

Evaluation of DFT-D3 dispersion corrections for various structural benchmark sets

Heiner Schröder, Jens Hühnert, and Tobias Schwabe

Citation: *J. Chem. Phys.* **146**, 044115 (2017); doi: 10.1063/1.4974840

View online: <http://dx.doi.org/10.1063/1.4974840>

View Table of Contents: <http://aip.scitation.org/toc/jcp/146/4>

Published by the American Institute of Physics

COMPLETELY
REDESIGNED!



PHYSICS
TODAY

Physics Today Buyer's Guide
Search with a purpose.

Evaluation of DFT-D3 dispersion corrections for various structural benchmark sets

Heiner Schröder, Jens Hünnert, and Tobias Schwabe

Center for Bioinformatics and Institute of Physical Chemistry, University of Hamburg, Bundesstraße 43, 20146 Hamburg, Germany

(Received 18 November 2016; accepted 11 January 2017; published online 30 January 2017)

We present an evaluation of our newly developed density functional theory (DFT)-D3 dispersion correction D3(CSO) in comparison to its predecessor D3(BJ) for geometry optimizations. Therefore, various benchmark sets covering bond lengths, rotational constants, and center of mass distances of supramolecular complexes have been chosen. Overall both corrections give accurate structures and show no systematic differences. Additionally, we present an optimized algorithm for the computation of the DFT-D3 gradient, which reduces the formal scaling of the gradient calculation from $\mathcal{O}(N^3)$ to $\mathcal{O}(N^2)$. Published by AIP Publishing. [<http://dx.doi.org/10.1063/1.4974840>]

I. INTRODUCTION

Density functional theory (DFT) has emerged as the workhorse of quantum chemistry due to its favorable scaling behavior in comparison to wavefunction based methods. Even in those cases, where properties are computed via wavefunction methods, this is usually done on the basis of a DFT-(pre-)optimized structure. Unfortunately, standard density functional approximations (DFAs) have some serious shortcomings like the self-interaction error and the inability to correctly describe long-ranged London dispersion interactions.^{1,2} These dispersion interactions, however, play an important role in the formation of larger structures, e.g., such as in bulky organic molecules,³ biomolecules,^{4–6} and polymers and thus an accurate description is critical in order to obtain meaningful equilibrium structures.

Various schemes for the correction of this shortcoming have been proposed (for a review see, e.g., Refs. 7 and 8). In this work, we focus on the simple, yet accurate DFT-D3 method.⁹ It has been shown, that the DFT-D3 dispersion correction significantly improves the performance of standard DFAs and related semiempirical methods for the optimization of equilibrium structures.^{10–13}

Recently, we introduced a reformulation¹⁴ of the D3(BJ) dispersion correction,¹⁰ which only relies on C_6 dispersion coefficients (CSO variant) and also requires only one single empirical parameter instead of three as for D3(BJ), while keeping the same accuracy for thermochemistry. Note that an alternative linear soft damping has been introduced just recently, which also shows good results for some DFAs when only C_6 dispersion coefficients are used.¹⁵

This work extends the evaluation of D3(CSO) to geometry optimizations. Therefore, an analytical gradient is introduced for D3(CSO). The performance of the new correction for geometry optimization is evaluated for several benchmark sets, using seven standard DFAs from different functional classes. Additionally, an optimized algorithm for the computation of the DFT-D3 gradient is presented.

II. THEORY

In the DFT-D3 scheme, the total energy E_{tot} of a system is the sum of the Kohn-Sham energy E_{KS} and the dispersion correction E_{disp} ,

$$E_{tot} = E_{KS} + E_{disp}. \quad (1)$$

The total dispersion energy E_{disp} here is the sum over all individual attractive atom pair (AB) contributions, which depends on the dispersion coefficients C_n^{AB} up to n th order and the interatomic distance R_{AB} . Interactions with an odd value of n and many body terms are usually neglected even though they may have non-negligible contributions to the dispersion interaction in some cases. In general, the former are neglected, since they cancel out in freely rotating systems¹⁶ and the latter because the contributions are usually small (<5%–10%).⁹ Furthermore, DFT-D3 is restricted to molecules in the electronic ground state, but dispersion effects and related excitonic coupling play an important role in solvatochromism.¹⁷

The general formula for empirical dispersion corrections is thus given as

$$E_{disp} = - \sum_{AB} \sum_{n=6,8,10,\dots} \frac{C_n^{AB}}{R_{AB}^n} f_{damp}(R_{AB}, A, B). \quad (2)$$

The leading term of the expansion in Equation (2) with $n = 6$ ensures the correct asymptotic behavior of the potential, while the higher order terms influence its shape at shorter distances. To avoid near-singularities at small R_{AB} and double counting effects of electron correlation at medium distances, a damping function f_{damp} is employed, which is fitted individually to empirical data for every DFA.

An analysis of Koide showed that the dispersion energy does not vanish, but becomes constant for small interatomic distances,¹⁸ which the damping function may take into account. This was first incorporated by Becke and Johnson into their parameter-free dispersion correction¹⁹ and later successfully adopted for the DFT-D3 method.¹⁰

Our analysis¹⁴ of the DFT-D3(BJ) correction showed that in this method the higher order term acts as an interpolation function between the $R_{AB} \rightarrow 0$ and $R_{AB} \rightarrow \infty$ limits which is specific to every DFA. The C_8^{AB} coefficients used there are calculated as a simple linear function of the C_6^{AB} coefficients. Instead of the higher order term, a sigmoidal interpolation function is employed for our D3(CSO)-correction, which requires only one DFA-specific empirical parameter instead of three for D3(BJ) without loosing accuracy for a huge benchmark set on thermochemistry. A similar *ansatz* incorporating only the $n = 6$ term was already made by Jurečka *et al.*²⁰ However, various schemes determining and incorporating higher order coefficients from first principles exist.^{19,21,22}

A general expression for the energy of all DFT-D3 variants can be reformulated as follows:

$$E_{\text{disp}}^{\text{D3}} = - \sum_{AB} C_6^{AB} f_{\text{remain}}(R_{AB}, A, B). \quad (3)$$

Notice that the following steps do not rely on the specific form of the remaining function f_{remain} , which constitutes the difference between the individual D3-variants and also includes the C_8^{AB} -based terms. In detail, the expressions are as follows (see also Appendix B):

$$\begin{aligned} \text{D3(Zero): } f_{\text{remain}}^{\text{Zero}}(R_{AB}, A, B) &= \left(f_{\text{damp},6}(R_{AB}) \frac{s_6}{(R_{AB})^6} \right. \\ &\quad \left. + f_{\text{damp},8}(R_{AB}) \frac{s_8 k_{6,8}^{AB}}{(R_{AB})^8} \right), \end{aligned} \quad (4)$$

$$\begin{aligned} \text{D3(BJ): } f_{\text{remain}}^{\text{BJ}}(R_{AB}, A, B) &= \frac{s_6}{(R_{AB})^6 + (a_1 R_0^{AB} + a_2)^6} \\ &\quad + \frac{s_8 k_{6,8}^{AB}}{(R_{AB})^8 + (a_1 R_0^{AB} + a_2)^8}, \end{aligned} \quad (5)$$

$$\begin{aligned} \text{D3(CSO): } f_{\text{remain}}^{\text{CSO}}(R_{AB}, A, B) &= \left(s_6 + \frac{a_1}{1 + \exp(R_{AB} - 2.5 R_0^{AB})} \right) \\ &\quad \times \frac{1}{R_{AB}^6 + (2.5^2)^6}. \end{aligned} \quad (6)$$

In contrast to its predecessors, effects of the neighboring atoms are taken into account for the D3-method. To distinguish, e.g., between a carbon atom in acetylene or ethane, coordination number (CN) dependent dispersion coefficients are used. The determination of CNs was designed in such a way, that—following chemical intuition—CNs close to integer values are obtained. Higher coordination numbers correspond to a more squeezed electron density and thus to a lower polarizability from which the dispersion coefficients are derived. Going the other way round, we calculated atomic polarizabilities from D3-dispersion coefficients and used them successfully for polarizable embedding calculations.²³ The same relations are exploited by Stöhr *et al.*, who recently presented a method to derive C_6 coefficients and polarizabilities from charge-population analysis, which allows their determination using semiempirical or tight-binding methods.²⁴

The value of the dispersion coefficients C_6^{AB} in DFT-D3 is interpolated from precomputed dispersion coefficients $C_{6,\text{rev}}^{AB}$ of simple reference systems at points (i,j) by a Gaussian-distance weighted average. The parameters k_{1-3} were chosen so that in conjunction with the covalent radii R_{cov}^X ²⁵ chemically reasonable CNs are obtained as follows:

$$C_6^{AB}(CN^A, CN^B) = \frac{\sum_i \sum_j L_i^A L_j^B C_{6,\text{rev}}^{AB}(i,j)}{\sum_i \sum_j L_i^A L_j^B}. \quad (7)$$

The Gaussian distance L^A is given by

$$L_i^A = \exp(-k_3(CN^A - CN_i^A)^2) \quad (8)$$

depending on the coordination number CN^A ,

$$CN^A = \sum_{B \neq A} \frac{1}{1 + \exp(-k_1(k_2 \frac{(R_{\text{cov}}^A + R_{\text{cov}}^B)}{R_{AB}} - 1))}. \quad (9)$$

The general gradient expression for the D3 energy with respect to a displacement of nucleus A along a Cartesian coordinate α ($\partial E_{\text{disp}}^{\text{D3}} / \partial r_\alpha^A$) is given by

$$\frac{\partial E_{\text{disp}}^{\text{D3}}}{\partial r_\alpha^A} = - \sum_B \sum_{C>B} \frac{\partial E^{BC}}{r_\alpha^A} \quad (10)$$

$$\begin{aligned} &= - \sum_{B=1}^{A-1} \left(\frac{\partial E^{BA}}{r_\alpha^A} + \sum_{C>B}^{A-1} \frac{\partial E^{BC}}{r_\alpha^A} + \sum_{C>A} \frac{\partial E^{BC}}{r_\alpha^A} \right) \\ &\quad - \sum_{B>A} \left(\frac{\partial E^{AB}}{r_\alpha^A} + \sum_{C>B} \frac{\partial E^{BC}}{r_\alpha^A} \right), \end{aligned} \quad (11)$$

which may be reorganized to

$$\frac{\partial E_{\text{disp}}^{\text{D3}}}{\partial r_\alpha^A} = - \sum_{B \neq A} \frac{\partial E^{AB}}{r_\alpha^A} - \sum_{\substack{B \neq A \\ C>B \\ C \neq A}} \sum_{\substack{C>B \\ C \neq A}} \frac{\partial E^{BC}}{r_\alpha^A}, \quad (12)$$

since $\partial E^{AB} / \partial r_\alpha^A$ equals $\partial E^{BA} / \partial r_\alpha^A$. Thus the displacement does not only effect the dispersion energy of atom pairs including atom A but also all other atom pairs (BC) due to a change in the coordination numbers.

By inserting Equation (7) into Equation (3), the expression can be differentiated by applying the product and chain rules to obtain

$$\begin{aligned} \frac{\partial E_{\text{disp}}^{\text{D3}}}{\partial r_\alpha^A} &= - \sum_{B \neq A} \left(\frac{\partial CN^A}{\partial r_\alpha^A} \frac{\partial C_6^{AB}}{\partial CN^A} f_{\text{remain}}(R_{AB}, A, B) \right. \\ &\quad + \frac{\partial CN^B}{\partial r_\alpha^A} \frac{\partial C_6^{AB}}{\partial CN^B} f_{\text{remain}}(R_{AB}, A, B) \\ &\quad \left. + C_6^{AB} \frac{\partial f_{\text{remain}}(R_{AB}, A, B)}{\partial r_\alpha^A} \right) \\ &\quad - \sum_{\substack{B \neq A \\ C>B \\ C \neq A}} \sum_{\substack{C>B \\ C \neq A}} \frac{\partial CN^B}{\partial r_\alpha^A} \frac{\partial C_6^{BC}}{\partial CN^B} f_{\text{remain}}(R_{BC}, B, C) \\ &\quad - \sum_{\substack{B \neq A \\ C>B \\ C \neq A}} \sum_{\substack{C>B \\ C \neq A}} \frac{\partial CN^C}{\partial r_\alpha^A} \frac{\partial C_6^{CB}}{\partial CN^C} f_{\text{remain}}(R_{BC}, C, B). \end{aligned} \quad (13)$$

By taking out and splitting up the second term in the first summation and exchanging the indices of the last term in Equation (13), this can be rearranged to obtain

$$\begin{aligned} \frac{\partial E_{\text{disp}}^{D3}}{\partial r_\alpha^A} = & - \sum_{B \neq A} \left(\frac{\partial CN^A}{\partial r_\alpha^A} \frac{\partial C_6^{AB}}{\partial CN^A} f_{\text{remain}}(R_{AB}, A, B) + C_6^{AB} \frac{\partial f_{\text{remain}}(R_{AB}, A, B)}{\partial r_\alpha^A} \right) - \sum_{B=1}^{B < A} \frac{\partial CN^B}{\partial r_\alpha^A} \frac{\partial C_6^{BA}}{\partial CN^B} f_{\text{remain}}(R_{AB}, A, B) \\ & - \sum_{\substack{B \neq A \\ C > B \\ \wedge \\ C \neq A}} \frac{\partial CN^B}{\partial r_\alpha^A} \frac{\partial C_6^{BC}}{\partial CN^B} f_{\text{remain}}(R_{BC}, B, C) - \sum_{\substack{B > A \\ \wedge \\ C \neq A}} \frac{\partial CN^B}{\partial r_\alpha^A} \frac{\partial C_6^{AB}}{\partial CN^B} f_{\text{remain}}(R_{AB}, A, B) \\ & - \sum_{\substack{B \neq A \\ C < B \\ \wedge \\ C \neq A}} \frac{\partial CN^B}{\partial r_\alpha^A} \frac{\partial C_6^{CB}}{\partial CN^B} f_{\text{remain}}(R_{BC}, C, B). \end{aligned} \quad (14)$$

This leaves us with derivative terms that are independent with respect to index C. The single sums containing $\partial CN^B / \partial r_\alpha^A$ can be combined with the respective following double sums so that the restrictions in the double sums are lifted. Merging those double sums then gives a double sum, whose inner sum is independent of A, and we arrive at the final form of the D3-gradient,

$$\frac{\partial E_{\text{disp}}^{D3}}{\partial r_\alpha^A} = - \sum_{B \neq A} \left(\frac{\partial CN^A}{\partial r_\alpha^A} \frac{\partial C_6^{AB}}{\partial CN^A} f_{\text{remain}}(R_{AB}, A, B) + C_6^{AB} \frac{\partial f_{\text{remain}}(R_{AB}, A, B)}{\partial r_\alpha^A} + \frac{\partial CN^B}{\partial r_\alpha^A} \sum_{C \neq B} \frac{\partial C_6^{BC}}{\partial CN^B} f_{\text{remain}}(R_{BC}, B, C) \right). \quad (15)$$

Direct application of this formula to compute the gradient for all N nuclei scales by $\mathcal{O}(N^3)$, but the terms $\sum_{C \neq B} (\partial C_6^{BC} / \partial CN^B) f_{\text{remain}}(R_{BC}, B, C)$ and $\partial CN^A / \partial r_\alpha^A$ (see Appendix A) can be computed first and stored in memory to reach $\mathcal{O}(N^2)$ scaling. Explicit expressions for the remaining terms as well as the derivatives for all D3-variants can be found in Appendixes A and B.

To demonstrate the benefit of the optimized algorithm, a comparison between the naive and the new variant has been made. Therefore, the gradient for a $N \times 2 \times 2$ lattice of diamond primitive cells has been computed, where N goes from 1 to 10. To have full control of comparability, a pilot implementation in Python run in the PyPy framework has been used. Working implementations based on compiled programming languages would be considerably faster but would not change the outcome of the comparison.

The most reasonable application scenario, where the gradient computing times are relevant, is the application of D3-variants in sophisticated force-fields or semi-empirical methods for molecular dynamic simulations. For those applications, cutoff radii for the long-range interactions are used. Therefore, a typical cutoff radius of 10 Å is also considered here.

Computing times have been determined on an Intel Core i7-2600 CPU with 3.40 GHz and are averaged over 3 runs.

Results are shown in Fig. 1. As expected, the optimized algorithm is considerably faster and shows a linear speedup with the number of atoms. This even holds for the computations including a cutoff radius although the larger systems exceed this cutoff radius. Note also that the test system grows only in 1 dimension. For real applications, the overall speedup factor should even be larger but would become a constant for very large systems due to the use of a cutoff.

Finally, although the D3(CSO) variant is shown here, the results are representative for any D3-variant.

III. TECHNICAL DETAILS

Structure optimizations have been performed for a variety of different standard functionals from different classes, three generalized gradient approximations (GGAs): BP86,^{26–28} BLYP,^{26,29} PBE,³⁰ one meta-GGA: TPSS,³¹ and three global hybrid GGAs: B3LYP,^{32,33} PBE0,³⁴ PW6B95.³⁵ In contrast to our preceding paper, which covered the thermodynamic benchmark, the B2PLYP³⁶ double-hybrid functional has not been investigated, since there is no analytical gradient implemented in TURBOMOLE³⁷ yet.

All structures were optimized using DFT³⁸ within a locally modified version of TURBOMOLE in which D3(CSO)

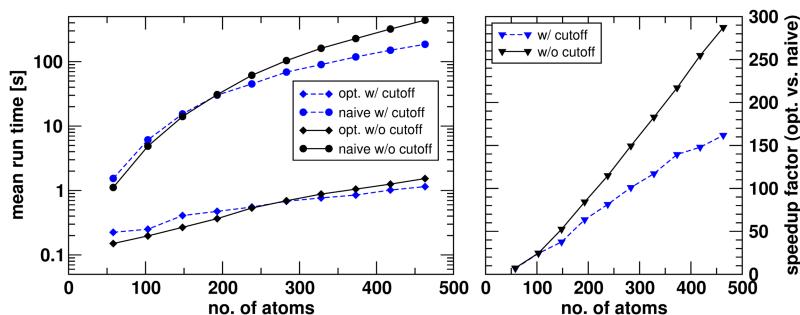


FIG. 1. Comparison of the naive ($\mathcal{O}(N^3)$ scaling) and the optimized gradient algorithm ($\mathcal{O}(N^2)$ scaling) for the D3(CSO) gradient. Shown are mean run times and speedup factors based on computations for a $N \times 2 \times 2$ lattice of diamond primitive cells, where N goes from 1 to 10. An optional use of a cutoff radius of 10 Å is also considered.

and its analytical gradient were implemented. All computations are using natural internal coordinates.³⁹ The def2-TZVP⁴⁰ basis set and the integration grid m4⁴¹ were used for all calculations. In two cases, the m-grid option was switched off, since the optimization ended up oscillating. Both are marked in the [supplementary material](#). Reference structures from the benchmark sets were used as a starting point for all optimizations. Molecular point group symmetry was exploited whenever possible.

IV. RESULTS AND DISCUSSION

Various equilibrium structures have been optimized and evaluated. We start this section with the evaluation of bond lengths in small and medium sized molecules, going over to rotational constant of larger molecules (20–30 atoms), and finally come to the center of mass distances in supramolecular complexes of medium-sized organic molecules or fragments. The choice of benchmark sets ensures, on the one hand, that classical structure benchmarks of organic molecules are represented and that, on the other hand, unusual binding situations, transition metal complexes, and heavy main group elements are also covered. With the exception of the S66 benchmark set⁴² all reference data are obtained from the experiment and taken from the literature (LB12 and HMGB11 from Ref. 13, TMC32 from Ref. 43, and ROT34 from Refs. 11 and 12).

For the analysis of individual data points for a benchmark set, BLYP, TPSS, and PW6B95 have been chosen as representative functionals for their respective classes. All single data points are given in the [supplementary material](#). All distance errors given in Secs. IV A–IV C are the difference between the calculated value and the reference value, thus a positive error value corresponds to a very large distance.

A. Bond lengths

Bond lengths are evaluated by means of mainly inorganic molecules in this section. These include transition metal complexes (TMC32⁴³), simple molecules containing heavy main group elements (HMGB11¹³), and compounds with exceptionally long bonds (LB12¹³). Structures involving mainly organic molecules made of light main group elements are covered in Secs. IV B and IV C. The expected error of the experimental reference values is 2 pm for all test sets in this section. Statistical results for these evaluations are given in Table I.

Our first benchmark set, the LB12,¹³ includes 12 molecules with exceptionally long bonds or interatomic interactions. They are shown in Fig. 2. The labeling of the molecules was adopted from the reference. Those molecules are known to be difficult cases for standard DFAs, which shows the very high mean absolute deviation (MAD) compared to the other sets in this section (see Table I). D3(BJ) performs significantly better for all functionals in the overall statistics showing a lower MAD as well as a reduced error range.

The detailed analysis (Fig. 3) shows that the largest error as well as the largest deviation between both corrections can be found for molecule no. 5, which is the circular S_8^{2+} -cation. It is known to be highly problematic for standard DFAs.^{10,44} An enormous error of 69 pm is found for BLYP-D3(CSO) and still a very high error of 34 pm for BLYP-D3(BJ). While in this case, a switch to more sophisticated functionals like PBE0 and PW6B95 reduces the error significantly, no method got close to an acceptable error for molecule no. 3 (DTFS) and no. 4 (HAPPOD). However, whenever reasonable structures are obtained, both corrections only differ insignificantly.

The next benchmark set is the TMC32, which is composed of 50 metal-ligand bond lengths from 32 3d-transition

TABLE I. Statistical evaluation of the LB12, TMC32, and HMGB11 benchmark sets with several DFAs and either D3(BJ) or D3(CSO) correction. Given are the mean deviation (MD), mean absolute deviation (MAD), standard deviation (SD), and the maximal absolute deviation (MAX). All errors are in pm.

BP		BLYP		PBE		TPSS		B3LYP		PBE0		PW6B95		
BJ	CSO	BJ	CSO	BJ	CSO	BJ	CSO	BJ	CSO	BJ	CSO	BJ	CSO	
LB12														
MD	2.0	3.0	8.5	11.2	2.7	6.3	0.7	0.2	5.3	9.8	0.2	-0.5	-1.4	-1.8
MAD	6.1	7.8	9.4	11.2	6.3	7.8	4.8	6.1	5.9	9.8	3.3	3.6	4.5	5.1
SD	9.0	14.1	10.7	19.2	9.0	10.5	8.0	10.4	6.5	9.6	5.2	5.6	7.1	7.5
MAX	20.6	42.8	34.4	69.2	21.1	31.8	17.9	25.2	20.4	30.1	13.0	12.8	16.9	18.8
TMC32														
MD	-0.6	-0.6	1.4	1.4	-0.6	-0.4	-0.6	-0.7	-0.3	0.1	-2.2	-2.2	-1.7	-1.8
MAD	1.5	1.5	2.0	2.0	1.6	1.5	1.3	1.3	1.6	1.7	2.4	2.4	2.0	2.1
SD	1.9	1.9	2.1	2.1	1.9	1.9	1.7	1.6	1.9	2.1	1.7	1.7	1.7	1.6
MAX	5.7	5.8	6.3	6.2	5.8	5.6	5.2	5.2	4.5	4.6	6.5	6.6	5.2	5.3
HMGB11														
MD	2.5	1.3	5.4	4.1	2.6	2.9	1.9	0.8	2.6	3.2	0.1	-0.6	0.8	0.3
MAD	2.5	1.4	5.4	4.1	2.6	2.9	1.9	1.1	2.6	3.2	1.1	1.3	1.6	1.3
SD	1.3	1.3	2.1	1.4	1.7	1.8	1.3	1.2	1.2	1.5	1.4	1.5	2.2	1.9
MAX	4.5	3.3	8.4	6.6	5.1	6.0	3.7	2.7	4.5	6.2	2.7	2.7	6.2	4.5

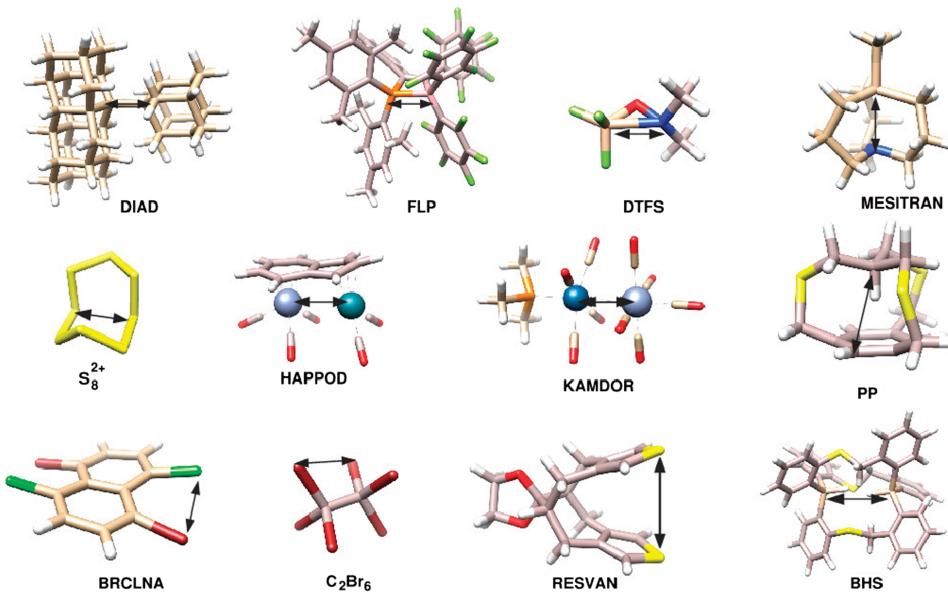


FIG. 2. Molecule structures of the LB12 set. Arrow indicates the considered interatomic distances. Reprinted from the work of J. Chem. Phys. **143**, 054107 (2015). Copyright 2015 AIP Publishing, LLC.

metal complexes. Mean errors as well as the error range are much lower than those for the preceding set. All functionals lie within or close to the estimated error range of the reference. Hardly any differences between both corrections can be found for any method, neither in the overall statistics nor for the individual structures, thus we abstain from a detailed discussion here.

The last part of this section deals with the bond lengths of heavy main group atoms. While the bond lengths are usually overestimated, D3(CSO) gives significantly shorter bond lengths in most cases but for B3LYP and PBE0, and thus improves the description. The deviation between both corrections increases with the nuclear charge and thus the bond distance between the central atoms (Fig. 4). D3(CSO) performs significantly better for the heaviest atoms in almost

all shown cases, though the small size of the test does not allow for a final conclusion, whether this result is transferable.

Overall, bond lengths obtained with both corrections are very similar. The largest deviations occur in rather exotic molecules, while there is hardly any difference for standard cases. Both dispersion corrections show very good results for standard bonding situations.

B. Rotational constants

Our next benchmark set deals with intramolecular effects. The statistical evaluation can be found in Table II. The set consists of 34 rotational constants B_e from 12 medium-sized organic molecules. In order to obtain accurate rotational constants, a well-balanced description of bond distances as well as non-bonding interactions is needed. Positive values of the

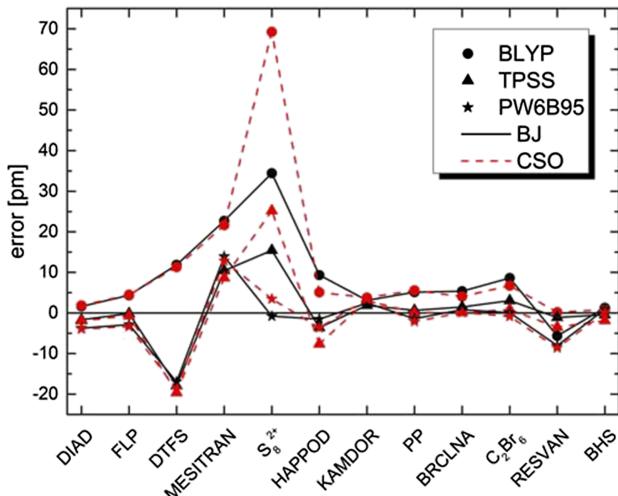


FIG. 3. Error of specific interatomic distances of the LB12 test set to the respective reference values.¹³ Molecules are ordered corresponding to the experimental reference values, beginning with the lowest. Labeling was adopted from Ref. 13. Error is in pm.

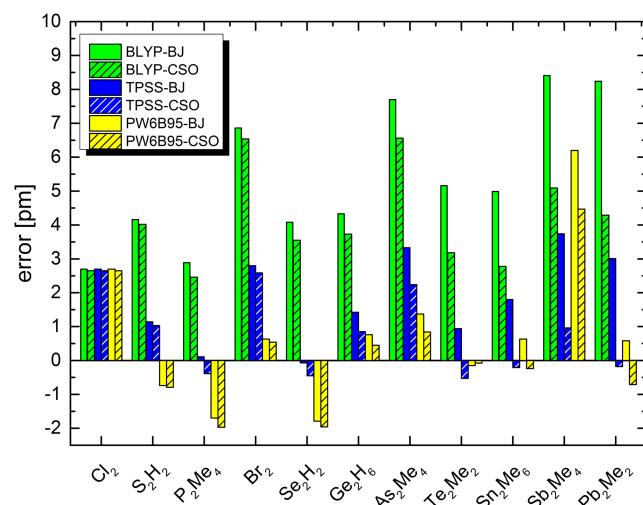


FIG. 4. Error of specific bond lengths between both heavy atoms of the HMGB11 test set to the respective reference values.¹³ Molecules are ordered corresponding to the experimental reference values, beginning with the lowest. Error is in pm.

TABLE II. Statistical relative error values of all functionals for computing the rotational constants in ROT34. Errors given are the mean deviation (MD), mean absolute deviation (MAD), standard deviation (SD), and maximum absolute deviation (MAX). All values are in %.

	BP		BLYP		PBE		TPSS		B3LYP		PBE0		PW6B95	
	BJ	CSO	BJ	CSO	BJ	CSO	BJ	CSO	BJ	CSO	BJ	CSO	BJ	CSO
MD	-1.3	-1.2	-2.3	-2.4	-1.3	-1.4	-1.2	-1.1	-0.8	-1.0	0.2	0.2	0.6	0.7
MAD	1.3	1.2	2.3	2.4	1.3	1.4	1.2	1.1	0.8	1.0	0.3	0.3	0.6	0.7
SD	0.6	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.4	0.4	0.3	0.2	0.4	0.4
MAX	3.5	3.2	4.1	4.3	2.5	2.7	3.3	3.2	2.0	2.1	0.6	0.7	1.7	2.0

TABLE III. Statistical error values of chosen functionals with D3(BJ) and D3(CSO) dispersion correction for the center of mass distances of the S66 benchmark set. Reference structures and numbering are taken from Ref. 42. Given are the mean deviation (MD), mean absolute deviation (MAD), standard deviation (SD), and maximal absolute deviation (MAX). All values are in pm.

	BP ^a		BLYP ^b		PBE ^c		TPSS ^c		B3LYP ^c		PBE0 ^c		PW6B95 ^d	
	BJ	CSO	BJ	CSO	BJ	CSO	BJ	CSO	BJ	CSO	BJ	CSO	BJ	CSO
MD	-3.3	-7.6	0.9	-1.2	2.4	4.7	5.4	3.9	-0.4	-0.3	1.1	0.4	1.9	-1.2
MAD	5.8	8.1	3.6	4.6	6.4	8.0	8.2	6.4	3.0	3.4	5.4	4.8	3.9	4.9
SD	8.1	5.2	6.2	6.8	9.5	10.2	10.4	7.2	4.6	4.7	7.2	5.9	4.9	7.0
MAX	47.9	20.8	32.3	26.0	54.3	49.6	55.7	25.0	19.3	15.0	36.7	15.7	13.8	27.5

^aComplexes 52–9,31 removed, see text.

^bComplexes 29,31 removed, see text.

^cComplexes 29,31,57 removed, see text.

^dComplex 31 removed, see text.

MD correspond to very bulky molecules. Thus molecular size is underestimated for all non-hybrid functionals and B3LYP, while they are slightly overestimated for PBE0 and PW6B95 for both corrections. Excellent results are obtained with PBE0 for both corrections with an MAD of only 0.3%. Overall the hybrid functionals give slightly better structures. The corrections themselves do not differ significantly for any functional. A maximum deviation of only 0.2% in the MAD can be found for B3LYP.

C. Supramolecular complexes

We conclude our evaluation with the S66 benchmark set⁴² of non-covalent interactions. It is composed of 66 van der Waals complexes containing two non-covalently bound molecules each. The optimized structures are evaluated by means of their center of mass distances. Because of the weak interactions, deviations of up to 10 pm are acceptable. Note that the center of mass distance is very sensitive to changes in the orientation of the molecules.

For three complexes, structures were obtained which differed qualitatively from the reference structures for at least some combination of DFA and D3 correction. Whenever this was the case, the values of both corrections were removed from the statistical evaluation (Table III) in order not to skew the overall result. These failures were distributed equally between both corrections. Detailed information can be found in the supplementary material. Especially interactions involving π -systems were particularly critical. It has been shown that MP2, which was used as the final optimization method for the S66 benchmark set, overbinds this kind of interactions⁴⁵ and might therefore be problematic in these cases.

Overall D3(BJ) performs slightly better than D3(CSO) for most functionals, but the deviations are again very low considering the weak interactions. The maximum difference in the MAD is 2.3 pm for BP. However, any combination of DFA and D3 correction gives results within the desired error range.

V. CONCLUSION

We evaluated our recently developed D3(CSO) dispersion correction for its performance for structure optimizations in comparison to its predecessor D3(BJ). An analytical gradient was therefore implemented.

Overall, there are only minor differences between both corrections for geometry optimizations for common test systems, like organic molecules or 3d-transition metals. Significant deviations only showed up in the LB12 set, where D3(CSO) performed significantly worse than D3(BJ) in the overall statistics, almost solely because of the S_8^{2+} -cation. Contrary, D3(CSO) improved the bond distance between heavy main group elements in the HMGB11 set significantly in most cases and even more so for the higher main group atoms. These findings are only supported by a few molecules though and more test systems are needed for a final conclusion. Finally, minor deviations could also be found in the S66 set of non-covalent interactions with slight advantages for D3(BJ).

Furthermore, an optimized algorithm for the computation of the D3-gradient was presented here, which reduces the formal scaling from $\mathcal{O}(N^3)$ to $\mathcal{O}(N^2)$. This is not restricted to our CSO-variant but applies to all D3-corrections. While the considerable speedup is insignificant for DFT calculations, where the computation of the DFT part of the gradient dominates,

semi-empirical methods or force fields using the D3-gradient may noticeably benefit from it.

SUPPLEMENTARY MATERIAL

See [supplementary material](#) for all individual data points for every benchmark set.

APPENDIX A: REMAINING TERMS

The remaining terms needed for an implementation of the D3-gradient are given as follows:

$$\frac{\partial CN^A}{\partial r_\alpha^A} = \sum_{A \neq B} \frac{\partial R^{AB}}{\partial r_\alpha^A} \frac{\partial CN^A}{\partial R^{AB}}, \quad (\text{A1})$$

$$\frac{\partial CN^B}{\partial r_\alpha^A} = \frac{\partial R^{AB}}{\partial r_\alpha^A} \frac{\partial CN^B}{\partial R^{AB}} \text{ for } B \neq A, \quad (\text{A2})$$

$$\frac{\partial CN^A}{\partial R^{AB}} = \frac{-k_1(k_2 \frac{R_{cov}^A + R_{cov}^B}{(R^{AB})^2}) \exp\left(-k_1 \left(k_2 \frac{R_{cov}^A + R_{cov}^B}{R^{AB}} - 1\right)\right)}{\left(1 + \exp\left(-k_1 \left(k_2 \frac{R_{cov}^A + R_{cov}^B}{R^{AB}} - 1\right)\right)\right)^2}, \quad (\text{A3})$$

$$\frac{\partial R^{AB}}{\partial r_\alpha^A} = -\frac{\alpha^B - \alpha^A}{R^{AB}}, \quad (\text{A4})$$

$$\frac{\partial C_6^{AB}}{\partial CN^A} = \sum_i \frac{\partial L_i^A}{\partial CN^A} \frac{\partial C_6^{AB}}{\partial L_i^A}, \quad (\text{A5})$$

$$\frac{\partial L_i^A}{\partial CN^A} = -2k_3(CN^A - CN_i^A) \exp(-k_3(CN^A - CN_i^A)^2), \quad (\text{A6})$$

$$\frac{\partial C_6^{AB}}{\partial L_i^A} = \frac{W^{AB} \sum_j L_j^B C_{6rev}^{AB}(i,j) - Z^{AB} \sum_j L_j^B}{(W^{AB})^2}, \quad (\text{A7})$$

$$W^{AB} = \sum_i \sum_j L_i^A L_j^B. \quad (\text{A8})$$

APPENDIX B: DERIVATIVES OF D3-VARIANTS

The partial derivative of the remaining function $f_{\text{remain}}(R_{AB}, A, B)$, which constitutes the difference between the D3 variants, occurring in Equation (15), is given by

$$\frac{\partial f_{\text{remain}}(R_{AB}, A, B)}{\partial r^A} = \frac{\partial R^{AB}}{\partial r^A} \frac{\partial f_{\text{remain}}(R_{AB}, A, B)}{\partial R^{AB}}. \quad (\text{B1})$$

In the following, the terms for the three corrections are given with their respective derivatives.

1. D3(Zero)

In the original D3 correction,⁹ the dispersion correction vanishes for $R_{AB} \rightarrow 0$, which is also known as zero-damping. Two DFA-specific empirical parameters $s_{r,6}$ and s_8 are needed. The C_8^{AB} coefficient is recursively calculated from the C_6^{AB} coefficient by multiplication with a distance independent factor $k_{6,8}^{AB}$. The remaining term has the form

$$f_{\text{remain}}^{\text{Zero}}(R_{AB}, A, B) = \left(f_{\text{damp},6}(R_{AB}) \frac{s_6}{(R_{AB})^6} + f_{\text{damp},8}(R_{AB}) \frac{s_8 k_{6,8}^{AB}}{(R_{AB})^8} \right). \quad (\text{B2})$$

The partial derivative with respect to R_{AB} is given by

$$\frac{\partial f_{\text{remain}}^{\text{Zero}}}{\partial R_{AB}} = f_{\text{damp},6} \left(\frac{-6s_6(R_{AB})^5}{((R_{AB})^6)^2} \right) + f_{\text{damp},8} \left(\frac{-8s_8 k_{6,8}^{AB} (R_{AB})^7}{((R_{AB})^8)^2} \right) + \frac{\partial f_{\text{damp},6}}{\partial R_{AB}} \frac{s_6}{(R_{AB})^6} + \frac{\partial f_{\text{damp},8}}{\partial R_{AB}} \frac{s_8 k_{6,8}^{AB}}{(R_{AB})^8}. \quad (\text{B3})$$

The n -order damping function $f_{\text{damp},n}$ contains the steepness parameter α_n , which is $\alpha_6 = 14$ for $n=6$ and $\alpha_{n+2} = \alpha_n + 2$ for higher even values of n . The partial derivative is

$$\frac{\partial f_{\text{damp},6}}{\partial R_{AB}} = \frac{6\alpha_6(s_{r,6}R_0^{AB})^{\alpha_6} (R^{AB})^{-\alpha_6-1}}{\left(1 + 6\left(\frac{R^{AB}}{s_{r,6}R_0^{AB}}\right)^{-\alpha_6}\right)^2} \quad (\text{B4})$$

$$\frac{\partial f_{\text{damp},8}}{\partial R_{AB}} = \frac{6\alpha_8(R_0^{AB})^{\alpha_8} (R^{AB})^{-\alpha_8-1}}{\left(1 + 6\left(\frac{R^{AB}}{R_0^{AB}}\right)^{-\alpha_8}\right)^2}. \quad (\text{B5})$$

2. D3(BJ)

In the D3(BJ) correction,¹⁰ the dispersion corrections become constant for $R_{AB} \rightarrow 0$, which is known as Becke-Johnson-damping¹⁹ (BJ-damping). Three DFA-specific parameters a_1, a_2 , and s_8 are needed. The parameter s_6 is unity for most functionals (for exceptions see Ref. 10). The C_8^{AB} coefficient is calculated as in D3(zero). The remaining term has the form

$$f_{\text{remain}}^{\text{BJ}}(R_{AB}, A, B) = \frac{s_6}{(R^{AB})^6 + (a_1 R_0^{AB} + a_2)^6} + \frac{s_8 k_{6,8}^{AB}}{(R^{AB})^8 + (a_1 R_0^{AB} + a_2)^8} \quad (\text{B6})$$

and the respective partial derivate with respect to R^{AB} is

$$\frac{\partial f_{\text{remain}}^{\text{BJ}}}{\partial R^{AB}} = \left(\frac{-6s_6(R^{AB})^5}{((R^{AB})^6 + (a_1 R_0^{AB} + a_2)^6)^2} + \frac{-8s_8 k_{6,8}^{AB} (R^{AB})^7}{((R^{AB})^8 + (a_1 R_0^{AB} + a_2)^8)^2} \right). \quad (\text{B7})$$

3. D3(CSO)

The D3(CSO) scheme¹⁴ also uses BJ-damping. One DFA-specific parameter a_1 is needed. The parameter s_6 is unity for all functionals without a non-local correlation part. In this case, it is lowered by the percentage of non-local correlation. The remaining term has the form

$$f_{\text{remain}}^{\text{CSO}}(R_{AB}, A, B) = \left(s_6 + \frac{a_1}{1 + \exp(R_{AB} - 2.5 R_0^{AB})} \right) \times \frac{1}{R_{AB}^6 + (2.5^2)^6} \quad (\text{B8})$$

and the corresponding partial derivative with respect to R^{AB} is

$$\frac{\partial f_{remain}^{CSO}(R_{AB}, A, B)}{\partial R^{AB}} = \frac{-a_1 \exp(R_{AB} - 2.5 R_0^{AB})}{\left(1 + \exp(R_{AB} - 2.5 R_0^{AB})\right)^2} \frac{1}{R_{AB}^6 + (2.5^2)^6} + \left(s_6 + \frac{a_1}{1 + \exp(R_{AB} - 2.5 R_0^{AB})}\right) \frac{-6s_6 R^{AB^5}}{\left(R_{AB}^6 + (2.5^2)^6\right)^2}. \quad (\text{B9})$$

- ¹A. D. Becke, *J. Chem. Phys.* **140**, 18A301 (2014).
²S. Kristyán and P. Pulay, *Chem. Phys. Lett.* **229**, 175 (1994).
³S. Grimme and P. R. Schreiner, *Angew. Chem., Int. Ed.* **50**, 12639 (2011).
⁴M. Kolář, T. Kubář, and P. Hobza, *J. Phys. Chem. B* **115**, 8038 (2011).
⁵L. Goerigk and R. Reimers, *J. Chem. Theory Comput.* **9**, 3240 (2013).
⁶L. Goerigk, C. A. Collyer, and R. Reimers, *J. Phys. Chem. B* **118**, 14612 (2014).
⁷S. Grimme, A. Hansen, J. G. Brandenburg, and C. Bannwarth, *Chem. Rev.* **116**, 5105 (2016).
⁸J. Klimeš and A. Michaelides, *J. Chem. Phys.* **137**, 120901 (2012).
⁹S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, *J. Chem. Phys.* **132**, 154104 (2010).
¹⁰S. Grimme, S. Ehrlich, and L. Goerigk, *J. Comput. Chem.* **32**, 1456 (2011).
¹¹S. Grimme and M. Steinmetz, *Phys. Chem. Chem. Phys.* **15**, 16031 (2013).
¹²T. Risthaus, M. Steinmetz, and S. Grimme, *J. Comput. Chem.* **35**, 1509 (2014).
¹³S. Grimme, J. G. Brandenburg, C. Bannwarth, and A. Hansen, *J. Chem. Phys.* **143**, 054107 (2015).
¹⁴H. Schröder, A. Creon, and T. Schwabe, *J. Chem. Theory Comput.* **11**, 3163 (2015).
¹⁵U. Vucak, H. Ji, Y. Singh, and Y. Jung, *J. Chem. Phys.* **145**, 174104 (2016).
¹⁶E. B. Guidez, P. Xu, and M. S. Gordon, *J. Phys. Chem. A* **120**, 639 (2016).
¹⁷T. Schwabe, *J. Chem. Phys.* **145**, 154105 (2016).
¹⁸A. Koide, *J. Phys. B: At. Mol. Phys.* **9**, 3173 (1976).
¹⁹E. R. Johnson and A. D. Becke, *J. Chem. Phys.* **124**, 174104 (2006).
²⁰P. Jurečka, J. Černý, P. Hobza, and D. R. Salahub, *J. Comput. Chem.* **28**, 555 (2007).
²¹J. Tao, J. P. Perdew, and A. Ruzsinszky, *Proc. Natl. Acad. Sci. U. S. A.* **109**, 18 (2012).
²²J. Tao, Y. Fang, P. Hao, G. E. Scuseria, A. Ruzsinszky, and J. P. Perdew, *J. Chem. Phys.* **142**, 024312 (2015).
²³H. Schröder and T. Schwabe, *J. Comput. Chem.* **37**, 2052 (2016).
²⁴M. Stöhr, G. S. Michelitsch, J. C. Tully, K. Reuter, and R. J. Maurer, *J. Chem. Phys.* **144**, 151101 (2016).
²⁵P. Pyykkö and M. Atsumi, *Chem. - Eur. J.* **15**, 186 (2009).
²⁶A. D. Becke, *Phys. Rev. A* **38**, 3098 (1988).
²⁷J. P. Perdew, *Phys. Rev. B* **33**, 8822 (1986).
²⁸J. P. Perdew, *Phys. Rev. B* **34**, 7406 (1986).
²⁹C. Lee, W. Yang, and R. G. Parr, *Phys. Rev. B* **37**, 785 (1988).
³⁰J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996).
³¹J. Tao, J. P. Perdew, V. N. Staroverov, and G. E. Scuseria, *Phys. Rev. Lett.* **91**, 146401 (2003).
³²A. D. Becke, *J. Chem. Phys.* **98**, 5648 (1993).
³³P. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch, *J. Phys. Chem.* **98**, 11623 (1994).
³⁴C. Adamo and V. Barone, *J. Chem. Phys.* **110**, 6158 (1999).
³⁵Y. Zhao and D. G. Truhlar, *J. Phys. Chem. A* **109**, 5656 (2005).
³⁶S. Grimme and *J. Chem. Phys.* **124**, 034108 (2006).
³⁷TURBOMOLE, V7.1, 2016, developed by University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, available from <http://www.turbomole.com>.
³⁸M. Häser and R. Ahlrichs, *J. Comput. Chem.* **10**, 104 (1989).
³⁹M. von Arnim and R. Ahlrichs, *J. Chem. Phys.* **111**, 9183 (1999).
⁴⁰F. Weigend and R. Ahlrichs, *Phys. Chem. Chem. Phys.* **7**, 3297 (2005).
⁴¹O. Treutler and R. Ahlrichs, *J. Chem. Phys.* **102**, 346 (1995).
⁴²J. Rezáč, K. E. Riley, and P. Hobza, *J. Chem. Theory Comput.* **7**, 2427 (2011).
⁴³M. Bühl and H. Kabrede, *J. Chem. Theory Comput.* **2**, 1282 (2006).
⁴⁴T. S. Cameron, R. J. Deeth, I. Dionne, H. Du, H. Donald, B. Jenkins, I. Krossing, J. Passmore, and H. K. Roobottom, *Inorg. Chem.* **39**, 5614 (2000).
⁴⁵T. Schwabe and S. Grimme, *J. Phys. Chem. A* **113**, 3005 (2009).