

Accurate non-covalent interactions via dispersion-corrected second-order Møller-Plesset perturbation theory.

Jan Řezáč,^{*,†} Chandler Greenwell,[‡] and Gregory J. O. Beran^{*,‡}

[†]*Institute of Organic Chemistry and Biochemistry, Czech Academy of Sciences, 166 10*

Prague, Czech Republic

[‡]*Department of Chemistry, University of California, Riverside, California 92521 USA*

E-mail: rezac@uochb.cas.cz; gregory.beran@ucr.edu

August 1, 2018

Abstract

Non-covalent interactions govern many important areas of chemistry, ranging from biomolecules to molecular crystals. Here, an accurate and computationally inexpensive dispersion-corrected second-order Møller-Plesset perturbation theory model (MP2D) is presented. MP2D recasts the highly successful dispersion-corrected MP2C model in a framework based on Grimme’s D3 dispersion correction, combining Grimme’s D3 dispersion coefficients with new analogous uncoupled Hartree-Fock ones and five global empirical parameters. MP2D is faster than MP2C, and unlike MP2C, it is suitable for geometry optimizations and can describe both intra- and intermolecular non-covalent interactions with high accuracy. MP2D approaches the accuracy of higher-level *ab initio* wavefunction techniques and out-performs a widely-used hybrid dispersion-corrected density functional on a range of intermolecular, intramolecular, and thermochemical benchmarks.

1 Introduction

Non-covalent interactions govern protein folding, chemistry in solution, molecular crystal polymorphism, and many other important phenomena. Simulating such systems requires theoretical models capable of accurately reproducing the often delicate balances among the different types of non-covalent interactions both within molecules and between them. Large-basis coupled cluster methods can achieve this accuracy for small systems,¹ but they are computationally prohibitive for larger ones. Dispersion-corrected density functional theory (DFT) models provide a much more affordable option,² though DFT cannot always provide the requisite accuracy due to self-interaction error and other inherent limitations in the functionals.³ Here, we report a new, computationally practical dispersion-corrected second-order Møller-Plesset perturbation theory (MP2) model which provides high-quality energetics and structures in systems where non-covalent interactions are important, filling an important gap between DFT and higher-level techniques.

With formal computational cost scaling with the fifth power with system size (though this scaling can be reduced via Laplace transform, local correlation models, etc⁴), MP2 provides a valuable and computationally affordable alternative to DFT for organic systems, but it has well-known problems describing van der Waals dispersion interactions. It overestimates the interaction energy in the π -stacked benzene dimer by a factor of two, for example.⁵ From the perspective of intermolecular perturbation theory, this deficiency in MP2 stems from its uncoupled Hartree-Fock (UCHF) treatment of intermolecular dispersion^{6,7} which approximates the excited states and excitation energies that contribute to the dispersion energy using unrelaxed ground-state Hartree-Fock orbitals.

Various models empirically scale the same-spin and opposite-spin correlation components in MP2 to improve its performance,⁸⁻¹³ though the optimal parameters often vary with the nature of the chemistry being modeled. The very successful non-empirical MP2C method^{7,14} replaces the problematic UCHF dispersion with an improved coupled Kohn-Sham (CKS)

treatment of dispersion,

$$E^{MP2C} = E^{MP2} - E_{disp}^{UCHF} + E_{disp}^{CKS} \quad (1)$$

effectively using time-dependent density functional theory to obtain an improved description of the excited states. The excellent performance of MP2C for intermolecular interactions earned it “the bronze-standard of quantum chemistry” moniker.¹⁵ Unfortunately, MP2C has two major limitations. First, the dispersion correction is derived from intermolecular perturbation theory and is not defined for intramolecular interactions. Intramolecular dispersion can be crucial in larger molecules. Second, MP2C is not currently used for structure optimization due to the complexity of its analytical nuclear gradients.

Here, we combine the ideas of MP2C with Grimme’s DFT-D3 dispersion correction¹⁶ to develop a new dispersion-corrected MP2D model. Recasting MP2C in terms of atom-centered two-body dispersion coefficients offers clear advantages. Atomic dispersion coefficients can be applied to both intra- and intermolecular atom-atom interactions. Furthermore, both the energy and analytical gradients of the dispersion correction can be computed with trivial computational cost. On the other hand, it introduces some empiricism to the model in the form of five global parameters. MP2D is also similar to the MP2+ Δ vdW model,¹⁷ but it improves upon that model in several important ways. It includes both the C_6 and C_8 terms, instead of only C_6 like MP2+ Δ vdW. More significantly, MP2D solves the problem of how to determine the atomic C_6 dispersion coefficients for different chemical environments by adopting the D3 dispersion correction approach.¹⁶

The following sections present the MP2D model, including how the dispersion coefficients were obtained, modifications to the short-range damping necessary to treat both covalent- and non-covalent chemistry, several minor changes to the D3 procedure, and the strategy used to ensure physically appropriate parameters were obtained. We then demonstrate that MP2D performs very well across thousands of benchmark energies, including intermolecular

interactions, conformation energies, and thermochemistry. We examine in detail the performance of MP2D on the challenging anthracene photodimerization, in which inter- and intramolecular interactions compete strongly. Finally, we study several examples of geometry optimization where dispersion effects play a major role. Throughout these tests, MP2D significantly improves MP2 in cases where van der Waals dispersion is important, and it does so with negligible additional computational cost. At the same time, the MP2D dispersion correction has little impact on MP2 in cases where dispersion does not contribute significantly.

2 Theory

2.1 MP2D overview

MP2D corrects MP2 by subtracting out the pairwise interatomic UCHF dispersion energy and replacing it with the equivalent contribution calculated at the CKS level of theory,

$$E^{MP2D} = E^{MP2} - \tilde{E}_{disp}^{UCHF} + \tilde{E}_{disp}^{CKS} \quad (2)$$

where

$$\tilde{E}_{disp} = s_6 \sum_{a,b} f_6(R_{AB}) \frac{C_{6,ab}}{R_{AB}^6} + s_8 \sum_{a,b} f_8(R_{AB}) \frac{C_{8,ab}}{R_{AB}^8} \quad (3)$$

In these expressions, C_6 and C_8 are the interatomic two-body dispersion coefficients calculated at either the UCHF or CKS levels of theory, R_{AB} is the distance between atoms A and B , f_n are short-range damping functions, and s_n are empirical scaling factors. One might further augment MP2D with a 3-body dispersion term,¹⁸ since those contributions are missing in MP2¹⁹ and can become significant in large systems,^{20,21} but that is not done here.

MP2D adopts Grimme’s D3 model¹⁶ to compute the UCHF and CKS dispersion contributions. In fact, MP2D uses Grimme’s existing D3 C_6 coefficients for the CKS dispersion

energy. New UCHF dispersion coefficients are computed here. Several other minor modifications are made to the D3 approach with regard to the damping at short non-covalent and covalent distances and the evaluation of the continuous coordination approach, as described below.

2.2 Review of the D3 approach

It is worthwhile to review Grimme’s D3 approach briefly before discussing the MP2D-specific changes. D3 computes frequency-dependent dipole-dipole polarizabilities $\alpha(i\omega)$ for a series of different hydrides with different coordination numbers (CNs). For carbon, for instance, it computes them for C ($CN = 0$), CH ($CN \approx 1$), C₂H₂ ($CN \approx 2$), C₂H₄ ($CN \approx 3$), and C₂H₆ ($CN \approx 4$). The C_6 dispersion coefficients for all possible pairwise combinations of atoms and coordination numbers are calculated via Casimir-Polder integration with these polarizabilities after subtracting out the approximate hydrogen contribution,

$$C_{6,ab}(CN_i^a, CN_j^b) = \frac{3}{\pi} \int_0^\infty d\omega \frac{1}{m} \left[\alpha^{A_m H_n}(i\omega) - \frac{n}{2} \alpha^{H_2}(i\omega) \right] \frac{1}{k} \left[\alpha^{B_k H_l}(i\omega) - \frac{l}{2} \alpha^{H_2}(i\omega) \right] \quad (4)$$

In this expression, $\alpha^{A_m H_n}(i\omega)$ and $\alpha^{B_k H_l}(i\omega)$ are the frequency dependent polarizabilities for the reference hydrides, and $\alpha^{H_2}(i\omega)$ is the corresponding value for H₂.

Key to the success of the D3 model is how it interpolates the pre-tabulated C_6 coefficients to adapt them to the current chemical environment via these coordination numbers. It computes continuous coordination numbers for each atom a in a given system as,

$$CN^a = \sum_{b \neq a}^N \frac{1}{1 + e^{-16 \left(\frac{4}{3} \frac{R_{ab}^{cov}}{R_{ab}} - 1 \right)}} \quad (5)$$

where R_{ab}^{cov} are sums of pre-tabulated, modified covalent radii for each element pair. Refer to the original D3 paper for details on the set of the radii used.¹⁶ The C_6 dispersion coefficients for atom a interacting with atom b in their current coordination environments

is then computed via a weighted average of the C_6 values from the tabulated coordination environments,

$$C_{6,ab} = \frac{\sum_i \sum_j C_{6,ab}^{ref}(CN_i^a, CN_j^b) L_{ij}}{\sum_i \sum_j L_{ij}} \quad \text{where} \quad L_{ij} = e^{-4[(CN^a - CN_i^a)^2 + (CN^b - CN_j^b)^2]} \quad (6)$$

Indices i and j sum over the all reference hydrides for the given element. Once the C_6 coefficients are obtained for an atom-pair, the C_8 coefficients are estimated according to,

$$C_{8,ab} = 3C_{6,ab} \sqrt{Q_a Q_b} \quad (7)$$

and

$$Q_a = \frac{1}{2} \sqrt{Z_a} \frac{\langle r^4 \rangle_a}{\langle r^2 \rangle_a} \quad (8)$$

where Z_a is the nuclear charge, and $\langle r^4 \rangle_a$ and $\langle r^2 \rangle_a$ are pretabulated multipole expectation values for the element. See Grimme's original work for more details.¹⁶

In MP2D, Grimme's reference hydride C_6 coefficients (as implemented in Cuby4²²) are used for the CKS portion of the model. Other D3 parameters, such as the multipole expectation values $\langle r^n \rangle_a$, covalent radii R_{ab}^{cov} , and cutoff radii $R_{0,ab}$ are also employed in MP2D without modification. However, several modifications and new ingredients were added in the development of MP2D, as described in the following sections.

2.3 UCHF dispersion coefficients

Before adding the CKS dispersion to MP2, one must subtract out the UCHF dispersion that is already present. This requires computing UCHF dispersion coefficients that are analogous to the existing D3 CKS ones. The general expression for the frequency dependent dipole-dipole polarizability tensor $\alpha_{\lambda\sigma}(i\omega)$ from intermolecular perturbation theory is,

$$\alpha_{\lambda\sigma}(i\omega) = \sum_{M \neq 0} \frac{\omega_M [\langle 0 | \hat{\mu}_\lambda | M \rangle \langle M | \hat{\mu}_\sigma | 0 \rangle + \langle 0 | \hat{\mu}_\sigma | M \rangle \langle M | \hat{\mu}_\lambda | 0 \rangle]}{\hbar(\omega_M^2 + \omega^2)} \quad (9)$$

where 0 and M refer to ground and excited states, ω_M is the excitation energy, and λ and σ refer to different Cartesian components of the dipole operator $\hat{\mu}$. The resulting polarizability tensor $\alpha_{\lambda\sigma}(i\omega)$ is a symmetric 3×3 matrix with unique xx , xy , xz , yy , yz , and zz elements.

At the UCHF level of theory, the excited state wavefunctions involve vertical excitation of an electron from occupied orbital i to virtual orbital a with no orbital relaxation. In that case, the matrix elements simplify to matrix elements of the dipole operator involving occupied orbital i and virtual orbital a , and the excitation energy ω_M reduces to the energy difference between orbitals i and a , $\omega_M = \epsilon_i - \epsilon_a = \epsilon_{ia}$. Employing these simplifications and recognizing that the molecular orbitals are real, Eq 9 becomes,

$$\alpha_{\lambda\sigma}(i\omega) = 2 \sum_{ia} \frac{\epsilon_{ia} \langle i | \hat{\mu}_\lambda | a \rangle \langle a | \hat{\mu}_\sigma | i \rangle}{\hbar(\epsilon_{ia}^2 + \omega^2)} \quad (10)$$

where the sums run over all spin orbitals i and a . Spin integration yields the following spin-unrestricted expression,

$$\alpha_{\lambda\sigma}(i\omega) = 2 \sum_{ia}^{\alpha \text{ spin}} \frac{\epsilon_{ia} \langle i | \hat{\mu}_\lambda | a \rangle \langle a | \hat{\mu}_\sigma | i \rangle}{\hbar(\epsilon_{ia}^2 + \omega^2)} + 2 \sum_{\bar{i}\bar{a}}^{\beta \text{ spin}} \frac{\epsilon_{\bar{i}\bar{a}} \langle \bar{i} | \hat{\mu}_\lambda | \bar{a} \rangle \langle \bar{a} | \hat{\mu}_\sigma | \bar{i} \rangle}{\hbar(\epsilon_{\bar{i}\bar{a}}^2 + \omega^2)} \quad (11)$$

where i and a refer to α spin orbitals and $\bar{i}$ and $\bar{a}$ refer to β spin orbitals. In the spin restricted case, the expression for the frequency-dependent polarizability further simplifies to,

$$\alpha_{\lambda\sigma}(i\omega) = 4 \sum_{ia} \frac{\epsilon_{ia} \langle i | \hat{\mu}_\lambda | a \rangle \langle a | \hat{\mu}_\sigma | i \rangle}{\hbar(\epsilon_{ia}^2 + \omega^2)} \quad (12)$$

Finally, the isotropic frequency-dependent polarizabilities used as inputs for the D3 model are computed as the trace of the frequency dependent polarizability tensors $\alpha_{\lambda\sigma}(i\omega)$.

From these isotropic UCHF polarizabilities, the C_6 coefficients were computed by subtracting out the approximate hydrogen contribution and performing Casimir-Polder integration over imaginary frequency according to Eq 4. The integration was performed via

quadrature at ten frequencies given by,

$$i\omega_j = \frac{i}{\tan \left[\frac{\pi}{4N}(2j-1) \right]} \quad (13)$$

for $j = 1, 2, \dots, 10$ and with integration weights g_j :

$$g_j = \frac{\pi}{2N \sin^2 \left[\frac{\pi}{4N}(2j-1) \right]} \quad (14)$$

Empirical testing indicates that ten quadrature points is sufficient to obtain well-converged dispersion coefficients.

The frequency-dependent polarizabilities $\alpha(i\omega)$ were evaluated using a modified version of Molpro 2012.²³ While most of the hydrides used in the D3 model involve closed-shell species (spin restricted wavefunctions), there are some open-shell species for which the unrestricted spin formalism is necessary. Open-shell species include many bare elements (e.g. H, C, N, O) and low-coordination number hydrides (e.g. CH, OH). The spin-restricted expressions were already available in Molpro as part of the MP2C implementation, and the spin-unrestricted variant was implemented in a local version of Molpro.

At present, UCHF frequency-dependent polarizabilities have been computed for all necessary hydrides of H, B, C, N, O, F, Ne, P, S, Cl, Ar, and Br. These represent some of the most common elements occurring in organic chemistry. Extending the list of elements further would be straightforward, though not all elements would be well-described with an MP2-based model (e.g. transition metals). The hydride geometries and basis sets used to obtain these frequency-dependent polarizabilities are identical to those used in the original D3 work.¹⁶

Table 1 presents several sample C_6 coefficients for C-C interactions. Broadly speaking, as the level of hydrogen saturation decreases, the atomic polarizability and therefore magnitude of the contribution to the dispersion energy should increase. Accordingly, the dispersion coefficients increase in magnitude from ethane to ethene and ethyne. The dispersion energy

Table 1: Sample UCHF and CKS C_6 coefficients for carbon-carbon interactions between two identical carbons for different species/coordination environments as computed according to the D3 scheme.

	UCHF C_6	CKS C_6
Ethane	24.1	18.3
Ethene	34.9	25.7
Ethyne	41.2	29.5
Benzene	34.8	25.6

for such systems is typically overestimated at the UCHF level. This manifests in the UCHF coefficients listed in Table 1, which are 30–40% larger than the CKS ones.

It is notable that the local coordination number scheme used to interpolate the C_6 coefficients for the given chemical environment barely differentiates between the aromatic bonds in a species like benzene and the double-bond environment of ethene. The resulting dispersion coefficients are nearly identical for both cases. In reality, the dispersion coefficients should be somewhat larger for the aromatic species. This translates to MP2D underestimating the magnitude of the dispersion correction in the benzene π dimer at the S66x8 equilibrium separation, for example. MP2D reduces the MP2 binding energy from 4.9 kcal/mol to 3.3 kcal/mol, versus 2.8 kcal/mol for MP2C and 2.7 kcal/mol for the CCSD(T) benchmark. The dispersion correction here also would not capture the sorts of system-size-dependent changes in the C_6 coefficients observed for large carbon nanotubes or graphene,²⁴ for example. Nevertheless, the results in Section 4 will demonstrate that MP2D performs well overall across a broad range of chemical systems.

2.4 Short-range damping

The MP2D dispersion correction must be damped at short-ranges to avoid unphysical behavior. Here, Tang-Toennies damping,²⁵ is used to attenuate the dispersion correction at short interatomic separations instead of the Becke-Johnson or zero-damping used in D3²⁶. The physically-motivated Tang-Toennies damping function is well-suited for reproducing the dispersion energy in correlated methods.²⁷ The Tang-Toennies damping expression is given

by,

$$f_N(R_{ab}) = 1 - \exp(s_R R_{ab}) \sum_{k=0}^N \frac{(s_R R_{ab})^k}{k!} \quad (15)$$

where N is the order of the dispersion term, i.e. 6 and 8, and s_R is a distance scaling factor calculated from the cutoff radius $R_{0,ab}$ (taken from the D3 dispersion correction) using two empirical parameters:

$$s_R = a_1 R_{0,ab} + a_2. \quad (16)$$

Fitting the damping functions to UCHF and CKS dispersion energies separately, we found that optimal damping parameters a_1 and a_2 for the UCHF dispersion energies differ from those for the CKS ones. This would give rise to four parameters, a_1^{UCHF} , a_2^{UCHF} , a_1^{CKS} and a_2^{CKS} . However, the number of parameters can be reduced as described in Section 2.6 below.

Although the Tang-Toennies damping provides a physically sound treatment in the non-covalent regime, it damps insufficiently at covalent distances. This leads to deteriorated MP2D description of reaction energies, for example. We thus introduce a secondary short-ranged damping that ensures that dispersion correction becomes constant at covalent distances. This damping is achieved by modifying the value of the interatomic distance R_{ab} that enters Eq 3 as

$$R'_{ab} = \begin{cases} r_{cut} R_{0,ab} & \text{if } R_{ab} \leq R_{0,ab}(r_{cut} - w/2) \\ R_{ab} & \text{if } R_{ab} \geq R_{0,ab}(r_{cut} + w/2) \\ r_{cut} R_{0,ab} + f(R_{ab}, R_{0,ab} r_{cut}, R_{0,ab} w) & \text{otherwise;} \end{cases} \quad (17)$$

$$f(R_{ab}, r'_{cut}, w') = (-2.5x^8 + 10x^7 - 14x^6 + 7x^5) * w'; \quad (18)$$

$$x = \frac{R_{ab} - (r'_{cut} - w'/2)}{w'} \quad (19)$$

where two more parameters are introduced, r_{cut} and w , which are defined as dimensionless factors scaling the radius $R_{0,ab}$ taken from the D3 correction. Eq 17 leaves R_{ab} untouched

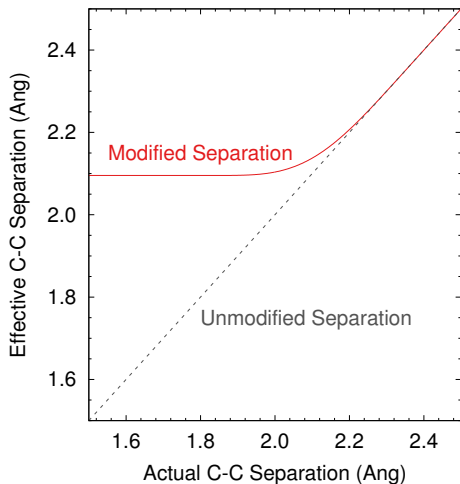


Figure 1: Example of how the secondary short-range damping modifies the effective C-C interatomic separation based on a cutoff radius of $R_{0,ab} = 2.9103 \text{ \AA}$, $r_{cut} = 0.72$, and $w = 0.2$.

for distances greater than $R_{0,ab}(r_{cut} + w/2)$, and it fixes R_{ab} at a constant fraction of $R_{0,ab}$ for distances that are shorter than $R_{0,ab}(r_{cut} - w/2)$. The third portion of the function in Eq 17 smooths the transition between these two regimes. The high-order polynomial ensures smooth first, second and third derivatives at the end points of the switching interval.

This damping is applied at very short distances so that it practically does not affect intermolecular non-covalent interactions. For example, using the final optimized parameters described in Section 2.6, this damping smoothly alters the effective interatomic separation for two carbon atoms from the actual separation to a fixed value near $2 \text{ \AA}$ and below (Figure 1). The fractional nature of r_{cut} and w means that these distances adapt depending on the threshold radius $R_{0,ab}$ for the given atom pair. The overall dispersion energy is therefore doubly damped: first by Tang-Toennies in the non-covalent regime, and second by this short-range damping in the covalent regime.

2.5 Modified C_6 interpolation

In D3, the C_6 coefficients are interpolated using a continuous coordination number CN calculated from distances to all other atoms using a switching function (Eq 5). Although this switching function decays quickly, it yields small but nonzero contributions even at non-

covalent distances. This makes the C_6 coefficients in a dimer slightly different to these in isolated monomers in the same geometry, and this propagates also to the C_8 coefficients. This proves problematic at short distances (where the C_8 term becomes important) when a weaker damping function is used.²⁷ In MP2D, we eliminated this issue by replacing the switching function with one that drops exactly to zero at larger separations. Again, the polynomial interpolating between the short- and long-range regimes was designed to have smooth first and second derivatives at the end points of the switching interval. The scaling factors in this function were fitted to closely reproduce the original D3 one. The MP2D coordination number is now calculated as,

$$CN = \sum_{B \neq A}^N f(R_{ab}^{cov}, R_{ab}); \quad (20)$$

$$f(R_{ab}^{cov}, R_{ab}) = \begin{cases} 1.0 & \text{if } R_{ab} \leq 0.95R_{ab}^{cov} \\ 0.0 & \text{if } R_{ab} \geq 1.75R_{ab}^{cov} \\ f'(x) & \text{otherwise;} \end{cases} \quad (21)$$

$$f'(x) = 1.0 - (-20x^7 + 70x^6 - 84x^5 + 35x^4); \quad (22)$$

$$x = \frac{R_{ab} - 0.95R_{ab}^{cov}}{1.75R_{ab}^{cov} - 0.95R_{ab}^{cov}}. \quad (23)$$

In the geometry optimizations reported here, we used integer coordination numbers to simplify the calculation of the gradient. All the studied systems have well-defined geometries where the continuous valence numbers differ only negligibly from integer ones, so this assumption does not introduce any appreciable error. One could implement gradients for the continuous valence coordination numbers if desired.

2.6 MP2D Parameterization

The MP2D model described thus far could conceivably employ up to ten potential global parameters: separate s_6 and s_8 scaling terms for each of the UCHF and CKS C_6 and C_8

dispersion energies (four parameters), two parameters for the UCHF Tang-Toennies damping function (a_1^{UCHF} and a_2^{UCHF}), two parameters for the CKS Tang-Toennies damping function (a_1^{CKS} and a_2^{CKS}), and two parameters for the secondary short-range damping (r_{cut} and w). To obtain correct dispersion energies at long distances, the s_6 parameter is set to unity for both UCHF and CKS dispersion. In exploring the parameterization, we found that the value of the s_8 parameter is similar in both the UCHF and CKS cases, so we use a single global s_8 parameter for both. This reduces the number of possible global fitting parameters from ten to seven. At this point, we tested multiple variants of the parameterization protocol, and analyzed the results obtained with different parameter sets, exploring the possibilities to simplify the method further.

First, we tested fitting the CKS and UCHF Tang-Toennies terms separately to the corresponding CKS and UCHF energies from MP2C calculations performed on the S66x8 benchmark data set.²⁸ The resulting dispersion coefficient model reproduced the original CKS and UCHF energies rather well. The optimal parameter values in the CKS and UCHF Tang-Toennies damping functions differed significantly. Next, to capture some higher-order contributions not covered by MP2C, the Tang-Toennies fit parameters were refined against CCSD(T)/CBS interaction energies on the same S66x8 data (i.e. fitting to the energy difference between MP2 and CCSD(T) instead of the raw MP2C energy components). When the parameterization was started from the UCHF and CKS parameter values fitted to the CKS and UCHF dispersion energy components separately, the resulting Tang-Toennies damping parameters changed only very slightly.

However, further testing found that an equally good fit to the post-MP2 correlation energy can be obtained when the same values of the parameters are used in both the CKS and UCHF damping functions. Using common parameters reduces the fidelity with which MP2D reduces the individual UCHF and CKS dispersion energy components, but it has no appreciable negative impact on the quality of the net dispersion correction. We decided that reduction in the number of adjustable parameters was more useful than reproducing

the individual dispersion energy components. Setting

$$a_1^{UCHF} = a_1^{CKS} = a_1, \text{ and} \quad (24)$$

$$a_2^{UCHF} = a_2^{CKS} = a_2. \quad (25)$$

reduces the number of global parameters down to five (s_8 , a_1 , a_2 , r_{cut} , and w). Figure S1 in the Supporting Information provides an sample comparison for MP2D with and without constraining the Tang-Toennies parameters to be identical.

During the parameterization procedure thus far, the S66x8 data set proved sufficient for a robust and transferable parameterization of the correction at non-covalent distances. The three parameters affecting the dispersion energy in this range of distances (a_1 , a_2 and s_8) were thus optimized first on the S66x8 data set with the short-ranged damping disabled. Subsequently, the initial values of the remaining two parameters in the short-range damping ($r_{cut} = 0.7$ and $w = 0.2$) were manually chosen to yield the best compromise between short intermolecular interactions (in the S66x10 data set) and conformation energies (using all the conformer data sets considered later in the paper). Overall, a fairly broad range of values for r_{cut} and w provide comparably good performance. Larger values of r_{cut} would be obtained if the method was optimized on thermochemistry data, but the description of non-covalent interactions at short distances would be compromised while reaction energies would improve only by about 0.5 kcal/mol.

Finally, all the five parameters were fitted again to the S66x8 data set, resulting in only small change to the values of r_{cut} . The resulting parameters thus represent a minimum with respect to non-covalent interactions around and above equilibrium distance (represented by the S66x8 data set) which can be expected to be the main application targets for the method. At the same time, the formulation of the model and the choice of the initial values of the parameters used additional information from conformational energies needed to provide a seamless connection between the covalent and non-covalent regime.

Table 2 lists the final values of the parameters. Figure 2 plots the MP2D dispersion correction energy (C_6 contributions only for simplicity) for two carbon atoms as a function of distance with no damping, just Tang-Toennies damping, and finally the actual doubly-damped model that also includes the covalent regime short-range damping.

Table 2: Optimized values of the five global parameters used in the MP2D method.

s_8	1.187	dimensionless
a_1	0.944	dimensionless
a_2	0.480	Ångstrom
r_{cut}	0.72	dimensionless
w	0.20	dimensionless

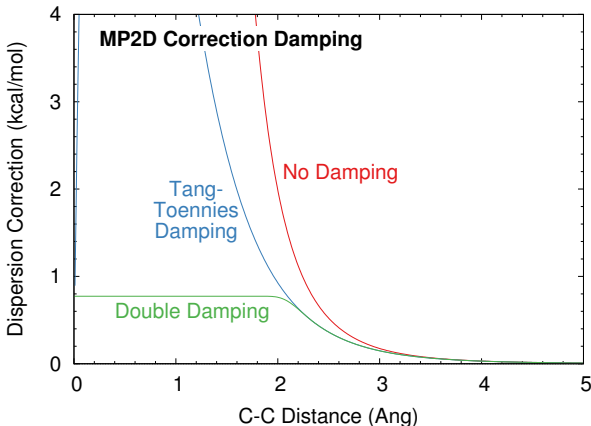


Figure 2: Behavior of the MP2D dispersion correction (C_6 only here) without damping, with Tang-Toennies damping, and the double Tang-Toennies/shorter-range damping for two sp^2 -hybridized carbon atoms.

The MP2D dispersion correction has been fitted to reproduce the counterpoise-corrected $\Delta\text{CCSD(T)}$ energy correction. The remaining question is how to treat the basis set superposition error (BSSE) in the MP2 part of the calculation. When the MP2 energy is calculated in a large basis set or extrapolated to the CBS limit, the MP2D method should be universally applicable to both inter- and intramolecular energies because the BSSE would be smaller than the error of the dispersion correction.

This can be demonstrated on the calculations of the interaction energies in the S66 data set performed with and without counterpoise (CP) correction. At the CBS limit, using

MP2D without CP correction yields smaller RMSE (0.19 kcal/mol) than when CP correction is applied (0.26 kcal/mol). The good performance without CP correction results from error cancellation, but it shows that the energy changes associated with the CP correction are several times smaller than the overall error. In the aug-cc-pVQZ basis, the CP-corrected and uncorrected results are very similar (RMSE 0.35 and 0.37 kcal/mol). In smaller basis sets, the CP uncorrected interaction energies become significantly worse.

3 Computational Methods

The dispersion correction was implemented in the freely available Cuby4 framework,²² which interfaces multiple computational chemistry packages that could provide the MP2 calculation. The MP2D dispersion correction implementation, including the dispersion coefficients for the 13 common first- and second-row elements noted above, documentation, and input examples, are provided at the Cuby website.²⁹

Electronic structure calculations were carried out using a mixture of PSI4,³⁰ Molpro 2012.1,²³ and TURBOMOLE.^{31,32} All single-point MP2, MP2D, MP2C, MP2.5, and CCSD(T) results reported here were extrapolated to the complete basis set (CBS) limit,³³ typically from the Dunning aug-cc-pVXZ basis sets.³⁴ For MP2.5 and CCSD(T), the CBS limit was estimated using the standard focal point technique which combines MP2/CBS with post-MP2 correlation estimated in a smaller basis set.^{4,35} B3LYP-D3(BJ) results employ the nearly-complete def2-QZVP basis. Geometry optimizations were performed in the def2-TZVP basis with no counterpoise correction. Integer coordination numbers of the atoms were employed for the dispersion coefficients in MP2D geometry optimizations. All calculations here employed density fitting with standard auxiliary basis sets throughout. Counterpoise corrections for basis set superposition error were employed in the MP2-based methods (excluding the geometry optimizations). Because the DFT-D3(BJ) damping parameters were fitted without counterpoise correction,³⁶ no counterpoise corrections were applied to the DFT re-

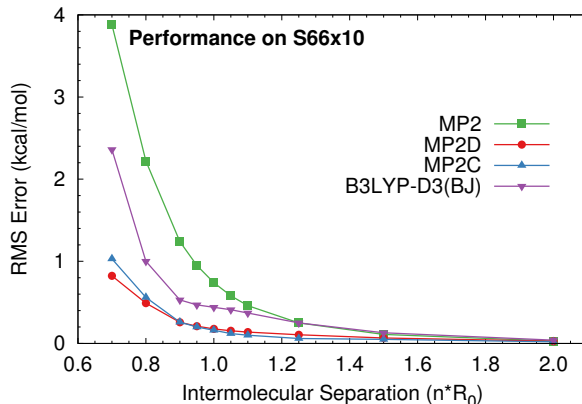


Figure 3: Performance of MP2 and other methods on the S66x10 benchmark test set used in fitting the empirical parameters.

sults here.

4 Results and Discussion

4.1 Energetics

The results here will demonstrate that MP2D performs very competitively with other techniques across a wide variety of systems, including intermolecular interactions, conformational energies, and thermochemistry. For comparison purposes, B3LYP-D3(BJ) was chosen as a representative, widely used density functional that generally performs well for non-covalent interactions.^{37–40} Comparison against other functionals (particularly double-hybrid density functionals⁴⁰) would make for an interesting subject of future work.

Non-covalent interactions at short range. Consider first the sixty-six dimers at ten intermolecular separations comprising the S66x10 benchmark test set.⁴¹ The eight largest intermolecular separations are identical to the S66x8 set used in the parameter fitting, while the shortest two distances ($0.7R_0$ and $0.8R_0$) were not part of the final fit. Figure 3 plots the root-mean-square (rms) errors versus intermolecular distance. At equilibrium distances $1.0R_0$, the 0.18 kcal/mol rms error for MP2D is much better than MP2 (1.24 kcal/mol), appreciably smaller than B3LYP-D3(BJ) (0.44 kcal/mol), and almost as good as MP2C

(0.16 kcal/mol).

The MP2D performance improves further relative to the other methods at shorter separations. As the dimer separation decreases, the fraction of the MP2C UCHF and CKS dispersion energies captured by MP2D decreases, with the short-range damping effectively mimicking some of the repulsive exchange-dispersion terms which are not corrected in MP2C. At $0.7R_0$ for instance, which was not included in the training data, the MP2D error is only 0.82 kcal/mol, versus 1.03 kcal/mol for MP2C and 2.36 kcal/mol for B3LYP-D3(BJ). Even the recently proposed B3LYP-D3M(BJ), which seeks to improve the short-range behavior of D3, exhibits a significantly larger error of 1.57 kcal/mol at this separation (those reported results were counterpoise corrected).⁴¹

Of course, MP2D is only useful if it also performs well broadly, beyond the systems the empirical parameters were fitted for. Figure 4 reports box plot error distributions for MP2, MP2D, MP2C, and B3LYP-D3(BJ) across 10 different benchmark sets consisting of several thousand diverse examples and lists the rms errors.

Interaction energies in more diverse systems. For the 3380 protein side chain-side chain interactions in SSI,³⁸ the MP2D errors are two-thirds smaller than MP2, comparable to MP2C, and half those for B3LYP-D3(BJ) (Figure 4a). For the halogen-containing dimers in X40⁴² (excluding iodine-containing species for which the Molpro MP2C implementation fails) and the S22 set,^{43,44} MP2D reduces the MP2 errors 2–4-fold, though they are larger than the MP2C ones and only moderately better than B3LYP-D3(BJ). Some of the largest MP2D errors occur for π -stacked cases, for which it corrects much of the MP2 error, but not as effectively as MP2C. Such π -stacked cases make up a disproportionately large fraction of the S22 set compared to S66x8 or SSI, which helps explain the larger difference between MP2C and MP2D for that set.

Conformation energies. Figure 4b examines conformational energy benchmarks for alkanes (ACONF⁴⁵), amino acids (Amino20x4⁴⁰), sugars (SCONF⁴⁰), and short peptides (from MPCONF196⁴⁶). MP2D improves upon MP2 in all cases, reducing the rms errors

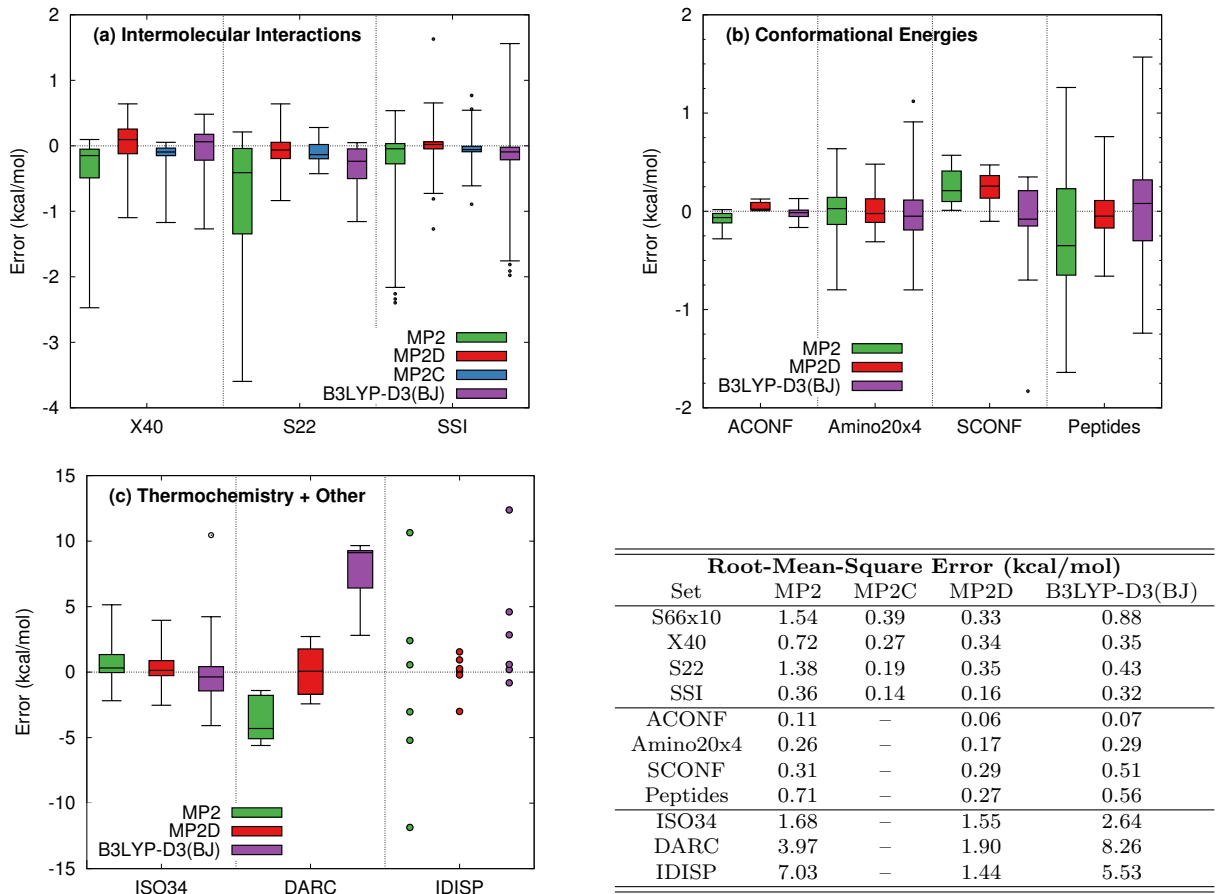


Figure 4: Performance of MP2D and other methods for benchmark sets focusing on (a) intermolecular interaction energies, (b) conformational energies, and (c) thermochemistry. MP2C is only defined for the intermolecular interactions. Boxes and whiskers contain 50% and 95% (or 99.9% for part (a)) of the data, respectively. The table lists root-mean-square errors for all methods and sets.

and generally narrowing the width of the error distributions. The largest improvements occur for the amino acids and peptides where van der Waals interactions are relatively large. Dispersion plays the smallest role the sugar conformers (SCONF), and the MP2D improvement there is correspondingly small. MP2D outperforms B3LYP-D3(BJ) appreciably for Amino20x4, SCONF, and the Peptides.

Thermochemistry. While intermolecular interactions and conformational energies represent traditional applications for methods like MP2D, it is equally important that MP2D should not interfere with thermochemistry. Figure 4c plots errors for small-molecule isomerizations (ISO34⁴⁷) and Diels-Alder reactions (DARC⁴⁸). Dispersion plays a minimal role in

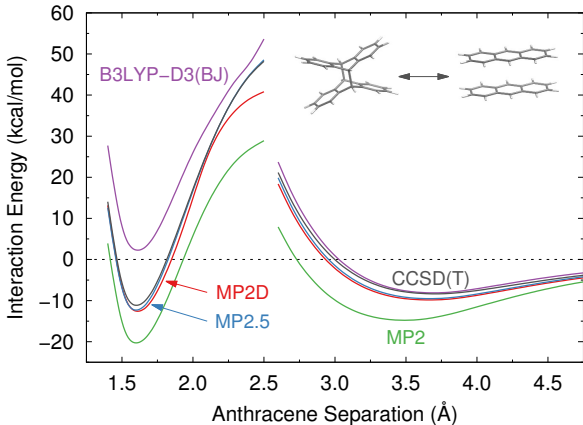


Figure 5: Potential energy basins for covalent anthracene photodimer (near 1.6 Å) and non-covalent π -dimer (near 3.6 Å). B3LYP-D3 employs def2-QZVP basis; all other methods at extrapolated CBS limit.

the small-molecule isomerization reactions found in ISO34, and the MP2D correction only reduces the MP2 error by 8%. Dispersion is much more important in the Diels-Alder reactions (DARC), and MP2D cuts the MP2 error in half and substantially out-performs B3LYP-D3(BJ). Also impressive, however, is the MP2D performance for IDISP,⁴⁰ which consists of four chemical reactions and two conformational changes involving intramolecular dispersion. MP2D reduces the 7.0 kcal/mol rms error for MP2 to only 1.4 kcal/mol, compared to 5.5 kcal/mol for B3LYP-D3(BJ).

While a thorough comparison of density functionals is beyond the scope of this letter, many of the test sets considered here are part of the GMTKN55 suite,⁴⁰ for which results from many density functionals have been reported. Analysis of those results suggests that MP2D is competitive with or better than the best dispersion-corrected hybrid functionals. The best double hybrid functionals sometimes perform moderately better (with similar computational cost to MP2D), but those functionals frequently employ empirical spin-component scaling of the MP2 correlation (e.g. DSD-BLYP-D3(BJ)⁴⁹), which is not exploited here.

Anthracene dimerization. For further insight on how MP2D performs, consider the anthracene photodimerization reaction, which is the most challenging case in the IDISP set. This reaction includes both a non-covalent π -stacked dimer with intermolecular separation

~ 3.6 Å and a covalent photoreacted dimer with intramolecular separation ~ 1.6 Å. It proves problematic for many electronic structure methods.⁵⁰ Whereas the MP2C dispersion correction is defined only for the intermolecular dimer, MP2D can describe both regimes. A benchmark CCSD(T)/CBS potential energy scan was created for this reaction, as shown in Figure 5 and described in the Supporting Information.

Figure 5 shows that various methods perform well for the intermolecular dimer energy basin, with MP2 performing the worst due to its overestimation of the π - π interactions. MP2D and MP2.5^{51,52} perform very similarly, overbinding the π -dimer by 1.2–1.5 kcal/mol relative to CCSD(T). B3LYP-D3(BJ) does even better, underbinding it by 0.2 kcal/mol. The real challenge, however, occurs in modeling the covalent basin and the competition among the long covalent bonds between the two anthracenes, the anthracene ring distortion, and the very short-range dispersion interactions between the anthracene rings. Unsurprisingly, MP2 overbinds by 9.1 kcal/mol. At the other extreme, B3LYP-D3(BJ) underestimates the stability of the covalent dimer by 13.4 kcal/mol, predicting the photodimerization reaction to be significantly endothermic instead of slightly exothermic. In contrast, MP2.5 and MP2D reproduce CCSD(T) nicely across most of the potential energy surface, with 1.2–1.5 kcal/mol errors at the minima. Both methods reproduce the energy difference between the two minima to within less than 0.1 kcal/mol. Only near ~ 2.5 Å does MP2D perform appreciably worse than CCSD(T) or MP2.5, where the increasingly significant static correlation arising from the stretching of the two inter-anthracene covalent bonds is ill-described by spin-restricted MP2. The ability to accurately describe both intra- and intermolecular interactions simultaneously is a key feature of MP2D.

4.2 Geometries

Facile geometry optimization represents another advantage of MP2D over MP2C for systems with significant non-covalent interactions. Consider the challenging example of [7]helicene—seven fused benzene rings arranged in a helix. Figure 6 overlays the structures optimized with

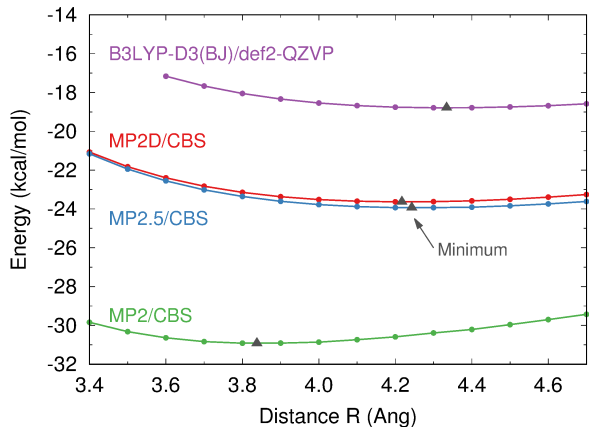
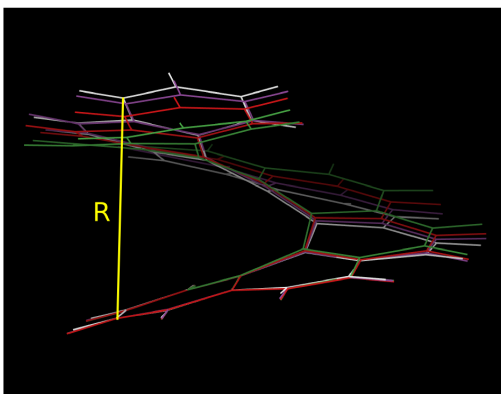


Figure 6: Overlay of [7]helicene crystal structure⁵³ with that predicted by MP2 (green), MP2D (red), and B3LYP-D3(BJ) (purple) in the def2-TZVP basis. Single-point energies relative to linear heptacene as a function of the distance R .

MP2, MP2D, and B3LYP-D3(BJ) in the def2-TZVP basis with the X-ray crystal structure. MP2 overestimates the van der Waals interactions and artificially compresses the helical spacing R at 3.72 Å, versus ~ 4.4 – 4.6 Å from the experimental crystal structures.^{53–55} MP2D significantly corrects this to 4.08 Å, and B3LYP-D3(BJ) predicts a seemingly even better 4.32 Å.

However, the potential energy surface for compressing [7]helicene is very flat, and the experimental crystal structure may differ from the gas-phase electronic energy one due to solid-state packing forces. To investigate, a one-dimensional B3LYP-D3(BJ)/def2-TZVP relaxed scan over R was performed (see Supporting Information for details). Large-basis single-point energies were computed with B3LYP-D3(BJ), MP2, MP2D, and MP2.5 and plotted relative to the energy of the isomeric linear heptacene in Figure 6. Linear heptacene provides a useful reference structure, since the planar molecule has no significant non-covalent interactions. Taking MP2.5 as the reference energy, MP2 significantly overestimates the interactions and underestimates the optimal R . B3LYP-D3(BJ) underestimates the interaction energy by a quarter, and it overestimates the distance R , while MP2D reproduces MP2.5 quite well and drastically lower computational cost. These results imply that the MP2D optimized structure is actually closer than the B3LYP-D3(BJ) one to the true gas-phase structure.

Table 3: Root-mean-square deviations (rmsd) in pm for the optimized geometries for 10 benzene dimer stationary points relative to reference structures obtained with a mixed DFT/CCSD(T) approach.⁵⁶ Structure S4 causes significant problems for MP2 and B3LYP-D3(BJ), so the rmsd is reported with and without S4 included. Counterpoise corrections were not employed.

Structure	MP2 def2-TZVP	MP2D def2-TZVP	B3LYP-D3(BJ) def2-TZVP
M1	11.4	3.1	2.7
M2	6.2	2.3	0.7
S1	7.2	4.9	1.2
S2	11.7	3.4	2.9
S3	8.3	3.9	2.7
S4	90.5	3.5	12.2
S5	3.8	1.0	1.0
S6	2.5	1.3	0.9
S7	11.1	3.6	4.7
S8	10.4	2.9	4.6
RMSD (all)	29.8	3.2	4.7
RMSD (excluding S4)	8.7	3.2	2.8

The ten stationary points on the benzene dimer potential energy surface⁵⁶ provide another interesting case. For nine of the ten structures, B3LYP-D3(BJ) and MP2D reproduce the reference DFT + CCSD(T) structures well, with root-mean-square RMSD of 2.8 and 3.2 pm respectively (Table 3). However, the geometry of the S4 structure varies strongly with the method, as shown in Figure 7. In the reference structure, the benzene molecules interact at a 53.8° angle. MP2 fails completely for this structure, optimizing to a parallel π -stacked structure, while B3LYP-D3(BJ) underestimates the angle at 45.2° (RMSD 12.2 pm). In contrast, MP2D predicts a 51.8° angle of the reference structure and gives an RMSD of only 3.5 pm.

5 Conclusions

In summary, MP2D largely corrects the key dispersion-related flaws of MP2 with trivial computational cost using a dispersion correction based on the Grimme D3 scheme. It relies on pre-tabulated *ab initio* dispersion coefficients and five universal empirical parameters de-



Figure 7: Overlay of benzene dimer S4, showing how MP2D (red) reproduces the correct angle between the molecules, unlike MP2 (green) and B3LYP-D3(BJ) (purple).

signed to attenuate the correction at short range and compensate for higher-order dispersion contributions. Unlike MP2C, MP2D improves the description of intramolecular dispersion and can be used for geometry optimizations. The results here indicate that MP2D provides a valuable alternative to DFT in systems where van der Waals interactions are important, ranging from organics to biomolecules. A software implementation of the MP2D dispersion correction that can be easily coupled with MP2 calculations in many computational chemistry programs is freely available.²⁹

Future work should compare MP2D against a broader suite of density functionals. It would also be interesting to pursue a spin-component-scaled version of MP2D. One of the key problems in spin-component-scaled methods has been the difficulty of finding parameters that simultaneously improve the treatment of thermochemistry and non-covalent interactions. The MP2D dispersion correction addresses the non-covalent interactions, which would allow the spin-component scaling to correct the thermochemistry errors. Such an approach could provide an interesting alternative to some of the best-performing double-hybrid density functionals which also employ spin-component-scaling and a similar number of global

empirical parameters.

6 Acknowledgments

Funding for this work from the National Science Foundation (CHE-1665212), Czech Science Foundation (P208/12/G016) and supercomputer time from XSEDE (TG-CHE110064) are gratefully acknowledged. The work is part of research project RVO:61388963 of the IOCB AS CR. We thank Prof. Filipp Furche for help validating the UCHF frequency-dependent polarizability implementation, and Prof. Stefan Grimme for providing the basis sets used to compute dispersion coefficients in the original D3 work.

Supporting Information Available

Additional details on the formulation of the method, values of the empirical parameters used, and newly calculated geometries and benchmark values for the validation systems are provided as a Supporting Information.

References

- (1) Řezáč, J.; Hobza, P. Benchmark Calculations of Interaction Energies in Noncovalent Complexes and Their Applications. *Chem. Rev.* **2016**, *116*, 5038–5071.
- (2) Grimme, S.; Hansen, A.; Brandenburg, J. G.; Bannwarth, C. Dispersion-Corrected Mean-Field Electronic Structure Methods. *Chem. Rev.* **2016**, *116*, 5105–5154.
- (3) Cohen, A. J.; Mori-Sanchez, P.; Yang, W. Challenges for Density Functional Theory. *Chem. Rev.* **2012**, *112*, 289–320.
- (4) Riley, K. E.; Pitonak, M.; Jurecka, P.; Hobza, P. Stabilization and structure calculations

- for noncovalent interactions in extended molecular systems based on wave function and density functional theories. *Chem. Rev.* **2010**, *110*, 5023–63.
- (5) Sinnokrot, M. O.; Sherrill, C. D. Highly accurate coupled cluster potential energy curves for the benzene dimer: Sandwich, T-shaped, and parallel-displaced configurations. *J. Phys. Chem. A* **2004**, *108*, 10200–10207.
 - (6) Cybulski, S. M.; Lytle, M. L. The origin of deficiency of the supermolecule second-order Møller-Plesset approach for evaluating interaction energies. *J. Chem. Phys.* **2007**, *127*, 141102.
 - (7) Hesselmann, A. Improved supermolecular second order Møller-Plesset intermolecular interaction energies using time-dependent density functional response theory. *J. Chem. Phys.* **2008**, *128*, 144112.
 - (8) Gerenkamp, M.; Grimme, S. Spin-component scaled second-order Møller-Plesset perturbation theory for the calculation of molecular geometries and harmonic vibrational frequencies. *Chem. Phys. Lett.* **2004**, *392*, 229–235.
 - (9) Hill, J. G.; Platts, J. A. Spin-component scaling methods for weak and stacking interactions. *J. Chem. Theory Comput.* **2007**, *3*, 80.
 - (10) Distasio, 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.
 - (11) Marchetti, O.; Werner, H.-J. Accurate Calculations of Intermolecular Interaction Energies Using Explicitly Correlated Coupled Cluster Wave Functions and a Dispersion-Weighted MP2 Method. *J. Phys. Chem. A* **2009**, *113*, 11580–11585.
 - (12) Tan, S.; Barrera Acevedo, S.; Izgorodina, E. I. Generalized spin-ratio scaled MP2

- method for accurate prediction of intermolecular interactions for neutral and ionic species. *J. Chem. Phys.* **2017**, *146*, 064108.
- (13) McGibbon, R. T.; Taube, A. G.; Donchev, A. G.; Siva, K.; Hernández, F.; Hargus, C.; Law, K.-H.; Klepeis, J. L.; Shaw, D. E. Improving the accuracy of Møller-Plesset perturbation theory with neural networks. *The Journal of Chemical Physics* **2017**, *147*, 161725.
 - (14) Pitonak, M.; Hesselmann, A. Accurate Intermolecular Interaction Energies from a Combination of MP2 and TDDFT Response Theory. *J. Chem. Theory Comput.* **2010**, *6*, 168–178.
 - (15) Burns, L. A.; Marshall, M. S.; Sherrill, C. D. Appointing silver and bronze standards for noncovalent interactions: A comparison of spin-component-scaled (SCS), explicitly correlated (F12), and specialized wavefunction approaches. *J. Chem. Phys.* **2014**, *141*, 234111.
 - (16) 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.
 - (17) Tkatchenko, A.; DiStasio, R. A.; Head-Gordon, M.; Scheffler, M. Dispersion-corrected Møller-Plesset second-order perturbation theory. *J. Chem. Phys.* **2009**, *131*, 094106.
 - (18) Huang, Y.; Beran, G. J. O. Reliable prediction of three-body intermolecular interactions using dispersion-corrected second-order Møller-Plesset perturbation theory. *J. Chem. Phys.* **2015**, *143*, 044113.
 - (19) Chalasinski, G.; Szczesniak, M. M.; Cybulski, S. M. Calculations of nonadditive effects by means of supermolecular Møller-Plesset perturbation theory approach: Ar₃ and Ar₄. *J. Chem. Phys.* **1990**, *92*, 2481–2487.

- (20) von Lilienfeld, O. A.; Tkatchenko, A. Two- and three-body interatomic dispersion energy contributions to binding in molecules and solids. *J. Chem. Phys.* **2010**, *132*, 234109.
- (21) Risthaus, T.; Grimme, S. Benchmarking of London Dispersion-Accounting Density Functional Theory Methods on Very Large Molecular Complexes. *J. Chem. Theory Comput.* **2013**, *9*, 1580–1591.
- (22) Řezáč, J. Cuby: An integrative framework for computational chemistry. *J. Comp. Chem.* **2016**, *37*, 1230–1237.
- (23) MOLPRO, version 2012.1, a package of ab initio programs, H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz, P. Celani, T. Korona, R. Lindh, A. Mitrushenkov, G. Rauhut, K. R. Shamasundar, T. B. Adler, R. D. Amos, A. Bernhardsson, A. Berning, D. L. Cooper, M. J. O. Deegan, A. J. Dobbyn, F. Eckert, E. Goll, C. Hampel, A. Hesselmann, G. Hetzer, T. Hrenar, G. Jansen, C. Köppl, Y. Liu, A. W. Lloyd, R. A. Mata, A. J. May, S. J. McNicholas, W. Meyer, M. E. Mura, A. Nicklass, D. P. O’Neill, P. Palmieri, D. Peng, K. Pflüger, R. Pitzer, M. Reiher, T. Shiozaki, H. Stoll, A. J. Stone, R. Tarroni, T. Thorsteinsson, and M. Wang, see <http://www.molpro.net>.
- (24) Gobre, V. V.; Tkatchenko, A. Scaling laws for van der Waals interactions in nanostructured materials. *Nature Commun.* **2013**, *4*, 2341.
- (25) Tang, K. T.; Toennies, J. P. An improved simple model for the van der Waals potential based on universal damping functions for the dispersion coefficients. *J. Chem. Phys.* **1984**, *80*, 3726–3741.
- (26) Goerigk, L.; Kruse, H.; Grimme, S. Benchmarking Density Functional Methods against the S66 and S66x8 Datasets for Non-Covalent Interactions. *ChemPhysChem* **2011**, *12*, 3421–3433.

- (27) Sedlak, R.; Řezáč, J. Empirical D3 Dispersion as a Replacement for ab Initio Dispersion Terms in Density Functional Theory-Based Symmetry-Adapted Perturbation Theory. *J. Chem. Theory Comput.* **2017**, *13*, 1638–1646.
- (28) Řezáč, J.; Riley, K. E.; Hobza, P. S66: A Well-Balanced Database of Benchmark Interaction Energies Relevant to Biomolecular Structures. *J. Chem. Theory Comput.* **2011**, *7*, 2427–2438.
- (29) Řezáč, J. Cuby 4, software framework for computational chemistry: MP2D method. 2015; <http://cuby4.molecular.cz/MP2D.html>.
- (30) Parrish, R. M.; Burns, L. A.; Smith, D. G. A.; Simmonett, A. C.; DePrince, A. E.; Hohenstein, E. G.; Bozkaya, U.; Sokolov, A. Y.; Di Remigio, R.; Richard, R. M.; Gonthier, J. F.; James, A. M.; McAlexander, H. R.; Kumar, A.; Saitow, M.; Wang, X.; Pritchard, B. P.; Verma, P.; Schaefer, H. F.; Patkowski, K.; King, R. A.; Valeev, E. F.; Evangelista, F. A.; Turney, J. M.; Crawford, T. D.; Sherrill, C. D. Psi4 1.1: An Open-Source Electronic Structure Program Emphasizing Automation, Advanced Libraries, and Interoperability. *Journal of Chemical Theory and Computation* **2017**, *13*, 3185–3197, PMID: 28489372.
- (31) TURBOMOLE V7.2 2017, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from <http://www.turbomole.com>.
- (32) Furche, F.; Ahlrichs, R.; Hättig, C.; Klopper, W.; Sierka, M.; Weigend, F. Turbomole. *WIREs Comput. Mol. Sci.* **2014**, *4*, 91–100.
- (33) Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. Basis-set convergence of correlated calculations on water. *J. Chem. Phys.* **1997**, *106*, 9639–9646.

- (34) Dunning, 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.
- (35) Császár, A. G.; Allen, W. D.; Schaefer, H. F. In pursuit of the ab initio limit for conformational energy prototypes. *The Journal of Chemical Physics* **1998**, *108*, 9751–9764.
- (36) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the Damping Function in Dispersion Corrected Density Functional Theory. *J. Comp. Chem.* **2011**, *32*, 1456–1465.
- (37) Burns, L. A.; Mayagoitia, A. V.; Sumpter, B. G.; Sherrill, C. D. Density-functional approaches to noncovalent interactions: A comparison of dispersion corrections (DFT-D), exchange-hole dipole moment (XDM) theory, and specialized functionals. *J. Chem. Phys.* **2011**, *134*, 084107.
- (38) Burns, L. A.; Faver, J. C.; Zheng, Z.; Marshall, M. S.; Smith, D. G.; Vanommeslaeghe, K.; MacKerell, A. D.; Merz, K. M.; Sherrill, C. D. The BioFragment Database (BFDdb): An open-data platform for computational chemistry analysis of noncovalent interactions. *J. Chem. Phys.* **2017**, *147*.
- (39) 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*, 6670–88.
- (40) 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.
- (41) Smith, D. G. A.; Burns, L. A.; Patkowski, K.; Sherrill, C. D. Revised Damping Parameters for the D3 Dispersion Correction to Density Functional Theory. *J. Phys. Chem. Lett.* **2016**, *7*, 2197–2203.

- (42) Řezáč, J.; Riley, K. E.; Hobza, P. Benchmark Calculations of Noncovalent Interactions of Halogenated Molecules. *J. Chem. Theory Comput.* **2012**, *8*, 4285–4292.
- (43) Jurečka, P.; Šponer, J.; Černý, 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.
- (44) Marshall, M. S.; Burns, L. A.; Sherrill, C. D. Basis set convergence of the coupled-cluster correction, δ MP2CCSD(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.
- (45) Gruzman, D.; Karton, A.; Martin, J. M. L. Performance of Ab Initio and Density Functional Methods for Conformational Equilibria of C_nH_{2n+2} Alkane Isomers ($n = 4-8$). *J. Phys. Chem. A* **2009**, *113*, 11974–11983.
- (46) Řezáč, J.; Bím, D.; Gutten, O.; Rulíšek, L. Toward Accurate Conformational Energies of Smaller Peptides and Medium-Sized Macrocycles: MPCONF196 Benchmark Energy Data Set. *J. Chem. Theory Comput.* **2018**, *14*, 1254–1266.
- (47) Grimme, S.; Steinmetz, M.; Korth, M. How to Compute Isomerization Energies of Organic Molecules with Quantum Chemical Methods. *J. Org. Chem.* **2007**, *72*, 2118–2126.
- (48) Johnson, E. R.; Mori-Sánchez, P.; Cohen, A. J.; Yang, W. Delocalization errors in density functionals and implications for main-group thermochemistry. *The Journal of Chemical Physics* **2008**, *129*, 204112.
- (49) 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.

- (50) Grimme, S.; Diedrich, C.; Korth, M. The Importance of Inter- and Intramolecular van der Waals Interactions in Organic Reactions: the Dimerization of Anthracene Revisited. *Angew. Chem. Int. Ed.* **2006**, *45*, 625–629.
- (51) Pitonak, M.; Neogady, P.; Cerny, J.; Grimme, S.; Hobza, P. Scaled MP3 non-covalent interaction energies agree closely with accurate CCSD(T) benchmark data. *ChemPhysChem* **2009**, *10*, 282–289.
- (52) Sedlak, R.; Riley, K. E.; Řezáč, J.; Pitoák, M.; Hobza, P. MP2.5 and MP2.X: approaching CCSD(T) quality description of noncovalent interaction at the cost of a single CCSD iteration. *ChemPhysChem* **2013**, *14*, 698–707.
- (53) Fuchter, M. J.; Weimar, M.; Yang, X.; Judge, D. K.; White, A. J. P. An unusual oxidative rearrangement of [7]-helicene. *Tetrahedron Lett.* **2012**, *53*, 1108.
- (54) Joly, M.; Defay, N.; Martin, R. H.; Declercq, J. P.; Germain, G.; Soubrier-Payen, B.; Van Meerssche, M. Hlicnes ponts: Ethano3,15 et oxa2propano3,15[7]hlicne. Synthse, spectrographie 1HRMN. et tude par diffraction aux rayons X. *Helv. Chim. Acta* **1977**, *60*, 537–560.
- (55) Fujino, S.; Yamaji, M.; Okamoto, H.; Mutai, T.; Yoshikawa, I.; Houjou, H.; Tani, F. Systematic investigations on fused π -system compounds of seven benzene rings prepared by photocyclization of diphenanthrylenes. *Photochem. Photobio. Sci.* **2017**, *16*, 925–934.
- (56) Bludský, O.; Rubeš, M.; Soldán, P.; Nachtigall, P. Investigation of the benzene-dimer potential energy surface: DFT/CCSD(T) correction scheme. *J. Chem. Phys.* **2008**, *128*, 114102.

Graphical TOC Entry

