comput.* **2014**, *10*, 968–980.
- [70] A. D. Becke, E. R. Johnson *J. Chem. Phys.* **2005**, *123*, 024101.
- [71] E. R. Johnson, A. D. Becke *J. Chem. Phys.* **2005**, *123*, 154101.
- [72] E. R. Johnson, A. D. Becke *J. Chem. Phys.* **2006**, *124*, 174104.
- [73] S. Grimme, M. Steinmetz *Phys. Chem. Chem. Phys.* **2013**, *15*, 16031–16042.
- [74] L. Goerigk, J. R. Reimers *J. Chem. Theory Comput.* **2013**, *9*, 3240–3251.
- [75] L. Goerigk, C. A. Collyer, J. R. Reimers *J. Phys. Chem. B* **2014**, *118*, 14612–14626.
- [76] D. Gruzman, A. Karton, J. M. L. Martin *J. Phys. Chem. A* **2009**, *113*, 11974–11983.
- [77] G. I. Csonka, A. D. French, G. P. Johnson, C. A. Stortz *J. Chem. Theor. Comput.* **2009**, *5*, 679–692.

- [78] J. J. Wilke, M. C. Lind, H. F. Schaefer III, A. G. Csaszar, W. D. Allen *J. Chem. Theor. Comput.* **2009**, *5*, 1511–1523.
- [79] D. Reha, H. Valdes, J. Vondrasek, P. Hobza, A. Abu-Riziq, B. Crews, M. S. de Vries *Chem. Eur. J.* **2005**, *11*, 6803–6817.
- [80] P. Jurecka, J. Sponer, J. Cerny, P. Hobza *Phys. Chem. Chem. Phys.* **2006**, *8*, 1985–1993.
- [81] T. Takatani, E. G. Hohenstein, M. Malagoli, M. S. Marshall, C. D. Sherrill *J. Chem. Phys.* **2010**, *132*, 144104.
- [82] S. Grimme, C. Mück-Lichtenfeld, E.-U. Würthwein, A. W. Ehlers, T. P. M. Goumans, K. Lammertsma *J. Phys. Chem. A* **2006**, *110*, 2583–2586.
- [83] M. Piacenza, S. Grimme *J. Comput. Chem.* **2004**, *25*, 83–99.
- [84] H. L. Woodcock, H. F. Schaefer III, P. R. Schreiner *J. Phys. Chem. A* **2002**, *106*, 11923–11931.
- [85] P. R. Schreiner, A. A. Fokin, R. A. Pascal, A. de Meijere, *Org. Lett.* **2006**, *8*, 3635–3638.
- [86] C. Lepetit, H. Chermette, M. Gicquel, J.-L. Heully, R. Chauvin *J. Phys. Chem. A* **2007**, *111*, 136–149.
- [87] J. S. Lee *J. Phys. Chem. A* **2005**, *109*, 11927–11932.
- [88] E. R. Johnson, P. Mori-Sanchez, A. J. Cohen, W. Yang *J. Chem. Phys.* **2008**, *129*, 204112.