A Comprehensive Assessment of the Effectiveness of Orbital Optimization in Double-Hybrid Density Functionals in the Treatment of Thermochemistry, Kinetics, and Noncovalent Interactions

Asim Najibi and Lars Goerigk*

School of Chemistry, The University of Melbourne, Victoria 3010, Australia

E-mail: lars.goerigk@unimelb.edu.au

Abstract

Orbital optimization (OO) has been suggested as a way to solve some shortcomings of second-order Møller-Plesset (MP2) variants and double-hybrid density functionals (DHDFs). A closer inspection of the literature, however, shows that the only two studies on OO-DHDFs were limited to three non-empirical PBE-based functionals, which are known to be of only mediocre accuracy. Herein, we provide a more in-depth analysis of OO-DHDFs with the main focus being on main-group thermochemistry, kinetics, and noncovalent interactions. We re-analyze two PBE-based OO-DHDFs and present four new OO-DHDF variants, two of which make use of the spin-component-scaling idea in their nonlocal correlation part. We also provide a more thorough analysis of three OO-MP2 variants. After assessing more than 621 reference points, we come to the conclusion that the benefits of OO are not as straightforward as previously thought. Results heavily depend on the underlying parent method. While OO-SCS/SOS-MP2 usually provide improved results — including for noncovalently bound systems — the opposite is true for OO-MP2. OO-DHDFs, like their non-optimized counterparts, still require London-dispersion corrections. Among the DHDFs, the largest effect of OO on thermochemical properties is seen for PBE0-2 and the smallest for PBE0-DH. However, results can both worsen and improve with OO. If the latter is the case, the resulting OO-DHDF is still outperformed by the currently most accurate conventional DHDFs, namely DSD-BLYP and DSD-PBEP86. We therefore recommend the OO technique only to be used in specialized cases. For the general method user we re-emphasize using conventional dispersion-corrected DHDFs for robust reliable results. Our findings also indicate that entirely different strategies seem to be required in order to obtain a substantial improvement over the currently best DHDFs.

Introduction

Double-hybrid density functionals (DHDFs) have become indispensable tools in modern thermochemistry. In the spirit of hybrid density functional theory (DFT),¹ nonlocal wavefunction theory (WFT) type terms substitute parts of semi-local DFT components, with the difference that in DHDFs this substitution occurs not only for the exchange (X), but also for the correlation (C) energy. In its simplest form, as introduced by Grimme in 2006,² the DHDF XC-energy expression (E_{XC}^{DHDF}) is:

$$E_{XC}^{DHDF} = (1 - a_X) E_X^{DFT} + a_X E_X^{HF} + (1 - a_C) E_C^{DFT} + a_C E_C^{MP2} \quad (1)$$

where $E_{X/C}^{DFT}$ are the semi-local DFT exchange and correlation energy density functional approximations (DFAs), respectively, E_X^{HF} is a nonlocal WFT-type Fock-exchange energy term, E_C^{MP2} a nonlocal second-order perturbative correlation-energy term, and a_X and a_C are scale parameters. The first three components in Eq. 1 resemble a hybrid DFA and they are used to obtain a set of Kohn-Sham molecular orbitals (KS-MOs) and MO energies within the familiar self-consistent-field (SCF) framework. The resulting occupied and virtual KS-MOs are then used to obtain the perturbative correlation-energy contribution. Thus, DHDFs are formally related with second-order Møller-Plesset perturbation theory (MP2),³ which means that known techniques used to accelerate or improve the results of MP2 treatments can also be applied to DHDFs.

Due to the inclusion of virtual KS-MOs in the correlation-energy treatment, DHDFs belong to the fifth and highest rung on Perdew’s “Jacob’s Ladder”,⁴ which is simultaneously attributed to higher accuracy and reliability compared to lower rungs. That this is indeed the case was first conclusively shown for main-group thermochemistry, kinetics, and noncovalent interactions in a study of 45 DFAs performed in 2011,⁵ and very recently re-confirmed in a much broader investigation of nearly 220 dispersion-corrected

and -uncorrected DFA variations.⁶ Not only are DHDFs on average the most accurate and robust DFT methods, they also outperform WFT-based MP2 variants.⁵ The demonstrated accuracy of the DHDF idea has sparked high activity amongst method developers, with most advancements being made for electronic ground-state thermochemistry,^{7–20} while only a few studies also pursued DHDF-based strategies for the treatment of electronic excited states.^{21–23} Herein, we will only provide a compact overview of recent developments in this field and we recommend Ref. 24 for a detailed review.

DHDFs that strictly follow Eq. 1 can either contain empirically fitted scale parameters, such as the B2PLYP² and B2GP-PLYP⁹ DFAs, or they can be determined through the adiabatic connection formalism and a series of approximations; examples for those so-called “non-empirical” DHDFs are PBE0-DH¹⁶ and PBE0-2.¹⁷ Such semi- and non-empirical DHDFs were further improved by employing the now well-established concepts of spin-component^{25,26} (SCS) and spin-opposite²⁷ (SOS) scaling in the evaluation of the MP2-type term.^{8,10–14,19,20} In fact, Kozuch and Martin’s spin-scaled “DSD” methods,^{12–14} such as DSD-BLYP¹² and DSD-PBEP86,¹³ are currently by far the most accurate and reliable DFT methods for ground-state thermochemistry, kinetics and noncovalent interactions.⁶ In recent years, other strategies to further improve the nonlocal WFT-type correlation term were investigated. However, success in these endeavours has varied. Replacing the spin-opposite-scaled MP2 portion in PWPB95¹¹ with a scaled random-phase-approximation (RPA) type term showed promising results and made the reliable treatment of low-gap systems and transition metal-compounds with DHDFs a possibility;²⁸ for loosely related RPA-based approaches, see Refs. 29–33. In 2016, Chan et al. explored the inclusion of third- and fourth-order perturbative terms and of a perturbative (T)-type³⁴ triples correction in DHDFs.³⁵ Their conclusion was that such higher-order terms did not provide any significant improvements that warrant the additional computational cost and that the use of the MP2-based correlation term alone was sufficient for DHDFs.

Conventional MP2 treatments are carried out as a post-SCF approach to obtain the perturbative correlation-energy contribution, which means that the initial Hartree-Fock (HF) orbitals and energies are not further changed during this energy calculation. A poor HF reference can consequently lead to poor results. In 2007, Lochan and Head-Gordon introduced the “orbital-optimized opposite-spin scaled second-order correlation” concept³⁶ (O2 method), which we will refer to as “orbital-optimized” (OO) SOS-MP2 later. In 2009, Neese and Grimme followed with a discussion of OO-MP2 and OO-SCS-MP2.³⁷ In OO-MP2 type approaches, the total MP2 energy of a system is minimized with respect to MP2 amplitudes and orbital-relaxation parameters, see Ref. 37 for details. The O2 method was tested on small closed-shell systems and doublet radicals for which atomization energies, structures and vibrational frequencies were calculated. It was concluded that it gave results similar to SOS-MP2 for the closed-shell systems, but showed improvements compared to SOS-MP2 and conventional MP2 for the properties of the doublet radicals, including an almost complete elimination of spin contamination in those systems.³⁶ The OO-(SCS-)MP2 study by Neese and Grimme focussed on thermochemistry and kinetics of a small selection of organic closed-shell molecules, organometallic transition-metal compounds, and doublet radicals. For the latter two molecular classes, the authors reported smaller errors for the OO methods compared to their unoptimized counterparts. Other studies using OO-MP2 and related variants can be found in Refs. 38–40; they include, for instance, further analyses on noncovalent interaction (NCI) energies, aromatic bond-dissociation energies, and barrier heights (BHs) in open-shell systems. Closely related works on variations of the OO idea or its application to other WFT approaches can be found in Refs. 41–43.

Due to their formal resemblance with wave-function MP2, DHDFs can also be applied within the OO framework (OO-DHDFs). However, work in this area has been limited to two studies. In 2013, Head-Gordon and Peverati conducted initial studies, testing OO versions of the “non-empirical” PBE0-DH and PBE0-2 approaches.⁴⁴ A clear advantage of

OO-DHDFs over conventional DHDFs was shown for calculations of dipole-moment and nuclear-force dissociation curves of diatomics. PBE0-DH and PBE0-2 displayed unphysical jumps in these curves, whereas their OO versions showed smooth dissociation behaviors close to coupled-cluster singles doubles (CCSD) reference curves. Similarly to the related OO-MP2 variants, OO-DHDFs suffered less from spin contamination in the dissociation of the open-shell diatomic O_2^+ . Improved results were also reported for bond lengths in 12 diatomic doublet radicals compared to their parent DHDFs and (OO-)MP2. The OO-DHDFs were also tested on a small database that contained 262 data points ranging from total atomization energies (TAEs), to ionization potentials (IPs), electron affinities (EAs), proton affinities (PAs), bond dissociation energies, and NCI energies. However, across the entire database, the OO strategy seemed not to be a guarantee for improved results, with OO-PBE0-DH being similar to PBE0-DH, while OO-PBE0-2 showed a slight improvement of about 0.2 kcal/mol for its final mean absolute deviation (MAD). In 2016, the OO scheme was tested on the “parameter-free” PBE-QIDH¹⁸ approach on a thermochemistry and kinetics set comprising 46 reactions, the O_2^+ dissociation curve, NCI energies in six dimers of water with small open-shell systems, as well as singlet-triplet splittings in oligoacenes.⁴⁵ OO-PBE-QIDH was compared with its underlying parent PBE-QIDH, the conventional B2PLYP DHDF, and the PBE0 hybrid DFA. Marginal improvements compared to PBE-QIDH were seen for systems that were prone to the self-interaction error (SIE), as well as for BHs and singlet-triplet splittings. However, after close inspection of those results, we noticed that already the non-orbital optimized B2PLYP method gave better singlet-triplet splittings than PBE-QIDH. Therefore, it is not entirely clear whether the OO scheme would only be beneficial to PBE-QIDH or also to other DHDFs.

The two studies in 2013 and 2016 have both shifted the focus of DHDF development on the OO idea. They also showed that OO can be beneficial for electronically difficult systems. However, there is a need to further investigate the usefulness of this technique

for applications to general thermochemistry, kinetics, and NCIs. Both studies were limited to entirely PBE-based DHDFs. Moreover, it was pointed out in 2014 that so-called “non-empirical” DHDFs, such as PBE0-DH and PBE0-2, were less robust and accurate than semi-empirical versions.²⁴ Finally, while NCIs were assessed for OO-DHDFs, the question whether OO would obviate the need to apply London-dispersion corrections has never been addressed.

Herein, we present a comprehensive investigation of the OO strategy for a variety of DHDFs with our main goal being to assess the robustness of OO-DHDFs for main-group thermochemistry, kinetics, and NCIs. To answer the question if OO is an effective strategy in these areas, we argue that one should start the investigation with the most accurate and robust DHDFs available. Only if their OO versions outperform their unoptimized parent DFAs, one can suggest OO as a valuable enhancement of current DHDFs. For this purpose, we will assess a database of 621 reference energies split into two categories covering closed- and open-shell systems. We will assess a variety of semi-empirical and non-empirical DHDFs, namely DSD-PBEP86, DSD-BLYP, B2PLYP, B2GP-PLYP, PBE0-2, and PBE0-DH. In this context, we will also address the aforementioned topic of London-dispersion corrections; note that such corrections are necessary for conventional DHDFs — even for thermochemistry — despite their nonlocal correlation component.^{5,6,11,24,46} We also use this occasion to re-investigate the related OO versions of MP2, SCS-MP2, and SOS-MP2, and provide a more comprehensive assessment of these. Orbital optimization for the DHDFs and MP2 methods increases the computing time by a factor of about 10 compared to the underlying methods,³⁷ which is why it is necessary to also assess its accuracy when smaller atomic-orbital (AO) basis sets are used, something that has not been investigated before. Therefore, the dependence of the OO-DHDFs and OO-MP2 methods and their underlying DHDFs and MP2 methods on the basis-set size is also investigated for the various test cases sets in this study. Overall, we hope our analysis will provide a more comprehensive

understanding of the OO idea and new insights for potential future method developments and applications.

Computational Details

Assessed Chemical Systems

This study focusses on twenty benchmark sets, which combined provide a representative picture of the usefulness and robustness of the OO-DHDF idea. Detailed information on each set is given in Tab. 1. All are established sets in the literature, with most of them also being part of the recently published GMTKN55⁶ database for general main-group thermochemistry, kinetics, and NCIs. Only two sets were taken from the older GMTKN24⁴⁷ and GMTKN30¹¹ databases (RG6 and PA), while the O23 set was taken from a study investigating OO in the WFT context. The chemical processes represented by eight sets involve open-shell species (O23, BH76, G21IP, G21EA, MB16-43, RSE43, W4-11, and SIE4x4), while the remaining twelve contain exclusively closed-shell systems. All molecular geometries and reference values were adopted from the literature without any adjustments. All reference values are nonrelativistic and zero-point-vibrational-energy exclusive. As noted in footnotes of Tab. 1 and the Supporting Information (SI), some systems had to be discarded from this study due to convergence problems (also see the next section). We decided to keep the originally suggested names for the affected benchmark sets for consistency reasons.

In addition, we investigated the dissociation of the argon dimer, with W2-F12⁴⁸ reference values taken from Ref. 49, as well as adiabatic singlet-triplet energy gaps of the linear oligoacene series from benzene to hexacene, taken from the OO-PBE-QIDH study in 2016.⁴⁵ For better comparison, we adopt the same experimental reference values from Ref. 50.

Table 1: Overview of the assessed benchmark sets.

set	description	# ^a
RG6 ¹¹	NCI energies of rare-gas dimers	6
S22 ^{51,52}	NCI energies in noncovalently bound dimers	22
SCONF ^{6,47,53}	relative energies of sugar conformers	17
WATER27 ^{54,55}	NCI energies in neutral and charged water clusters	27
PCONF21 ^{6,56,57}	relative energies of tri- and tetrapeptide conformers	18
O23 ³⁹	NCI energies in open-shell systems	15 ^b
BH76 ^{6,47,58,59}	barrier heights in reactions of small molecules	76 ^c
BHPERI ^{47,60}	barrier heights in pericyclic reactions	26
PX13 ^{6,61}	proton-exchange barriers	13
BHROT27 ⁶	barrier heights for rotation around single bonds	27
INV24 ⁶²	inversion barrier heights	24
RSE43 ^{6,37}	radical-stabilization energies	43
W4-11 ⁶³	total atomization energies	140 ^d
MB16-43 ⁶	decomposition energies of artificial molecules	43
SIE4x4 ⁶	SIE-related problems	16
G21EA ^{47,64}	adiabatic electron affinities	25 ^e
G21IP ^{47,64}	adiabatic ionization potentials	36 ^f
DIPCS10 ⁶	adiabatic double-ionization potentials	10
G2RC ^{6,47}	reaction energies for small systems	25
PA ^{47,65,66}	adiabatic proton affinities	12

^a Number of relative energies. ^b 15 instead of 23 systems due to convergence problems for 8 systems for all OO methods (see Tabs. S7 and S45). ^c 74 systems for approaches based on PBE0-2, PBE0-DH, DSD-PBEP86, and MP2 due to convergence problems for their OO versions (see Tabs. S8 and S46). ^d The OO versions of some methods showed convergence issues and required adjustment of the statistics (see Tabs. S14 and S52): PBE0-2 (134 systems included in the statistics), PBE0-DH (133 systems), DSD-PBEP86 (133 for def2-QZVP, 126 for def2-TZVP), SCS-MP2 (139 systems). ^e The OO versions of some methods showed convergence issues and required adjustment of the statistics (see Tabs. S17 and S55): PBE0-DH (23 systems included in the statistics), all MP2 variants (24 for def2-QZVP, 23 for def2-TZVP). ^f The OO versions of some methods showed convergence issues and required adjustment of the statistics (see Tabs. S18 and S56): PBE0-2 (28 systems included in the statistics), PBE0-DH (29 systems for def2-QZVP, 28 for def2-TZVP), DSD-PBEP86 (27 for def2-QZVP, 26 for def2-TZVP).

Technical Details

The general-purpose quantum chemistry program ORCA (Versions 3.0.3 and 4.0.1)^{67,68} was used to carry out all calculations. To the best of our knowledge, this is the only program which allows OO for DHDFs. The underlying DHDFs used in this study are B2PLYP,² B2GP-PLYP,⁹ PBE0-2,¹⁷ PBE0-DH,¹⁶ DSD-PBEP86,¹³ and DSD-BLYP.¹²

B2PLYP, B2GP-PLYP, and DSD-BLYP are semi-empirical DHDFs containing Becke88⁶⁹ (B88) exchange and Lee-Yang-Parr^{70,71} (LYP) correlation. PBE0-DH and PBE0-2 are so-called “parameter-free” or “non-empirical” DHDFs with Perdew-Burke-Ernzerhof⁷² (PBE) exchange and correlation. DSD-PBEP86 is a semi-empirical DHDF with PBE exchange and Perdew-Wang86^{73,74} (P86) correlation. The DHDFs’ scaling parameters for Fock-exchange and the MP2 components are given in Tab. 2. In addition to the DHDFs, the MP2, SCS-MP2,²⁵ and SOS-MP2²⁷ methods are used to compare the OO technique between WFT and DHDFs.

The analysis of the performance of the various levels of theory for the thermochemistry, kinetics and NCI benchmark sets was carried out with the Ahlrichs-type split-valence triple- (TZ) and quadruple- ζ (QZ) atomic-orbital (AO) basis-sets def2-TZVP and def2-QZVP. In line with previous works on the GMTKN databases, diffuse s- and p- functions taken from the related augmented Dunning basis sets⁷⁵ were combined with the “def2” ones for the assessment of the G21EA and WATER27 test sets; the resulting AO basis sets are dubbed aug’-def2-T/QZVP.^{5,6,11,47} The large aug-cc-pVQZ⁷⁵ QZ basis-set with polarization and

Table 2: Scale factors of Fock exchange (a_X) and of second-order perturbative correlation (a_C) for the DHDFs used in this study. For the DSD functionals, the scale factors of the perturbative correlation contribution of opposite-spin electron pairs (c_O) and that of same-spin electron pairs (c_S) are shown, together with the scale factors of the DFT correlation (c_C); see Refs. 12–14.

functional	a_X^a	a_C^b	c_C	c_O	c_S
B2PLYP	0.53	0.27	—	—	—
B2GP-PLYP	0.65	0.36	—	—	—
PBE0-DH	0.50	0.125	—	—	—
PBE0-2	0.7937	0.50	—	—	—
DSD-BLYP	0.69	—	0.54	0.46	0.37
DSD-PBEP86	0.70	—	0.43	0.53	0.25

^a The scale factor for the DFT exchange component is $1-a_X$.

^b With the exception of the DSD methods, the DFT and MP2 correlation scale factors add to 1.

diffuse functions is used to compute the argon-dimer dissociation curves, since this AO basis set is close to the CBS limit and since very small relative energies are calculated in this case. Dunning’s cc-pVTZ⁷⁶ basis set was used to calculate singlet-triplet energy gaps of linear acenes in order to make a direct comparison with a previous OO-DHDF study which utilized this basis set. The resolution-of-the-identity (RI) approximation for Coulomb integrals in combination with the chain-of-sphere approximation for exchange integrals (RIJCOSX⁷⁷) was applied with ORCA’s default settings and the appropriate auxiliary basis functions.⁷⁸ The second-order perturbative treatment in all methods was also carried out with the RI approximation and the appropriate auxiliary basis functions.⁷⁹ ORCA’s routine ‘GRID4’ option was applied for the numerical integration of the exchange-correlation potentials in the (OO-)DHDFs. The frozen-core approximation was used for all (OO-)MP2 variants. The SCF convergence criterion was set to $10^{-7} E_h$ in all calculations.

Occasionally, OO calculations did not successfully converge. Some convergence problems were seen for the same systems across various OO methods and for both AO basis sets, but sometimes we also observed problems only for one method or one particular basis set. Unfortunately, only a very limited number of options to force convergence have been implemented into the OO part of ORCA. As a consequence, unconverged systems had to be left out of the analysis, as indicated in the previous section.

The DFT-D3(BJ)^{80,81} additive London-dispersion correction is applied to all DHDFs and their orbital-optimized counterparts. Damping parameters for B2PLYP, B2GP-PLYP and DSD-BLYP were taken from Ref. 5, for DSD-PBEP86 from Ref. 13, and for PBE0-2 and PBE0-DH from Ref. 24. Herein, we present new damping parameters for all OO-DHDFs that we determined in a least-squares fit as outlined in Section SI.1 in the SI. Therein, the values of all damping parameters used for this work are presented. In the following we will drop the acronym “(BJ)” and only use the suffix “-D3”.

In the following Results and Discussion section, the deviation of any relative energy

from the reference value is calculated as the difference between the evaluated method and the reference. This implies that a positive value of the deviation stands for an overestimation and a negative for an underestimation of the property in question. Based on these deviations, we will also discuss mean deviations (MDs) and mean absolute deviations (MADs). In many cases, a comparison between an OO method and its unoptimized version will be conducted by calculating the ratio between their respective MADs. A ratio of less than unity implies a systematic improvement of the OO approach, a value close to unity implies a negligible influence of OO, a value larger than unity signifies that OO leads to on average worse results. Note that we carry out our analysis using MADs, following historical tradition. Additionally, in order to verify the reliance of these values, we thoroughly checked root-mean-square deviations (RMSDs), but we did not find any changes to our conclusions. All RMSDs are included in the SI. For a discussion on the advantages and disadvantages of using MADs and RMSDs, see e.g. Ref. 82. For a recent example of analyses based on MADs giving the same trends as those based on RMSDs, see Ref. 6.

Results and Discussion

Singlet-Triplet Energy Gaps in Linear Acenes

In 2016, Sancho-García et al. tested the effect of OO on the PBE-QIDH functional for adiabatic singlet-triplet energy gaps in linear acenes from benzene to hexacene.⁴⁵ The authors showed that OO improved PBE-QIDH in these cases, as reflected by the MAD being significantly reduced from $\text{MAD} = 4.59 \text{ kcal/mol}$ (PBE-QIDH) to $\text{MAD} = 2.58 \text{ kcal/mol}$ (OO-PBE-QIDH).⁴⁵ We start our investigation of OO-DHDFs by extending the previously conducted work on these systems. Tab. 3 shows the adiabatic singlet-triplet energy gaps for various DHDFs and MP2 methods based on the same geometries and AO basis set as in 2016. For the DHDFs, OO improves the MADs of the DHDFs by at least 1.75 kcal/mol

Table 3: Adiabatic singlet-triplet energy gaps (kcal/mol) for linear acenes calculated with the cc-pVTZ AO basis set, and mean absolute deviations (MADs) from reference values (kcal/mol) for each assessed method.

method	benzene	naphthalene	anthracene	tetracene	pentacene	hexacene	MAD
reference ^a	84.34	60.90	42.60	29.45	19.83	12.43	—
PBE-QIDH ^b	93.95	67.89	45.73	32.18	22.12	15.21	4.59
OO-PBE-QIDH ^b	92.32	66.16	44.60	30.07	19.69	12.21	2.58
B2PLYP	94.57	66.76	45.48	31.86	22.13	15.66	4.48
OO-B2PLYP	92.98	65.23	44.23	30.07	20.15	13.29	2.73
B2GP-PLYP	96.21	68.67	46.88	33.57	23.79	17.28	6.14
OO-B2GP-PLYP	93.82	66.28	45.12	30.84	20.80	13.66	3.50
PBE0-2	98.74	71.59	48.60	36.10	26.43	20.12	8.67
OO-PBE0-2	94.86	67.45	45.95	31.52	21.34	13.87	4.24
PBE0-DH	89.69	64.38	42.86	28.41	18.13	10.67	2.26
OO-PBE0-DH	89.32	64.08	42.62	28.04	17.71	10.16	2.33
DSD-PBEP86	96.74	69.19	47.51	35.21	25.77	19.92	7.46
OO-DSD-PBEP86	92.98	65.65	45.05	31.16	21.33	14.52	3.52
DSD-BLYP	97.85	69.86	47.99	35.24	25.67	19.57	7.77
OO-DSD-BLYP	94.47	66.52	45.58	31.44	21.51	14.52	4.08
MP2	113.00	83.71	56.85	49.43	42.66	39.81	22.65
OO-MP2	101.46	70.01	48.72	34.25	24.25	17.00	7.69
SCS-MP2	109.61	80.64	55.90	48.71	41.81	38.72	20.97
OO-SCS-MP2	98.46	68.99	49.22	35.60	25.99	19.13	7.97
SOS-MP2	107.91	79.11	55.43	48.35	41.39	38.18	20.14
OO-SOS-MP2	96.93	68.39	49.42	36.24	26.78	20.06	8.04

^a experimental values taken from Ref. 50. ^b taken from Ref. 45.

(OO-B2PLYP), with the largest improvement being 4.43 kcal/mol (OO-PBE0-2). The only exception is PBE0-DH for which OO slightly worsens the average error ($\text{MAD}(\text{PBE0-DH}) = 2.26$ kcal/mol and $\text{MAD}(\text{OO-PBE0-DH}) = 2.33$ kcal/mol). Despite this exception, we can confirm that OO does indeed improve the description of singlet-triplet splittings in these systems. The improvement is most strongly pronounced in the MP2-based methods, where MADs for the unoptimized approaches are at least 20 kcal/mol, while a reduction to about 8 kcal/mol is seen when the OO scheme is applied. Note, however, that even the OO-MP2 variants are outperformed by all conventional DHDFs. It is also interesting to

note that the unoptimized B2PLYP approach seems to be slightly better than conventional PBE-QIDH ($\text{MAD}(\text{B2PLYP}) = 4.48 \text{ kcal/mol}$ and $\text{MAD}(\text{PBE-QIDH}) = 4.59 \text{ kcal/mol}$), while OO reverses this order, with $\text{MAD}(\text{OO-B2PLYP}) = 2.73 \text{ kcal/mol}$ and $\text{MAD}(\text{OO-PBE-QIDH}) = 2.58 \text{ kcal/mol}$. Interestingly, the two DSD methods are largely improved by OO, with MADs being reduced by 3.69 (OO-DSD-BLYP) and 3.94 kcal/mol (OO-DSD-PBEP86), however their OO versions are not competitive with those of the other DHDFs with the exception of OO-PBE0-2 (see Tab. 3).

It is important to recognise that although OO improves most of the DHDFs, the best method for the adiabatic singlet-triplet energy gaps of the linear acenes is in fact conventional PBE0-DH ($\text{MAD} = 2.26 \text{ kcal/mol}$). It is followed by OO-PBE0-DH ($\text{MAD} = 2.33 \text{ kcal/mol}$), OO-PBE-QIDH ($\text{MAD} = 2.58 \text{ kcal/mol}$), and OO-B2PLYP ($\text{MAD} = 2.73 \text{ kcal/mol}$). Out of the tested six OO-DHDFs, OO-PBE0-2 is the least accurate with an MAD of 4.24 kcal/mol.

Orbital Optimization and London-Dispersion Effects

Due to their scaled-down MP2 contribution, DHDFs also need dispersion corrections to properly describe van-der-Waals (vdW) effects.⁴⁶ In this context, we would like to point out that London dispersion does not only play a role in the stability of noncovalently bound systems or in conformers, but also in the determination of reaction energies and BHs.^{5,6} We note, however, that the two previous OO-DHDF studies did not include any appropriate dispersion corrections,^{44,45} which is why an analysis of the effect of OO on the magnitude of nonlocal electron-correlation effects, such as London dispersion, is warranted. Analysis of the argon-dimer dissociation allows us to have a closer look at such effects. Before discussing the OO-DHDFs, however, we first have a look at the MP2 approaches. It is well-known that conventional MP2 overestimates London-dispersion effects, while most of its spin-scaled versions underestimate them,^{26,51,83} apart from some variants that have

been specialized for the treatment of London dispersion effects.^{26,84-86} The same behavior can also be seen in Figs. 1a-c. OO causes further stabilization of the system. MP2 overestimates the interaction energy in the vdW-minimum region by 0.05 kcal/mol, with an underestimation of the Ar-Ar distance in the same region by 0.03 Å. OO overestimates this energy by an additional 0.10 kcal/mol and shortens the Ar-Ar distance by another 0.05 Å (Fig. 1a) . For the SCS- and SOS-MP2 methods, OO also lowers the interaction energy and the corresponding interatomic distance, so that they are closer to the W2-F12 reference. In particular, the OO-SCS-MP2 curve shows reasonable agreement with the reference (Fig. 1b), while OO-SOS-MP2 still underestimates the interaction energy and predicts the minimum structure to be located at a larger Ar-Ar distance (Fig. 1c). Thus, OO does have an effect on the dispersion forces in MP2-type treatments.

We discuss the performance of OO-DHDFs with B2GP-PLYP as a representative (Fig. 1d) As expected, the pure DFA underestimates dispersion interactions, something that can be fixed by adding Grimme’s DFT-D3(BJ) correction. In fact, B2GP-PLYP-D3 shows perfect agreement with the reference curve. OO, on the other hand, has no effect and the OO-B2GP-PLYP curve is nearly identical with the dispersion-uncorrected B2GP-PLYP one. The dissociation curves for all tested OO-DHDFs are shown in Fig. S1 in the SI, and they show the same behavior as discussed for OO-B2GP-PLYP. We can therefore conclude that also OO-DHDFs need dispersion corrections, which is why we determined DFT-D3(BJ) damping parameters for them (see SI). In the following discussion, we will always use dispersion corrections when necessary. However, sometimes we will also discuss uncorrected results when we want to separate the influence of OO from that of DFT-D3.

Figure 1: The interaction energy (kcal/mol) of the argon dimer as a function of interatomic distance ($\text{\AA}$), calculated using a) MP2 and OO-MP2, b) SCS-MP2 and OO-SCS-MP2, c) SOS-MP2 and OO-SOS-MP2 and d) DHDFs and OO-DHDFs (here B2GP-PLYP is used as an example). The reference curve (W2-F12) is shown in each case.

Analysis of Noncovalent Interaction Test Sets

Non-Covalent Interactions in Closed-Shell Systems

First, we discuss the results for the NCI test sets that only include closed-shell systems, namely RG6, S22, SCONF, WATER27, and PCONF21. For the argon-dimer dissociation curve, we already saw that OO brought the SCS- and SOS-MP2 curves closer to the reference, but that it overestimated the interaction energies for the MP2 curve. We also saw that the DHDF and OO-DHDF dissociation curves were virtually identical. It is not

surprising that the same trends can be observed for the RG6 set of six rare-gas dimers in their equilibrium geometries, which we will therefore only discuss very briefly; detailed results are presented in Tab. S2 in the SI. OO dramatically improves the SCS-MP2 and SOS-MP2 methods from MADs of 0.30 and 0.38 kcal/mol to 0.08 and 0.18 kcal/mol, respectively, while it mildly worsens the MP2 results from 0.13 to 0.15 kcal/mol. In DHDFs, OO has practically no effect on the MADs and both DHDFs and OO-DHDFs need dispersion corrections to obtain acceptable MADs. In almost all cases, both OO-DHDF-D3 and DHDF-D3 MADs are within or close to the accuracy limit of the reference values (about 0.1 kcal/mol), rendering OO as nonessential for this set.

Fig. 2 shows the MADs of the various methods for the S22, SCONF, WATER27 and PCONF21 sets. Based on our previous discussion, it is expected that the dispersion correction decreases the MADs, which is indeed the case in the majority of cases. Exceptions are in particular the two PBE-based DHDFs for the SCONF and WATER27 test sets; the observed overestimation of the NCIs for these functionals has already been reported elsewhere.²⁴

OO only slightly improves the DHDF results by on average 0.1 kcal/mol or less, and it has thus a significantly smaller effect than the dispersion correction (see Fig. 2). However, exceptions are observed for some functionals and we observe that inclusion of OO may lead to different outcomes depending on the test set. For instance, while OO shows improvements for DSD-PBEP86 for S22 and PCONF21, MADs for SCONF and WATER27 slightly increase. In particular, we notice that PBE0-2 is affected the most by OO and that the MADs increase by 64 % and 35 % for SCONF and WATER27, respectively. The related OO-PBE0-DH approach shows only an increase of 13 % for SCONF and a negligible improvement of 0.03 kcal/mol (<1%) for WATER27. PBE0-2 has with 50 % the largest fraction of MP2-type correlation of all tested DHDFs and its strikingly different results compared to the very similar PBE0-DH indicate that the stronger nonlocal corre-

Figure 2: Mean absolute deviations (MADs) in kcal/mol from reference values, for the a) calculated interaction energies in the S22 set (medium-sized dimers), b) relative conformer energies in the SCONF set (sugars), c) interaction energies in the WATER27 set (water clusters), and d) relative conformer energies in the PCONF21 set (peptides).

lation contribution determines how application of OO influences the outcome. Indeed, we observe a strong effect of OO for the three WFT methods, with results deteriorating significantly for OO-MP2. For instance, the MADs for OO-MP2 are three times larger than for MP2 for SCONF and four times larger for WATER27. Again, the two spin-component-scaled approaches seem to be more compatible with the OO strategy; with the exception of PCONF21, OO-SCS-MP2 shows a large reduction in the errors compared to SCS-MP2. SOS-MP2 is improved by the OO technique for all four benchmark sets (see Fig. 2).

Note that nearly all trends seen for S22 in Fig. 2a, can also be observed for the three sub-

Figure 3: Mean absolute deviations (MADs) in kcal/mol from reference values, for the calculated open-shell non-covalent interaction energies in the O23 set.

categories of S22, namely hydrogen-bonded, dispersion-bonded, and “mixed” complexes, with the exception of the dispersion-bonded dimers, for which OO-SCS-MP2 seems to be worse than SCS-MP2 (see Fig. S2). Overall, however, the results for NCIs in closed-shell systems observed for all methods reflect the same trends discussed for the rare-gas dimer cases and we can conclude that the effect of OO on NCIs strongly depends on the type of chosen approach, with DHDFs being the least affected. However, there may be an influence based on the amount of mixed-in nonlocal correlation.

Non-Covalent Interactions in Open-Shell Systems

The statistics for the interaction energies in open-shell dimers (O23 set) are shown in Fig. 3. We see a dramatic improvement for all three MP2 methods when OO is used: $\text{MAD}(\text{MP2}) = 0.90$ kcal/mol vs. $\text{MAD}(\text{OO-MP2}) = 0.39$, $\text{MAD}(\text{SCS-MP2}) = 1.27$ kcal/mol vs. $\text{MAD}(\text{OO-SCS-MP2}) = 0.39$, and $\text{MAD}(\text{SOS-MP2}) = 1.49$ kcal/mol vs. $\text{MAD}(\text{OO-SOS-MP2}) = 0.71$ kcal/mol (all statistical values are shown in Tab. S7). This observation agrees with previous studies by Bozkaya et al. where OO improved a variety of MP2 methods for the same test set.^{39,40} Here, the only DHDF that improves with OO is PBE0-2 (Fig. 3) with $\text{MAD}(\text{OO-PBE0-2}) = 0.43$ kcal/mol as opposed to $\text{MAD}(\text{PBE0-2}) = 0.56$ kcal/mol.

This again is an indication that the effectiveness of the OO scheme cannot be generalized.

Reaction Barrier Heights

The BH76, BHPERI and PX13 benchmark sets include BHs for processes that involve bond making and breaking, with BH76 being the only set containing open-shell species. The BHROT27 and INV24 benchmark sets consist of BHs that do not involve bond making and breaking, but instead rotations around single bonds or inversions of molecular shapes, respectively. Herein, we investigate if OO has a different effect on the nature of the process leading to a transition state and its related BH.

The results shown in Fig. 4 indicate that OO indeed changes the DHDF results for BH76. In most of the cases, however, MADs increase by between 0.3 and nearly 0.6 kcal/mol. The smallest increase of 0.09 kcal/mol is observed for PBE0-DH, the functional with the smallest amount of MP2 correlation. PBE0-2, the functional with the largest non-local correlation component, is the only DHDF that benefits from OO, with an MAD that decreased by 0.6 kcal/mol. However, its final MAD is still higher than that of conventional DSD-PBEP86 (Fig. 4). Results for the BHPERI set — despite involving bond-breaking processes — as well as results for BHROT27 and INV24 remain unaffected by OO, while the MADs for PX13 show an increase for all six tested methods, with the largest increase of 0.44 kcal/mol seen for OO-PBE0-2.

OO offers a way to improve the three MP2 variants for the BH76 benchmark set: $\text{MAD}(\text{MP2}) = 4.42$ kcal/mol and $\text{MAD}(\text{OO-MP2}) = 2.85$ kcal/mol, $\text{MAD}(\text{SCS-MP2}) = 5.47$ kcal/mol and $\text{MAD}(\text{OO-SCS-MP2}) = 1.89$ kcal/mol, and $\text{MAD}(\text{SOS-MP2}) = 6.17$ kcal/mol and $\text{MAD}(\text{OO-SOS-MP2}) = 2.35$ kcal/mol (Tab. S8). However, for the remaining four test sets we see deteriorating MADs even for processes that do not involve bond-breaking. The only exception is SCS-MP2, which improves significantly for PX13 when orbital optimized, with a reduction of the MAD by almost two thirds. Inversion barriers

Figure 4: Mean absolute deviations (MADs) in kcal/mol from reference values, for the calculated barrier heights of the a) BH76, b) BHPERI, c) PX13, d) BHROT27, and d) INV24 sets.

remain unaffected by OO for SCS-MP2 and SOS-MP2.

We can conclude that the effect of OO for the DHDFs and MP2 methods is not necessarily dependent on whether the BH under investigation is for a process that involves bond making/cleaving, or not. While there is a stronger effect on open-shell species, this finding cannot yet be generalized, as seen for the proton-exchange barriers in PX13, which are affected by OO despite involving closed-shell species only.

Thermochemistry

Thermochemistry for Reactions Involving Open-Shell Species

We begin our discussion of thermochemistry with the RSE43 set for radical stabilization energies, as it was introduced in the first study of the OO-MP2 and OO-SCS-MP2, for which improvements were reported thanks to OO.³⁷ We can confirm this finding for the recently updated reference values of RSE43,⁶ with improvements of at least 3.5 kcal/mol for MP2-type methods (Tab. 4). The final MADs of OO-MP2 and OO-SCS-MP2 are about 0.5 kcal/mol, which is similar to most dispersion-uncorrected DHDFs. The effect of OO on the latter is again strongly dependent on the underlying functional. Here, it only plays a significant role for PBE0-2 [MAD(PBE0-2)=0.98 kcal/mol and MAD(OO-PBE0-2)=0.25 kcal/mol] and DSD-PBEP86 [MAD(DSD-PBEP86)=0.80 kcal/mol and MAD(OO-DSD-PBEP86)=0.19 kcal/mol], for which the errors of the underlying DHDFs considerably decrease. PBE0-DH remains completely unaffected, while DSD-BLYP shows a marginal improvement of only 0.1 kcal/mol. B2PLYP and B2GP-PLYP results worsen by 0.1 to 0.3 kcal/mol.

Next, we turn our attention to the TAEs of the W4-11 test set. It is well-known that MP2-type methods perform badly in the calculation of such energies.⁶³ OO is no remedy for this problem and all three MP2 methods show significantly worse results with OO, with

MADs being doubled and final MADs increasing up to values as high as 14.19 kcal/mol (OO-MP2); see Tab. 4. This trend is the opposite to the expectation that OO-MP2 methods are better than MP2 for open-shell systems, as discussed for RSE43. DHDFs do not inherit the problems of MP2 for TAEs and, in fact, they are known to be the best DFT-based methods for this reaction type.^{6,11} Indeed, we also observe different trends for OO-DHDFs compared to MP2-type methods. In fact, there is noticeable improvement of some of the DHDFs: OO-PBE0-2, OO-DSD-BLYP, OO-B2GP-PLYP, and OO-DSD-PBEP86, with the latter two showing improvements of about 0.8 kcal/mol or more. In fact, OO-DSDPBEP86 seems to be one of the best DFT-based approaches for the calculation of TAEs when compared with the large number of data points collected for W4-11 in Ref. 6 (MAD = 1.66 kcal/mol).

The “mindless benchmark set” MB16-43 set consists of decomposition energies of randomly created molecules — involving both closed- and open-shell systems — and it serves as a reliable indicator for a method’s robustness. Given that the average absolute decomposition energy of its reference values is relatively high (414.73 kcal/mol), MADs of the assessed methods also usually appear larger; see Ref. 6 for examples. When analyzing the effect of OO only, we do see that it can improve MADs by about 1 kcal/mol for most DHDFs, which is relatively small given the average absolute reaction energies and relatively large MADs (Tab. 4). The two exceptions are again the PBE-based DHDFs. While PBE0-DH remains nearly unaffected by OO, we observe an increase in the MAD of more than 1 kcal/mol for OO-PBE0-2. The effect of increasing MADs is even more pronounced for OO-MP2 for which the MAD increases from 22.45 to 37.15 kcal/mol. Contrary to that, the remaining two MP2 variants improve: MAD(SCS-MP2) = 26.33 kcal/mol vs. MAD(OO-SCS-MP2) = 7.26 kcal/mol, and SOS-MP2 (MAD(SOS-MP2) = 44.17 kcal/mol vs. MAD(OO-SOS-MP2) = 24.33 kcal/mol. The difference between conventional MP2 and its spin-component-scaled versions can be better understood by analyzing their MDs. OO

Table 4: Mean absolute deviations (MADs) in kcal/mol for various DHDFs and MP2 methods for the thermochemistry benchmark sets.^a

method	RSE43	W4-11	MB16-43	SIE4x4	G21EA	G21IP	DIPCS10	G2RC	PA
B2PLYP	0.71 (0.99)	2.15 (2.22)	29.82 (28.82)	9.90 (10.12)	1.30 (1.36)	2.25 (2.31)	3.80 (3.89)	1.40 (1.67)	1.51 (1.58)
B2GP-PLYP	0.45 (0.56)	2.60 (1.86)	20.17 (18.93)	6.58 (6.94)	1.71 (1.48)	1.98 (1.97)	3.20 (3.33)	1.41 (1.47)	1.44 (1.52)
PBE0-2	0.98 (0.25)	3.60 (3.29)	17.37 (18.54)	1.80 (2.38)	2.21 (1.80)	2.15 (1.90)	2.87 (2.80)	4.81 (4.46)	1.83 (1.95)
PBE0-DH	0.60 (0.60)	5.52 (5.32)	13.09 (12.99)	7.41 (7.45)	3.13 (3.09)	3.11 (3.12)	2.61 (2.62)	6.17 (6.14)	3.00 (3.00)
DSD-PBEP86	0.80 (0.19)	2.42 (1.66)	10.49 (9.52)	4.94 (5.48)	1.52 (1.20)	2.22 (2.25)	3.66 (3.81)	1.85 (1.90)	1.43 (1.51)
DSD-BLYP	0.56 (0.45)	2.33 (1.95)	18.98 (17.58)	5.77 (6.27)	1.77 (1.57)	1.98 (2.01)	3.79 (3.96)	1.12 (1.40)	1.33 (1.45)
MP2	3.53 (0.48)	7.53 (14.19)	22.45 (37.15)	1.54 (3.17)	3.13 (2.63)	3.37 (3.10)	6.01 (4.91)	3.06 (4.01)	0.80 (0.69)
SCS-MP2	4.01 (0.53)	3.96 (8.63)	26.33 (7.26)	2.16 (1.51)	3.81 (3.20)	3.57 (3.08)	6.55 (6.45)	2.25 (2.81)	1.41 (1.83)
SOS-MP2	4.24 (0.87)	4.29 (6.78)	44.17 (24.33)	2.58 (1.04)	4.55 (3.72)	3.97 (3.33)	7.01 (7.58)	2.62 (3.22)	2.39 (2.78)

^a Numbers shown in brackets are for the OO methods

causes an increase of the decomposition energies for all assessed methods. MP2 already overestimates those with MD = 9.77 kcal/mol, which increases significantly for OO-MP2 (MD = 37.15 kcal/mol); see Tab. S15. SCS-MP2, for instance, underestimates the decomposition energies with MD = -26.22 kcal/mol, which improves to a value of -4.03 kcal/mol for OO-SCS-MP2 (Tab. S15).

The SIE4x4 set covers the dissociation energies of four positively charged dimers along four different points on their respective dissociation curves. These open-shell doublet radicals were chosen to serve as a test for the one-electron SIE. The one-electron SIE is reduced the more nonlocal Fock exchange is mixed into a functional, as can be seen for the series from DHDFs with about 50% of Fock exchange — PBE0-DH and B2PLYP with MADs of 7.41 and 9.90 kcal/mol, respectively — to the ones with nearly 80 %, for instance PBE0-2 with MAD = 1.80 kcal/mol (Tab. 4). While SIE reacts to Fock-exchange contributions, we deemed this set as interesting for OO due to the doublet-radical nature of the assessed systems. It turns out that for all tested DHDFs OO increases the deviations with PBE0-DH being the least affected (increase by 0.04 kcal/mol) and PBE0-2 the most affected (increase by nearly 0.6 kcal/mol). Again, it seems PBE0-2 is more strongly affected by OO due to its larger perturbative correlation component. Indeed, we see an even stronger effect when going to full MP2, for which the MAD increases from 1.54 to 3.17 kcal/mol. Interestingly, the opposite trend is observed for the two spin-scaled versions. OO-SOS-MP2, with an overall smaller effective perturbative correlation contribution, shows the lowest MAD of all tested methods (MAD=1.04 kcal/mol), as compared to its SOS-MP2 parent (MAD=2.58 kcal/mol).

We conclude this section with a discussion of the G21EA and G21IP sets. For the first set, OO has an almost negligible effect on B2PLYP and PBE0-DH (Tab. 4). For the remaining methods, we observe a larger improvement in the MADs of up to 0.41 kcal/mol (PBE0-2). Similarly, the OO-MP2 variants perform slightly better than their unoptimized

counterparts with MADs improving by up to 0.83 kcal/mol (OO-MP2). Relatively small effects are seen for the IPs. Most changes are below 0.1 kcal/mol with the exception of PBE0-2 whose MAD of 2.15 kcal/mol is reduced to 1.90 kcal/mol with OO. As observed before, MP2 methods show a stronger reaction to the OO scheme with improvements of up to 0.64 kcal/mol (OO-SOS-MP2). That being said, none of the conventional MP2 or OO-MP2 versions have MADs below 3 kcal/mol and they cannot compete with DHDFs.

In summary, we can conclude that for thermochemistry involving open-shell systems the effect of OO cannot be generalized. MP2-type methods depend the strongest on OO, followed by PBE0-2. In the case of the other DHDFs, any effects of OO are usually smaller compared to PBE0-2. However, even when a sizeable effect is observed, OO does not necessarily guarantee improved results. We will continue discussing this aspect after the following study of closed-shell thermochemistry.

Thermochemistry for Closed-Shell Systems

While G21IP in the previous section considered one-electron IPs, and thus a mixture of both open- and closed-shell systems, DIPCS10 involves solely double-ionization potentials of closed-shell species. While one could therefore assume that OO would have a smaller effect on the systems in DIPCS10 than in G21IP, we observe that this is not the case. For instance, while the MAD for G21IP stays nearly constant for (OO-)B2GP-PLYP, OO induces an increase in the MAD by 0.13 kcal/mol for DIPCS10 (Tab. 4). Results for the two DSD methods worsen on average by more than 0.15 kcal/mol. Only B2PLYP, PBE0-2, and PBE-0DH show changes below 0.1 kcal/mol. OO has only a beneficial influence on conventional MP2, for which we see a reduction in its MAD of more than 1 kcal/mol, while SCS-MP2 barely improves and the MAD for SOS-MP2 deteriorates by 0.57 kcal/mol.

The second set that only contains closed-shell species is G2RC with reaction energies for small hydrocarbons and inorganic compounds. For these reactions, the OO-DHDFs

perform slightly worse than their underlying DHDFs, with the exception of OO-PBE0-2 and OO-PBE0-DH. For instance, the B2PLYP result worsens by 0.27 kcal/mol and the DSD-BLYP one by 0.28 kcal/mol, while PBE0-DH remains largely unaffected and PBE0-2 is improved by on average 0.35 kcal/mol. However, the unoptimized and OO variants of PBE0-2 and PBE0-DH give noticeably larger errors than the remaining DHDFs by at least 2.35 kcal/mol. The benefit of OO is therefore questionable. OO-MP2 methods perform significantly worse than their unoptimized variants (see Tab. 4). Once again, we see that the influence of OO differs from method to method and that they cannot be generalized.

For all the DHDFs, OO has a small effect on the MADs for the PA test set of proton affinities with most MADs increasing. Differences between OO-DHDFs and DHDFs range from 0.00 kcal/mol (PBE0-DH) to 0.12 kcal/mol (PBE0-2 and DSD-BLYP). The SCS-MP2 and SOS-MP2 methods slightly worsen, but the MP2 method slightly improves (Tab. 4). In fact, the MAD for OO-MP2 is with 0.69 kcal/mol the best for this test set.

Overall Performance of Orbital-Optimized Methods

The previous sections showed results for a variety of different benchmark sets and we observed at times opposing trends with OO being beneficial in some cases, but having a worsening effect in others. To provide a better picture on whether OO induces an overall improvement or not, we calculated the ratio between the MAD of the OO method and its corresponding unoptimized counterpart for each benchmark set and each method. All ratios are reported in Section SI.4 in the SI. To further facilitate the identification of potential trends we then separately formed the average of those ratios for sets containing only closed-shell species and those containing a mixture of both. In the first case, these sets are all NCI test sets excluding O23, all BH test sets excluding BH76, as well as the DIPCS10, G2RC, and PA sets. The average ratios for all tested methods for those closed-shell sets are shown in Fig. 5a. The remaining test sets contain exclusively open-shell systems or

Figure 5: The ratios of the mean absolute deviations (MADs) of the results given by the orbital-optimized methods (OO-method) to that of their underlying methods (method), averaged over a) benchmark sets that only contain closed-shell species and b) benchmark sets that contain closed- and open-shell species.

reaction involving both closed- and open shell species. The corresponding averaged ratios for each method are depicted in Fig. 5b.

Already in the discussion of each individual set it was observed that conventional OO-MP2 results differed from those of the two spin-component variants. Indeed, OO only improved the results for the DIPCS10 and PA sets, while results deteriorated for the remaining sets; in particular, NCIs were described less accurately by OO-MP2 with deterioration of MADs often close to a factor of 2 or more. The resulting average ratio between OO-MP2 and MP2 is consequently the highest in the entire study with a value of 1.65 (Fig. 5a). In the majority of cases, particularly for the NCI sets, OO-SCS-MP2 and OO-SOS-MP2 perform better than their unoptimized counterparts with ratios of 0.91

and 0.85 kcal/mol, respectively. This comes somewhat as a surprise, as initially it was not expected that OO would have any significant effect on closed-shell species. In the case of test sets involving open-shell species, the results for OO-MP2 are slightly worse (1.07), while the average ratios for the spin-scaled methods are even lower than for the closed-shell systems, with values of 0.71 (OO-SCS-MP2) and 0.66 (OO-SOS-MP2). We can thus conclude that for the calculation of (relative) energies, OO provides no benefit for conventional MP2, but that it indeed improves the SCS-MP2 and SOS-MP2 variants. SCS-MP2 has been found to be the preferred MP2-type method for thermochemistry and we can add that OO-SCS-MP2 could be a useful alternative. That being said it should also be taken into consideration that DHDFs usually outperform SCS-MP2 except for cases in which SIE-related problems may manifest themselves.⁵

The picture for OO-DHDFs is a different one and strongly depends on the nature of the underlying XC component and the fraction of perturbative correlation. The average ratio between OO-PBE0-DH and PBE0-DH is unity in both cases, possibly due to the low amount of perturbative correlation. OO is therefore no viable strategy to obtain better energetic properties with this DFA. BLYP-based DHDFs remain on average unaffected for closed-shell systems and they slightly deteriorate for open-shell ones, with averaged ratios of up to 1.12 for the latter (OO-B2PLYP). It is also noteworthy that the inclusion of spin-component scaling, as done for DSD-BLYP, does not have any additional benefits from OO, as opposed to the aforementioned findings for SCS-MP2: the averaged ratio between OO-DSD-BLYP and DSD-BLYP for the open-shell test sets is 1.03 and thus comparable with the one for OO-B2GP-PLYP (1.06). DSD-PBEP86, which contains PBE exchange, is affected the most for the closed-shell cases with an average ratio of 1.13. OO-DSD-PBEP86, however, is a slight improvement for the open-shell cases (ratio of 0.88). PBE0-2, which bears formal resemblance to PBE0-DH, but contains 50% of perturbative correlation, shows similar trends as DSD-PBEP86 for the closed-shell cases (Fig. 5a). For

Figure 6: The ratios of the mean absolute deviations (MADs) of the results given by the methods using the (aug’)-def2-TZVP basis-set (method/TZ) to that of the methods using the (aug’)-def2-QZVP basis-set (method/QZ), averaged over all the benchmark sets. The same is done for the orbital-optimized methods (OO-methods).

open- shell cases, however, it shows the largest relative improvement with an average ratio of 0.84. While this average value looks promising, we remind the reader that OO-PBE0-2 is often outperformed by conventional DHDFs. The decision to favor OO-PBE0-2 over, for instance, DSD-PBEP86 has to be made on a case-by-case basis.

Basis Set Dependence of Orbital-Optimized Methods

Using a large AO basis set, such as (aug’)-def2-QZVP, allowed us to rule out any basis-set artifacts and to get a good understanding of the inherent effect of OO on DHDFs and MP2 variants. In this section, we briefly turn our attention to the basis-set dependence of these methods to see if dispersion-corrected OO-DHDFs combined with a TZ-quality AO basis set could be a viable option in applications. For that purpose, we carried out additional calculations with (aug’)-def2-TZVP; all results are reported in Section SI.5 in the SI. Usage of even smaller AO basis sets for conventional DHDFs has been discouraged,⁵ which is why we regard an analysis of OO-DHDFs with those as superfluous.

Similarly to the discussion in the previous section, we facilitate our analysis by pre-

senting ratios between the MADs obtained at the TZ and QZ levels. Individual ratios are shown in the SI in Tabs. S60-S68. In this section, we mostly focus on the average value of those ratios for one particular method, as shown in Fig. 6. In all cases, the average ratios are larger than unity, which indicates that, regardless of the method, the QZ basis set is the more accurate choice. With the exception of DSD-PBEP86-D3, the difference between the average ratios of the DHDFs and those of their OO counterparts is 0.1 or less, indicating that the OO variant inherits the basis-set dependence of the DHDF parent. The dispersion-corrected PBE0-DH variants show the smallest basis-set dependence of all tested methods (average ratios of 1.16 for PBE0-DH-D3 and 1.18 for OO-PBE0-DH-D3). This by itself was somewhat expected; Goerigk and Grimme already showed that DHDFs with small fractions of MP2 correlation have a small basis-set dependence similar to that of hybrids.⁵ The fact that the three BLYP-based DHDFs and DSD-PBEP86 show larger ratios confirm this picture, as their fractions of perturbative correlation are larger. Their average ratios range from 1.52 to 1.81 (Fig. 6). Somewhat contrary to that observation is that PBE0-2-D3 and OO-PBE0-2-D3 have averaged ratios similar to that of (OO)-PBE0-DH-D3 (1.22 for PBE0-2-D3 and 1.29 for OO-PBE0-2-D3). A closer inspection of Tab. S62 indeed shows that many benchmark sets show smaller MADs for TZ than for QZ, for instance MB16-43 and SIE4x4, thus counterbalancing those that have ratios larger than unity, such as WATER27 or O23. This compensating effect is not observed for methods such as (OO-)B2PLYP-D3, where the ratios are more balanced across the 20 sets (Tab. S60).

Turning to the MP2-based methods, we note that SOS-MP2 and OO-SOS-MP2 have average ratios that are as low as the ones for the PBE-based (OO-)DHDFs (1.20 for SOS-MP2 and 1.30 for OO-SOS-MP2). While the basis-set dependence for conventional SCS-MP2 is similar (1.30) it increases drastically to 1.73 for OO-SCS-MP2, which is why we do not recommend its usage with a TZ basis. MP2 and OO-MP2, on the other hand, show

the opposite trend, with OO-MP2 being comparable to OO-SOS-MP2 (average ratio of 1.27), while MP2 itself has an average ratio of 1.52. We can therefore conclude again that the behavior of OO-MP2 methods differs substantially from the behavior of OO-DHDFs.

In conclusion, we can say that not much is gained from using OO methods with AO basis sets smaller than QZ quality, with the following exceptions: OO-PBE0-DH-D3, OO-PBE02-D3, OO-SOS-MP2, and OO-MP2. However, bearing in mind the overall performance for thermochemistry for those methods, we seriously doubt the usefulness of this approach, except for a handful of special cases.

Summary and Conclusions

We presented a comprehensive analysis of the effect of orbital optimization (OO) on various double-hybrid density functionals (DHDFs) and MP2 variants. Previous studies on OO-DHDFs were limited to three non-empirical PBE-based DHDFs. Herein, we tested for the first time OO-DHDFs based on the best-performing⁶ DHDFs for main-group chemistry. Previous studies already demonstrated how the OO technique improves dipole-moment and nuclear-force dissociation curves of diatomics,^{41,45} which is why this work mainly focused on the treatment of main-group thermochemistry, kinetics, and noncovalent interactions represented by 20 different sets comprising 621 reference energies. We also complemented a previous analysis on OO-DHDFs applied to the singlet-triplet splitting in oligoacenes, and for the first time discussed the effect of OO on London-dispersion effects. The majority of our calculations were carried out with a large quadruple- ζ atomic-orbital (AO) basis set, but for the first time we also investigated the basis-set dependence of OO methods. Therefore, our study offered a detailed insight into the OO technique that allowed us to reach a conclusive verdict on its potential benefits.

OO improved most DHDFs for the singlet-triplet energy gaps of linear oligoacenes,

just confirming a limited study on a single PBE-based OO-DHDF.⁴⁵ However, despite this finding, we also observed that the best method for this problem was in fact the unoptimized PBE0-DH functional, thus rendering the application of OO unnecessary. An analysis of the argon-dimer dissociation showed completely different effects of OO for wave-function-based MP2 variants and DHDFs. OO-MP2 severely over-stabilized the argon dimer, while OO-SCS/SOS-MP2 were in very good agreement with a highly accurate *ab-initio* dissociation curve. OO-DHDFs, however, were not influenced at all by OO for the argon dimer. This means that they require London-dispersion corrections, which is why we presented new damping parameters to be used with Grimme’s DFT-D3(BJ) correction.

Previously, it was outlined how the OO technique had the strongest influence on open-shell systems for which they reduced the notorious spin-contamination problem.^{37,39,40} Our investigation of the 20 thermochemical benchmark sets confirmed how the effect of OO on DHDFs and MP2 methods strongly depended on whether the system under investigation had closed- or open-shell character. Our main findings were:

- 1) For thermochemical benchmark sets that only involved closed-shell systems, OO was mostly negligible for DHDFs. Exceptions were relative energies in conformers as well as binding energies in noncovalently bound dimers and in water clusters for which we observed improvements by up to 27 % for OO-B2PLYP, OO-B2GP-PLYP, and OO-DSD-BLYP. PBE-based OO-DHDFs, including the previously advertised OO-PBE0-2 and OO-PBE0-DH, did not show this trend. In fact, we observed a worsening for the same test sets by up to 64 %. Also, the treatment of the thermochemical closed-shell test sets showed a mixed picture; in the majority of cases, OO-DHDFs were very close to DHDFs. OO-MP2 showed some of the worst results of the entire study, while its spin-scaled variants showed marginal improvements.
- 2) The effect of OO on the DHDFs was more noticeable for the benchmark sets that con-

tain a mixture of closed- and open-shell systems. However, this does not imply that OO always improved the results. In fact, only OO-PBE0-2, OO-DSD-PBEP86, OO-SCS-MP2, and OO-SOS-MP2 performed on average significantly better than their unoptimized counterparts, with OO-SOS-MP2 displaying the biggest improvements. The remaining methods were relatively impartial to the application of OO in the majority of cases.

- 3) We observed that OO did not always improve the description of reaction barrier heights for the DHDFs. In fact, for reactions that involved open-shell species, functionals such as DSD-BLYP — one of the most accurate DFT methods — showed an increase of 36 % in the error for the BH76 set, while PBE0-2 — one of the worst DHDFs for thermochemistry and kinetics — showed an improvement of 31%. That being said, the average errors for OO-PBE0-2 were still worse than those of unoptimized DSD-BLYP. The only methods for which OO showed a consistent improvement for barrier heights of reactions involving open-shell species were the three MP2 variants.
- 4) A comparison between quadruple- ζ and triple- ζ AO basis sets revealed that in the majority of cases OO techniques inherited the basis-set dependence of their unoptimized counterparts, with results deteriorating on average by at least 18% (OO-PBE0-DH) and up to 81% (OO-DSD-BLYP) for the smaller basis set.
- 5) Initially, some of the results for the thermochemical benchmark sets indicated that the amount of perturbative correlation in a DHDF determined the impact of OO. For instance, PBE0-DH with a fraction of only 16% of nonlocal correlation was influenced the least by OO, while PBE0-2 with 50% of nonlocal correlation showed sometimes behaviors similar to the wave-function based MP2 variants. This initial hypothesis, however, could not be seen throughout the entire study. For instance, the basis-set dependence of OO-PBE0-2 was similar to that of OO-PBE0-DH.

6) Interestingly, the only OO-DHDFs with convergence problems were the PBE-based ones. Occasionally, also MP2-based OO methods had convergence issues, but never any BLYP-based OO-DHDFs (see Tab. 1 and SI). This indicated that PBE-based OO-DHDFs are less robust.

We conclude that the influence of OO is strongly method dependent. While a large difference between MP2 variants and DHDFs may have been expected, our comprehensive study also showed that there is a significant difference between the tested DHDFs. Even when results improved, there was no guarantee that the more costly OO method outperformed conventional DHDFs. We therefore conclude that OO should be reserved as a technique to solve specific problems, such as dipole-moment and nuclear-force dissociation curves; see Ref. 44. Moreover, our results suggest that OO is not a viable path to follow in the ongoing quest to improve DHDFs. In fact, we can confirm previous conclusions by Chan et al. that improving the already very good conventional DHDFs is a difficult task.³⁵ For practical purposes, we recommend to the mere method user to use the best-performing DHDFs DSD-BLYP and DSD-PBEP86, as shown in the recent comprehensive GMTKN55 benchmark study.⁶ We also hope that our in-depth discussion of OO provided new insights to method developers and encourage to seek for alternative strategies to improve on the currently best DHDFs.

Supporting Information Available

DFT-D3(BJ) damping parameters; argon-dimer dissociation curves; statistical analyses of all methods; basis-set dependence study.

Acknowledgement

This work was carried out as part of an Australian Research Council Discovery Early Career Researcher Award awarded to L. G. (project number DE140100550). We also acknowledge generous allocation of computing time from the National Computational Infrastructure (NCI) National Facility within the National Computational Merit Allocation Scheme (project fk5) and from Melbourne Bioinformatics, formerly known as the Victorian Life Science Computation Initiative (project RA0005).

References

- (1) Becke, A. D. A new mixing of Hartree-Fock and local-density-functional theories. *J. Chem. Phys.* **1993**, *98*, 1372–1377.
- (2) Grimme, S. Semiempirical hybrid density functional with perturbative second-order correlation. *J. Chem. Phys.* **2006**, *124*, 034108.
- (3) Møller, C.; Plesset, M. S. Note on an approximation treatment for many-electron systems. *Phys. Rev.* **1934**, *46*, 618–622.
- (4) Perdew, J. P.; Schmidt, K. In *Density Functional Theory and its applications to materials*, AIP Conference Proceedings; Doren, V. E. V., Alsenoy, C. V., Geerlings, P., Eds.; AIP Publishing: Melville, 2001; Vol. 577; pp 1–20.
- (5) Goerigk, L.; Grimme, S. A thorough benchmark of density functional methods for general main group thermochemistry, kinetics, and noncovalent interactions. *Phys. Chem. Chem. Phys.* **2011**, *13*, 6670–6688.
- (6) Goerigk, L.; Hansen, A.; Bauer, C.; Ehrlich, S.; Najibi, A.; Grimme, S. A look at the density functional theory zoo with the advanced GMTKN55 database for general main

- group thermochemistry, kinetics and noncovalent interactions. *Phys. Chem. Chem. Phys.* **2017**, *19*, 32184–32215.
- (7) Grimme, S.; Schwabe, T. Towards chemical accuracy for the thermodynamics of large molecules: new hybrid density functionals including non-local correlation effects. *Phys. Chem. Chem. Phys.* **2006**, *8*, 4398–4401.
- (8) Benighaus, T.; DiStasio Jr., R. A.; Lochan, R. C.; Chai, J.-D.; Head-Gordon, M. Semiempirical double-hybrid density functional with improved description of long-range correlation. *J. Phys. Chem. A* **2008**, *112*, 2702–2712.
- (9) Karton, A.; Tarnopolsky, A.; Lamere, J.-F.; Schatz, G. C.; Martin, J. M. L. Highly accurate first-principles benchmark data sets for the parametrization and validation of density functional and other approximate methods. Derivation of a robust, generally applicable double-hybrid functional for thermochemistry and thermochemical kinetics. *J. Phys. Chem. A* **2008**, *112*, 12868–12886.
- (10) Chai, J.-D.; Head-Gordon, M. Long-range corrected double-hybrid density functionals. *J. Chem. Phys.* **2009**, *131*, 174105.
- (11) Goerigk, L.; Grimme, S. Efficient and accurate double-hybrid-meta-GGA density functionals — evaluation with the extended GMTKN30 database for general main group thermochemistry, kinetics, and noncovalent interactions. *J. Chem. Theory Comput.* **2011**, *7*, 291–309.
- (12) Kozuch, S.; Gruzman, D.; Martin, J. M. L. DSD-BLYP: a general purpose double hybrid density functional including spin component scaling and dispersion correction. *J. Phys. Chem. C* **2010**, *114*, 20801–20808.
- (13) Kozuch, S.; Martin, J. M. L. DSD-PBEP86: in search of the best double-hybrid DFT

- with spin-component scaled MP2 and dispersion corrections. *Phys. Chem. Chem. Phys.* **2011**, *13*, 20104–20107.
- (14) Kozuch, S.; Martin, J. M. L. Spin-component-scaled double hybrids: an extensive search for the best fifth-rung functionals blending DFT and perturbation theory. *J. Comput. Chem.* **2013**, *34*, 2327–2344.
 - (15) Sharkas, K.; Toulouse, J.; Savin, A. Double-hybrid density-functional theory made rigorous. *J. Chem. Phys.* **2011**, *134*, 064113.
 - (16) Bremond, E.; Adamo, C. Seeking for parameter-free double-hybrid functionals: the PBE0-DH model. *J. Chem. Phys.* **2011**, *135*, 024106.
 - (17) Chai, J. D.; Mao, S. P. Seeking for reliable double-hybrid density functionals without fitting parameters: the PBE0-2 functional. *Chem. Phys. Lett.* **2012**, *538*, 121–125.
 - (18) Bremond, E.; Sancho-Garcia, J.; Perez-Jimenez, A.; Adamo, C. Communication: double-hybrid functionals from adiabatic-connection: the QIDH model. *J. Chem. Phys.* **2014**, *141*, 031101.
 - (19) Alipour, M. Seeking for spin-opposite-scaled double-hybrid models free of fitted parameters. *J. Phys. Chem. A* **2016**, *120*, 3726–3730.
 - (20) Roch, L. M.; Baldrige, K. K. General optimization procedure towards the design of a new family of minimal parameter spin-component-scaled double-hybrid density functional theory. *Phys. Chem. Chem. Phys.* **2017**, *19*, 26191–26200.
 - (21) Grimme, S.; Neese, F. Double-hybrid density functional theory for excited electronic states of molecules. *J. Chem. Phys.* **2007**, *127*, 154116.

- (22) Goerigk, L.; Moellmann, J.; Grimme, S. Computation of accurate excitation energies for large organic molecules with double-hybrid density functionals. *Phys. Chem. Chem. Phys.* **2009**, *11*, 4611–4620.
- (23) Schwabe, T.; Goerigk, L. Time-dependent double-hybrid density functionals with spin-component and spin-opposite scaling. *J. Chem. Theory Comput.* **2017**, *13*, 4307–4323.
- (24) Goerigk, L.; Grimme, S. Double-hybrid density functionals. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2014**, *4*, 576–600.
- (25) Grimme, S. Improved second-order MøllerPlesset perturbation theory by separate scaling of parallel- and antiparallel-spin pair correlation energies. *J. Chem. Phys.* **2003**, *118*, 9095–9102.
- (26) Grimme, S.; Goerigk, L.; Fink, R. F. Spin-component-scaled electron correlation methods. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 886–906.
- (27) Jung, Y. S.; Lochan, R. C.; Dutoi, A. D.; Head-Gordon, M. Scaled opposite-spin second order Moller-Plesset correlation energy: an economical electronic structure method. *J. Chem. Phys.* **2004**, *121*, 9793–9802.
- (28) Grimme, S.; Steinmetz, M. A computationally efficient double hybrid density functional based on the random phase approximation. *Phys. Chem. Chem. Phys.* **2016**, *18*, 20926–20937.
- (29) Janesko, B. G.; Henderson, T. M.; Scuseria, G. E. Long-range-corrected hybrids including random phase approximation correlation. *J. Chem. Phys.* **2009**, *130*, 081105.
- (30) Irelan, R. M.; Henderson, T. M.; Scuseria, G. E. Long-range-corrected hybrids using

- a range-separated Perdew-Burke-Ernzerhof functional and random phase approximation correlation. *J. Chem. Phys.* **2011**, *135*, 094105.
- (31) Garza, A. J.; Bulik, I. W.; Alencar, A. G. S.; Sun, J.; Perdew, J. P.; Scuseria, G. E. Combinations of coupled cluster, density functionals, and the random phase approximation for describing static and dynamic correlation, and van der Waals interactions. *Mol. Phys.* **2015**, *114*, 997–1018.
- (32) Mezei, P. D.; Csonka, G. I.; Ruzsinszky, A.; Kállay, M. Construction and application of a new dual-hybrid random phase approximation. *J. Chem. Theory Comput.* **2015**, *11*, 4615–4626.
- (33) Mezei, P. D.; Csonka, G. I.; Ruzsinszky, A.; Kállay, M. Construction of a spin-component scaled dual-hybrid random phase approximation. *J. Chem. Theory Comput.* **2017**, *13*, 796–803.
- (34) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. A fifth-order perturbation comparison of electron correlation theories. *Chem. Phys. Lett.* **1989**, *157*, 479–483.
- (35) Chan, B.; Goerigk, L.; Radom, L. On the inclusion of post-MP2 contributions to double-hybrid density functionals. *J. Comput. Chem.* **2016**, *37*, 183–193.
- (36) Lochan, R. C.; Head-Gordon, M. Orbital-optimized opposite-spin scaled second-order correlation: an economical method to improve the description of open-shell molecules. *J. Chem. Phys.* **2007**, *126*, 164101.
- (37) Neese, F.; Schwabe, T.; Kossmann, S.; Schirmer, B.; Grimme, S. Assessment of orbital-optimized, spin-component scaled second-order many-body perturbation theory for thermochemistry and kinetics. *J. Chem. Theory Comput.* **2009**, *5*, 3060–3073.

- (38) Bozkaya, U.; Turney, J. M.; Yamaguchi, Y.; Schaefer III, H. F.; Sherrill, C. D. Quadratically convergent algorithm for orbital optimization in the orbital-optimized coupled-cluster doubles method and in orbital-optimized second-order Møller-Plesset perturbation theory. *J. Chem. Phys.* **2011**, *135*, 104103.
- (39) Soydas, E.; Bozkaya, U. Accurate open-shell noncovalent interaction energies from the orbital-optimized Møller-Plesset perturbation theory: achieving CCSD quality at the MP2 level by orbital optimization. *J. Chem. Theory Comput.* **2013**, *9*, 4679–4683.
- (40) Bozkaya, U. Orbital-optimized second-order perturbation theory with density-fitting and Cholesky decomposition approximations: an efficient implementation. *J. Chem. Theory Comput.* **2014**, *10*, 2371–2378.
- (41) Stuck, D.; Head-Gordon, M. Regularized orbital-optimized second-order perturbation theory. *J. Chem. Phys.* **2013**, *139*, 244109.
- (42) Bozkaya, U. Orbital-optimized linearized coupled-cluster doubles with density-fitting and Cholesky decomposition approximations: an efficient implementation. *Phys. Chem. Chem. Phys.* **2016**, *18*, 11362–11373.
- (43) Myhre, R. H. Demonstrating that the nonorthogonal orbital optimized coupled cluster model converges to full configuration interaction. *J. Chem. Phys.* **2018**, *148*, 094110.
- (44) Peverati, R.; Head-Gordon, M. Orbital optimized double-hybrid density functionals. *J. Chem. Phys.* **2013**, *139*, 024110.
- (45) Sancho-Garcia, J. C.; Perez-Jimenez, A. J.; Savarese, M.; Bremond, E.; Adamo, C. Importance of orbital optimization for double-hybrid density functionals: application of the OO-PBE-QIDH model for closed- and open-shell systems. *J. Phys. Chem. A* **2016**, *120*, 1756–1762.

- (46) Schwabe, T.; Grimme, S. Double-hybrid density functionals with long-range dispersion corrections: higher accuracy and extended applicability. *Phys. Chem. Chem. Phys.* **2007**, *9*, 3397–3406.
- (47) Goerigk, L.; Grimme, S. A general database for main group thermochemistry, kinetics, and noncovalent interactions — assessment of common and reparameterized (meta-)GGA density functionals. *J. Chem. Theory Comput.* **2010**, *6*, 107–126.
- (48) Karton, A.; Martin, J. M. L. Explicitly correlated Wn theory: W1-F12 and W2-F12. *J. Chem. Phys.* **2012**, *136*, 124114.
- (49) Goerigk, L. Treating London-dispersion effects with the latest Minnesota density functionals: problems and possible solutions. *J. Phys. Chem. Lett.* **2015**, *6*, 3891–3896.
- (50) Hajgato, B.; Szieberth, D.; Geerlings, P.; Proft, F. D.; Deleuze, M. A benchmark theoretical study of the electronic ground state and of the singlet-triplet split of benzene and linear acenes. *J. Chem. Phys.* **2009**, *131*, 224321.
- (51) Jurečka, P.; Sponer, J.; Cerný, J.; Hobza, P. Benchmark database of accurate (MP2 and CCSD(T) complete basis set limit) interaction energies of small model complexes, DNA base pairs, and amino acid pairs. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1985–1993.
- (52) Marshall, M. S.; Burns, L. A.; Sherrill, C. D. Basis set convergence of the coupled-cluster correction, $\delta_{MP2}^{CCSD(T)}$: best practices for benchmarking non-covalent interactions and the attendant revision of the S22, NBC10, HBC6, and HSG databases. *J. Chem. Phys.* **2011**, *135*, 194102.
- (53) Csonka, G. I.; French, A. D.; Johnson, G. P.; Stortz, C. A. Evaluation of density functionals and basis sets for carbohydrates. *J. Chem. Theory Comput.* **2009**, *5*, 679–692.

- (54) Bryantsev, V. S.; Diallo, M. S.; van Duin, A. C. T.; Goddard III, W. A. Evaluation of B3LYP, X3LYP, and M06-class density functionals for predicting the binding energies of neutral, protonated, and deprotonated water clusters. *J. Chem. Theory Comput.* **2009**, *5*, 1016–1026.
- (55) Manna, D.; Kesharwani, M. K.; Sylvetsky, N.; Martin, J. M. L. Conventional and explicitly correlated ab Initio benchmark study on water clusters: revision of the BEGDB and WATER27 data sets. *J. Chem. Theory Comput.* **2017**, *13*, 3136–3152.
- (56) Řeha, D.; Valdes, H.; Vondrasek, J.; Hobza, P.; Abu-Riziq, A.; Crews, B.; de Vries, M. S. Structure and IR spectrum of phenylalanyl-glycyl-glycine tripeptide in the gas-phase: IR/UV experiments, ab initio quantum chemical calculations, and molecular dynamic simulations. *Chem. Eur. J.* **2005**, *11*, 6803–6817.
- (57) Goerigk, L.; Karton, A.; Martin, J. M. L.; Radom, L. Accurate quantum chemical energies for tetrapeptide conformations: why MP2 data with an insufficient basis set should be handled with caution. *Phys. Chem. Chem. Phys.* **2013**, *15*, 7028–7031.
- (58) Zhao, Y.; González-García, N.; Truhlar, D. G. Benchmark database of barrier heights for heavy atom transfer, nucleophilic substitution, association, and unimolecular reactions and its use to test theoretical methods. *J. Phys. Chem. A* **2005**, *109*, 2012–2018.
- (59) Zhao, Y.; Lynch, B. J.; Truhlar, D. G. Multi-coefficient extrapolated density functional theory for thermochemistry and thermochemical kinetics. *Phys. Chem. Chem. Phys.* **2005**, *17*, 43–52.
- (60) Karton, A.; Goerigk, L. Accurate reaction barrier heights of pericyclic reactions: surprisingly large deviations for the CBS-QB3 composite method and their consequences in DFT benchmark studies. *J. Comput. Chem.* **2015**, *36*, 622–632.

- (61) Karton, A.; O'Reilly, R. J.; Chan, B.; Radom, L. Determination of barrier heights for proton exchange in small water, ammonia, and hydrogen fluoride clusters with G4(MP2)-type, MPn, and SCS-MPn procedures — a caveat. *J. Chem. Theory Comput.* **2012**, *8*, 3128–3136.
- (62) Goerigk, L.; Sharma, R. The INV24 test set: how well do quantum-chemical methods describe inversion and racemization barriers? *Can. J. Chem.* **2016**, *94*, 1133–1143.
- (63) Karton, A.; Daon, S.; Martin, J. M. L.; Ruscic, B. W4-11: a high-confidence benchmark dataset for computational thermochemistry derived from first-principles W4 data. *Chem. Phys. Lett.* **2011**, *510*, 165–178.
- (64) Curtiss, L. A.; Raghavachari, K.; Trucks, G. W.; Pople, J. A. Gaussian-2 theory for molecular energies of first- and second-row compounds. *J. Chem. Phys.* **1991**, *94*, 7221–7230.
- (65) Parthiban, S.; Martin, J. M. L. Assessment of W1 and W2 theories for the computation of electron affinities, ionization potentials, heats of formation, and proton affinities. *J. Chem. Phys.* **2001**, *114*, 6014–6029.
- (66) Zhao, Y.; Truhlar, D. G. Assessment of density functionals for π systems: energy differences between cumulenes and poly-ynes; proton affinities, bond length alternation, and torsional potentials of conjugated polyenes; and proton affinities of conjugated Schiff bases. *J. Phys. Chem. A* **2006**, *110*, 10478–10486.
- (67) Neese, F. The ORCA program system. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 73–78.
- (68) ORCA — an ab initio, density functional and semiempirical program package, V. 4.0.1, F. Neese, MPI für Chemische Energiekonversion, Mülheim a. d. Ruhr (Germany), 2017.

- (69) Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **1988**, *38*, 3098–3100.
- (70) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **1988**, *37*, 785–789.
- (71) Miehlich, B.; Savin, A.; Stoll, H.; Preuss, H. Results obtained with the correlation energy density functionals of Becke and Lee, Yang and Parr. *Chem Phys. Lett.* **1989**, *157*, 200–206.
- (72) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev Lett* **1996**, *77*, 3865–3868.
- (73) Perdew, J. P. Density-functional approximation for the correlation-energy of the inhomogeneous electron gas. *Phys. Rev. B* **1986**, *33*, 8822–8824.
- (74) Perdew, J. P. Erratum: Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Phys. Rev. B* **1986**, *34*, 7406.
- (75) Kendall, R. A.; Dunning, T. H.; Harrison, R. J. Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. *J. Chem. Phys.* **1992**, *96*, 6796–6806.
- (76) Dunning, Jr., T. H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023.
- (77) Izsak, R.; Neese, F. An overlap fitted chain of spheres exchange method. *J. Chem. Phys.* **2011**, *135*, 144105.
- (78) Eichkorn, K.; Weigend, F.; Treutler, O.; Ahlrichs, R. Auxiliary basis sets for main row atoms and transition metals and their use to approximate Coulomb potentials. *Theor. Chem. Acc.* **1997**, *97*, 119–124.

- (79) Weigend, F.; Köhn, A.; Hättig, C. Efficient use of the correlation consistent basis sets in resolution of the identity MP2 calculations. *J. Chem. Phys.* **2002**, *116*, 3175–3183.
- (80) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.
- (81) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem* **2011**, *32*, 1456–1465.
- (82) Ruscic, B. Uncertainty quantification in thermochemistry, benchmarking electronic structure computations, and active thermochemical tables. *Int. J. Quantum Chem.* *114*, 1097–1101.
- (83) Antony, J.; Grimme, S. Is spin-component scaled second-order Møller-Plesset perturbation theory an appropriate method for the study of noncovalent interactions in molecules? *J. Phys. Chem. A* **2007**, *111*, 4862–4868.
- (84) Distasio, Jr., R. A.; Head-Gordon, M. Optimized spin-component scaled second-order Møller-Plesset perturbation theory for intermolecular interaction energies. *Mol. Phys.* **2007**, *105*, 1073–1083.
- (85) Hill, J. G.; Platts, J. A. Spin-component scaling methods for weak and stacking interactions. *J. Chem. Theory Comput.* **2007**, *3*, 80–85.
- (86) King, R. A. On the accuracy of spin-component-scaled perturbation theory (SCS-MP2) for the potential energy surface of the ethylene dimer. *Mol. Phys.* **2009**, *107*, 789–795.

TOC Graphic