Local decomposition of imaginary polarizabilities and dispersion coefficients

Ignat Harczuk,[†] Balazs Nagy,[‡] Frank Jensen,[‡] Olav Vahtras,[†] and Hans Ågren[*,†,¶]

† *KTH Royal Institute of Technology, School of Biotechnology, Division of Theoretical Chemistry and Biology, SE-106 91 Stockholm, Sweden*

‡ *Department of Chemistry, Langelandsgade 140, 8000 Aarhus C, Denmark*

¶ *Institute of Nanotechnology, Spectroscopy and Quantum Chemistry, Siberian Federal University, Svobodny pr. 79, 660041 Krasnoyarsk, Russia*

E-mail: hagren@kth.se

Abstract

We present a new way to compute the two-body contribution to the dispersion energy using *ab initio* theory. By combining the complex polarization propagator method and the LoProp transformation, local contributions to the Casimir-Polder interaction is obtained. The full dispersion energy in dimer systems consisting of pairs of molecules including H_2, N_2, CO, CH_4, pyridine, and benzene is investigated, where anisotropic as well as isotropic models of dispersion are obtained using a decomposition scheme for the dipole-dipole polarizability. It is found that the local minima structure of the π-cloud stacking of the benzene dimer is underestimated by the total molecular dispersion, but is alleviated by the inclusion of atomic interactions via the decomposition scheme. The dispersion energy in the T-shaped benzene dimer system is greatly underestimated by all dispersion models, as compared to high-level quantum calculations. The generalization of the decomposition scheme to higher order multipole polarizability interactions, representing higher order dispersion coefficients, is briefly discussed. It is argued that the incorporation of atomic C_6 coefficients in new atomic force fields may have important ramifications in molecular dynamics studies of biomolecular systems.

1 Introduction

Intermolecular forces describe how molecules attract or repel each other, and play a huge role in atomistic simulations. Since it is unfeasible to treat large systems using many-body theory, the partitioning of the total energy into atom-pair potentials has for a long time been essential for the accurate description of complex systems, with the purpose of obtaining statistical averages of properties in protein dynamics, surface adsorption, and other phenomena. The intermolecular forces which govern subtle mechanics in complex systems have thus been in focus, and as available computational power has steadily increased throughout the years, so has the possibility in evaluating energy contributions from *ab initio* theory rather than by empirical fitting. This has motivated the use of bottom-up approaches for force-field design.[1]

The attractive long-range intermolecular forces give rise to potentials which can be partitioned into electrostatic, induction, and dispersion energy contributions. The electrostatic and induction terms are direct results of molecular ground-state properties, i.e. multipole moments and real static polarizabilities, respectively. The effects of dispersion interactions are, in general, very subtle and minuscule at short ranges compared to the other terms, but dominate in systems, which have either zero static multipole moments, e.g., the interactions between noble gases, or the first non-vanishing moment is beyond the dipole moment, e.g., the interaction in a benzene dimer. Most of the additive force-fields used today [2–4] neglect the polarization induction, which nevertheless sometimes can be a crucial term. [5] Furthermore, the dispersion is often modeled by a Leonard-Jones potential, where the parameters are often obtained by fitting to heats of vaporization or solvation energies taken from experiments or calculations.

Early works by the pioneers van der Waals, London, and Buckingham focused on the physical description of intermolecular interactions. Later, the long-range contributions to the dispersion energy could be attributed to the vacuum fluctuation of the quantized field, and quantifiably be predicted by the Casimir-Polder integral. [6,7] Stone and Misquitta have investigated the various intermolecular contributions due to the electrostatic, polarization, and dispersion energy by the means of the SAPT(DFT) (Symmetry Adapted Perturbation Theory — Density Functional Theory) method for a large variety of systems. [8–10] Since the SAPT(DFT) method gives energy contributions of monomer-monomer perturbations to various degrees, the physical origin of each energy contribution of each order can be directly interpreted as of electrostatic, induction or dispersion origin, with the repulsive terms coming from the Pauli exchange interaction coupled to each respective term. The C_n dispersion coefficients between sites 1 and 2 depend on the integral

$$C_n \propto \int_0^\infty \alpha^{(1)}_{tt'}(\mathrm{i}\omega)\alpha^{(2)}_{uu'}(\mathrm{i}\omega)\mathrm{d}\omega \tag{1}$$

where $\alpha^{(1)}_{tt'}(\mathrm{i}\omega)$ is the polarizability for site 1 calculated at an imaginary frequency iω. The lower indices of α are angular momentum labels and the order n in the C_n coefficient is defined as $n = t + t' + u + u' + 2$, giving, e.g., the C_6 term from the linear dipole-dipole polarizability ($t = t' = u = u' = 1$). By calculating the lowest orders of C_n and localizing the contribution to the frequency dependent polarizabilities using the Williams-Stone-Misquitta method,[8,11] the dispersion energy has been investigated for a large number of systems.[12,13]

Norman et al.[14,15] implemented the complex polarization propagator (CPP) approach in order to calculate the first-leading isotropic C_6 dispersion coefficient with the help of analytical response theory. Since the general complex polarizability is a function of complex frequencies, the idea behind the CPP approach is to transform the expression of the polarizability as a function of real frequencies only, and extract the real and imaginary part via the Cauchy moments.[14] The CPP method has been applied to systems such as C_{60} fullerene,[16] Na^+-clusters,[17] and n-alkanes.[18]

In this work, the CPP approach is applied to calculate the Casimir-Polder integral of the isotropic C_6 coefficient for H_2, N_2, CO, methane, pyridine, and benzene. By using the LoProp transformation, we decompose the molecular C_6 coefficients into atomic sites, giving rise to a total dispersion in a dimer system at various relative orientations. The LoProp dispersion is then compared to the native dispersion only calculated between the centers of mass of the dimers. We compare the dispersion energy with the dispersion energy term obtained at the SAPT(DFT) level of theory, which serves as the reference method.

2 Theory

2.1 Anisotropic Dispersion Interaction

Thorough accounts of long-range intermolecular interactions have been given previously, e.g. by Magnasco et al.[19] We will here give the dispersion energy in a form which requires a few definitions in order to make the presentation self-contained. Consider two separated

systems (molecules or atoms) labeled by 1 and 2. They have each a center-of-mass at a position represented by vectors $\mathbf{R}_1$ and $\mathbf{R}_2$, respectively, with Cartesian representations in a common coordinate system. These vectors have Cartesian representations in some common coordinate system: X_1^i and X_2^i for $i = 1, 2, 3$. We refer to the vector connecting the two subsystems by

$$\mathbf{R}_{12} = \mathbf{R}_1 - \mathbf{R}_2 \tag{2}$$

i.e., the distance between the centers of mass is $R_{12} = |\mathbf{R}_{12}|$

We use $\mathbf{r}_1$ for an electronic coordinate associated with system 1 and $\mathbf{r}_2$ for electrons associated with system 2. The density-density response function of system 1 is defined by

$$\langle\langle \hat{\rho}(\mathbf{r}_1); \hat{\rho}(\mathbf{r}_1{}') \rangle\rangle_{\omega}^{(1)} = \sum_{n_1>0} \frac{\langle \Psi_0^{(1)} | \hat{\rho}(\mathbf{r}_1) | \Psi_{n_1}^{(1)} \rangle \langle \Psi_{n_1}^{(1)} | \hat{\rho}(\mathbf{r}_1{}') | \Psi_0^{(1)} \rangle}{\hbar(\omega - \omega_{n_1}^{(1)})} - \frac{\langle \Psi_0^{(1)} | \hat{\rho}(\mathbf{r}_1') | \Psi_{n_1}^{(1)} \rangle \langle \Psi_{n_1}^{(1)} | \hat{\rho}(\mathbf{r}_1) | \Psi_0^{(1)} \rangle}{\hbar(\omega + \omega_{n_1}^{(1)})} \tag{3}$$

where $\Psi_{n_1}^{(1)}$ is a stationary state of system 1 with $n_1 = 0$ being the ground state and $\hbar\omega_{n_1}^{(1)} \equiv E_{n_1}^{(1)} - E_0^{(1)}$ is excitation energy from the ground state of system 1 to its n_1:th excited state. The density operator $\hat{\rho}(\mathbf{r}_1)$ has the property that its expectation value is the electron density at a point in space. It is given in second-quantized form by expanding it in a given basis set on system 1, $\{\phi_p^{(1)}(\mathbf{r}_1)\}_{p=1}^{\infty}$:

$$\hat{\rho}(\mathbf{r}_1) = \sum_{pq} \phi_p^{(1)}(\mathbf{r}_1)^* \phi_q^{(1)}(\mathbf{r}_1) (a_{p\alpha}^\dagger a_{q\alpha} + a_{p\beta}^\dagger a_{q\beta}) \tag{4}$$

where $a_{p\alpha}^\dagger a_{q\alpha}$ is an operator representing an excitation of an α-electron (spin $+\frac{1}{2}$) from orbital q to p and $a_{p\beta}^\dagger a_{q\beta}$ the equivalent for a β-electron (spin $-\frac{1}{2}$).

In particular the dispersion energy of two interacting molecules (denoted by 1 and 2) is given by

$$E_{\text{disp}} = -\frac{\hbar}{2\pi} \iint \mathrm{d}\mathbf{r}_1 \mathrm{d}\mathbf{r}_2 \frac{1}{r_{12}} \iint \mathrm{d}\mathbf{r}_1' \mathrm{d}\mathbf{r}_2' \frac{1}{r_{12}'} \int_0^\infty \mathrm{d}\omega \langle\langle \hat{\rho}(\mathbf{r}_1); \hat{\rho}(\mathbf{r}_1') \rangle\rangle_{\mathrm{i}\omega}^{(1)} \langle\langle \hat{\rho}(\mathbf{r}_2); \hat{\rho}(\mathbf{r}_2') \rangle\rangle_{\mathrm{i}\omega}^{(1)} \tag{5}$$

where the linear response functions are evaluated at an imaginary frequency iω. The intermolecular Coulomb potential can further be expanded in multipole moments, e.g., in solid spherical harmonics. Another alternative is a real-valued Taylor expansion to obtain the Cartesian moments, by expanding around the intermolecular distance vector. Introducing the electronic coordinate vectors relative to the local molecular coordinate frames (see Fig.1) $\mathbf{r}_{1;1} = \mathbf{r}_1 - \mathbf{R}_1$ and $\mathbf{r}_{2;2} = \mathbf{r_2} - \mathbf{R}_2$ we have

$$\mathbf{r}_{12} = \mathbf{r}_{1;1} - \mathbf{r}_{2;2} + \mathbf{R}_{12} \tag{6}$$

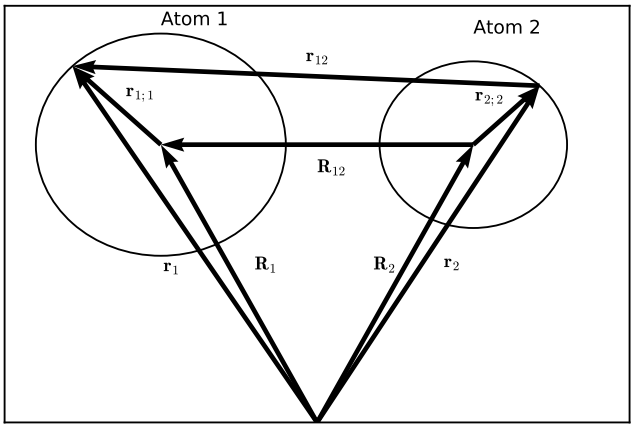


Figure 1: Definition of coordinates in section 2.1

When the first two terms are small compared to the last, as we can assume for long-range intermolecular interactions, we may expand the Coulomb potential in a Taylor series

$$\frac{1}{r_{12}} = \frac{1}{|\mathbf{r}_{1;1} - \mathbf{r}_{2;2} + \mathbf{R}_{12}|} = \sum_{n=0}^{\infty} \frac{1}{n!} ((\mathbf{r}_{1;1} - \mathbf{r}_{2;2}) \cdot \nabla_{R_{12}})^n \frac{1}{R_{12}} \tag{7}$$

where the derivative operator only acts on the internuclear distance

$$\nabla_{R_{12}} = \left(\frac{\partial}{\partial X_{12}^1}, \frac{\partial}{\partial X_{12}^2}, \frac{\partial}{\partial X_{12}^3} \right) \tag{8}$$

Furthermore we can use the binomial theorem to obtain

$$((\mathbf{r}_{1;1} - \mathbf{r}_{2;2}) \cdot \nabla_{R_{12}})^n = \sum_{k=0}^{n} \binom{n}{k} (\mathbf{r}_{1;1} \cdot \nabla_{R_{12}})^k (-\mathbf{r}_{2;2} \cdot \nabla_{R_{12}})^{n-k} \tag{9}$$

Substituting expansions in Eqs. 7 and 9 in the linear response function in Eq. 3, we can rewrite Eq. 5 as

$$\begin{aligned} E_{\text{disp}} = -\frac{\hbar}{2\pi} \sum_{nk} \sum_{n'k'} \frac{1}{n!n'!} \binom{n}{k} \binom{n'}{k'} (-1)^{n-k+n'-k'} \times \\ \int_0^\infty \mathrm{d}\omega \langle\!\langle \int d\mathrm{r}_1 \hat{\rho}(\mathbf{r}_1) (\mathbf{r}_{1;1} \cdot \nabla_{R_{12}})^k ; \int d\mathrm{r}_2 \hat{\rho}(\mathbf{r}'_1) (\mathbf{r}'_{1;1} \cdot \nabla_{R'_{12}})^{k'} \rangle\!\rangle_{\mathrm{i}\omega} \times \\ \langle\!\langle \int d\mathrm{r}'_1 \hat{\rho}(\mathbf{r}_2) (\mathbf{r}_{2;2} \cdot \nabla_{R_{12}})^{n-k} ; \int d\mathrm{r}'_2 \hat{\rho}(\mathbf{r}'_2) (\mathbf{r}'_{2;2} \cdot \nabla_{R'_{12}})^{n'-k'} \rangle\!\rangle_{\mathrm{i}\omega} \left. \frac{1}{R_{12}} \frac{1}{R'_{12}} \right|_{R_{12}=R'_{12}} \end{aligned} \tag{10}$$

where we can identify linear response functions of second-quantized formulations of multipole moments of various orders.

The lowest non-vanishing term is with $k = k' = 1$ and $n = n' = 2$, where we have electrical dipole operators

$$\hat{\mathbf{r}} = \int \mathrm{d}\mathbf{r} \hat{\rho}(\mathbf{r}) \mathbf{r} \tag{11}$$

and we can identify the well-known expression for dipole-dipole polarizabilities

$$\alpha(\mathrm{i}\omega) = -\langle\!\langle \hat{\mathbf{r}}; \hat{\mathbf{r}} \rangle\!\rangle_{\mathrm{i}\omega} \tag{12}$$

This expansion for a specific orientation of the molecules is an expression which couples

the molecular polarizability tensors with second derivatives of the intermolecular distance and which decays as R_{12}^{-6}

$$E_{\text{disp}} \approx -\frac{\hbar}{2\pi} \int_0^\infty \mathrm{d}\omega \sum_{ijkl} \alpha_{ij}^{(1)}(\mathrm{i}\omega)\alpha_{kl}^{(2)}(\mathrm{i}\omega)T_{ik}T_{jl} \tag{13}$$

where the summation is over cartesian coordinates in a common coordinate system, and T is the second-order derivative of the Coulomb potential

$$T_{ik} = \frac{\partial^2}{\partial X_{12}^i \partial X_{12}^k}\frac{1}{R_{12}} = \frac{3X_{12}^i X_{12}^k - \delta_{ik}R_{12}^2}{R_{12}^5} \tag{14}$$

2.2 Isotropic dispersion coefficient C_6

An isotropic expression for the dispersion energy is obtained by rotational averaging of the two subsystems independently. Molecular tensors such as polarizabilities are usually calculated in a coordinate system fixed by molecular nuclei (body-fix axes) whereas observations are carried out in a coordinate system defined by the observer (lab-fix axes). With $C_{\mu i}$ as the transformation matrix between body-fix (greek indices) and space-fix (roman indices) axes this averaging may be written

$$E_{\text{disp}}^{\text{iso}} \approx -\frac{\hbar}{2\pi} \int_0^\infty \mathrm{d}\omega \sum_{ijkl} \sum_{\mu\nu\sigma\tau} \alpha_{\mu\nu}^{(1)}(\mathrm{i}\omega)\alpha_{\sigma\tau}^{(2)}(\mathrm{i}\omega)\overline{C_{\mu i}C_{\nu j}} \quad \overline{C_{\sigma k}C_{\tau l}}T_{ik}T_{jl} \tag{15}$$

Detailed expressions for rotational averages of tensors can be found, e.g., in Ref. 20; in our case the average can be written as

$$\overline{C_{\mu i}C_{\nu j}} = \frac{1}{3}\delta_{\mu\nu}\delta_{ij} \tag{16}$$

Substituting with the isotropic polarizabilities

$$\bar{\alpha} = \frac{1}{3}\sum_k \alpha_{kk} \tag{17}$$

we obtain

$$\begin{aligned} E_{\text{disp}}^{\text{iso}}(R_{12}) &= -\frac{\hbar}{2\pi}\int_0^\infty \overline{\alpha}^{(1)}(\mathrm{i}\omega)\overline{\alpha}^{(2)}(\mathrm{i}\omega)\sum_{ik} T_{ik}T_{ik} \\ &= -\frac{\hbar}{2\pi}\int_0^\infty \overline{\alpha}^{(1)}(\mathrm{i}\omega)\overline{\alpha}^{(2)}(\mathrm{i}\omega)\frac{6}{R_{12}^6} \\ &= -\frac{3\hbar}{\pi}\int_0^\infty \bar{\alpha}^{(1)}(\mathrm{i}\omega)\bar{\alpha}^{(2)}(\mathrm{i}\omega)\mathrm{d}\omega\frac{1}{R_{12}^6} \\ &\equiv -\frac{C_6}{R_{12}^6} \end{aligned} \tag{18}$$

which defines the well-known C_6 coefficient for the dispersion interaction. By decomposing the molecular polarizabilities $\alpha^{(1)}$ and $\alpha^{(2)}$ into distributed atomic contributions $\alpha_i^{(1)}$ and $\alpha_j^{(2)}$, respectively, where i is an atomic site in molecule 1 and j an atomic site in molecule 2, the atom-atom dispersion coefficient C_6^{ij} for atoms i and j can be written as

$$C_6^{ij} = \frac{3\hbar}{\pi}\int_0^\infty \bar{\alpha}_i^{(1)}(\mathrm{i}\omega)\bar{\alpha}_j^{(2)}(\mathrm{i}\omega)\mathrm{d}\omega \tag{19}$$

where now the distributed polarizabilities for the atoms contributes to the atomic C_6^{ij} dispersion. Note that in order to calculate the total dispersion energy by our scheme LoProp (see below), each atom-atom pair has to be evaluated. For a water dimer with 3 atoms in each molecule, there is thus $3 \times 3 = 9$ atom-pairs to consider. This is due to the LoProp polarizability being additive and summing up to the molecular polarizability, and the sum of all LoProp C_6^{ij} elements being equal to the molecular C_6.

The LoProp approach[21] to obtain frequency dependent polarizabilities is based on diagonalizing the atomic overlap matrix S. This is done in sequential orthonormalization steps which gives a transformation matrix used to transform the property integrals and perturbed electronic densities. By summing the transformed integrals over the subspace of atom-pairs, the properties in the LoProp basis are obtained. In previous works,[22,23] it was shown how the frequency-dependent polarizability and hyperpolarizability could be obtained from linear response theory. Here, in this work, the natural extension to the extraction of the response

vectors from the CPP code is made in order to obtain the distributed real polarizabilities of atoms at imaginary frequencies.

3 Computational Details

All geometries except for the diatomics were optimized at the B3LYP/6–31++G** level of theory using Gaussian09.[24] For the H_2, N_2, and CO molecules the experimental bond lengths of 0.7414,[25] 1.0980,[25] and 1.1280[26] Å were used, respectively. All structures in Cartesian coordinates are shown in Tables S1-S9 in the Supporting Information.

The polarizabilities at a set of six imaginary frequencies were obtained at the TDHF level of theory with the aug-cc-pVTZ[27] basis set. For these calculations the CPP method[14,15] was invoked as implemented in DALTON.[28] The infinite integral of the polarizabilities was calculated numerically relying on a 6-point Gauss-Legendre integration scheme with the variable transformation

$$\omega = \omega_0 \frac{1-t}{1+t}, \tag{20}$$

where the constant ω_0 was set to 0.3 E_h according to Ref. 17. The atomic decomposition of the molecular polarizabilities was carried out with the LoProp script[29] interfaced to DALTON.

The accuracy of the 6-point Gauss-Legendre integration formula has been verified in Ref. 16 for benzene, where integrations of orders 4, 6, 8, and 10 have been performed to check the convergence of the molecular C_6 dispersion coefficient. The 6-point result was concluded to be fairly converged, therefore this was the preferred order throughout the study in Ref. 16. We have also evaluated the numerical integrals for the molecular C_6 coefficient of benzene using a 6-point and a 10-point integration scheme. The resulting coefficients were 1761.80 and 1761.77 a.u., respectively, with a negligible difference of 0.03 a.u. Based on this, the 6-point formula was also accepted in our calculations.

The SAPT2+[30] dispersion interaction energies for dimers of H_2, N_2, and CO were obtained with PSI4[31] at the density-fitted Hartree-Fock (DF-HF) level of theory using the

aug-cc-pVTZ basis set and its auxiliary equivalents aug-cc-pVTZ-JKFIT and aug-cc-pVTZ-RI for the HF and SAPT calculations, respectively.

4 Results and Discussion

We first present the basis set convergence of C_6 dispersion coefficients obtained with the isotropic molecular and LoProp decomposition models for formaldehyde, formamide, and glycine molecules. Then we discuss the dispersion energy for dimers of diatomic molecules H_2, N_2, and CO obtained with the anisotropic models. The results are compared with reference data calculated with the SAPT2+ method. Dispersion interaction energies for dimers of benzene, methane, and pyridine molecules at a number of random orientations calculated within the molecular isotropic and anisotropic as well as the LoProp isotropic and anisotropic frameworks are also presented and compared to SAPT(DFT) data.

4.1 Basis set dependence of C_6 coefficients

Figure 2 shows the convergence of the isotropic atomic and summed molecular C_6 coefficients as a function of the basis set used to obtain the polarizabilities. The molecular values show a fast convergence with the basis set size, the results are essentially converged with the aug-cc-pVDZ set. It is well-known that the augmentation of basis sets with diffuse functions is essential when calculating higher-order molecular properties, e.g., hyperpolarizabilities.[32] It can be seen in Figure 2 that the augmentation is also necessary for dispersion coefficients and consequently, polarizabilities. There is no apparent benefit of using a basis set beyond aug-cc-pVTZ, therefore this is the basis set of choice in all further calculations in this work.

Considering the LoProp decomposed atomic C_6 coefficients in Figure 2, it is interesting to note that in contrast to the well-converged molecular results, these values show some particular trends as the basis set increases. The localized hydrogen and carbon polarizabilities are enhanced by including more polarization functions, while the oxygen and nitrogen

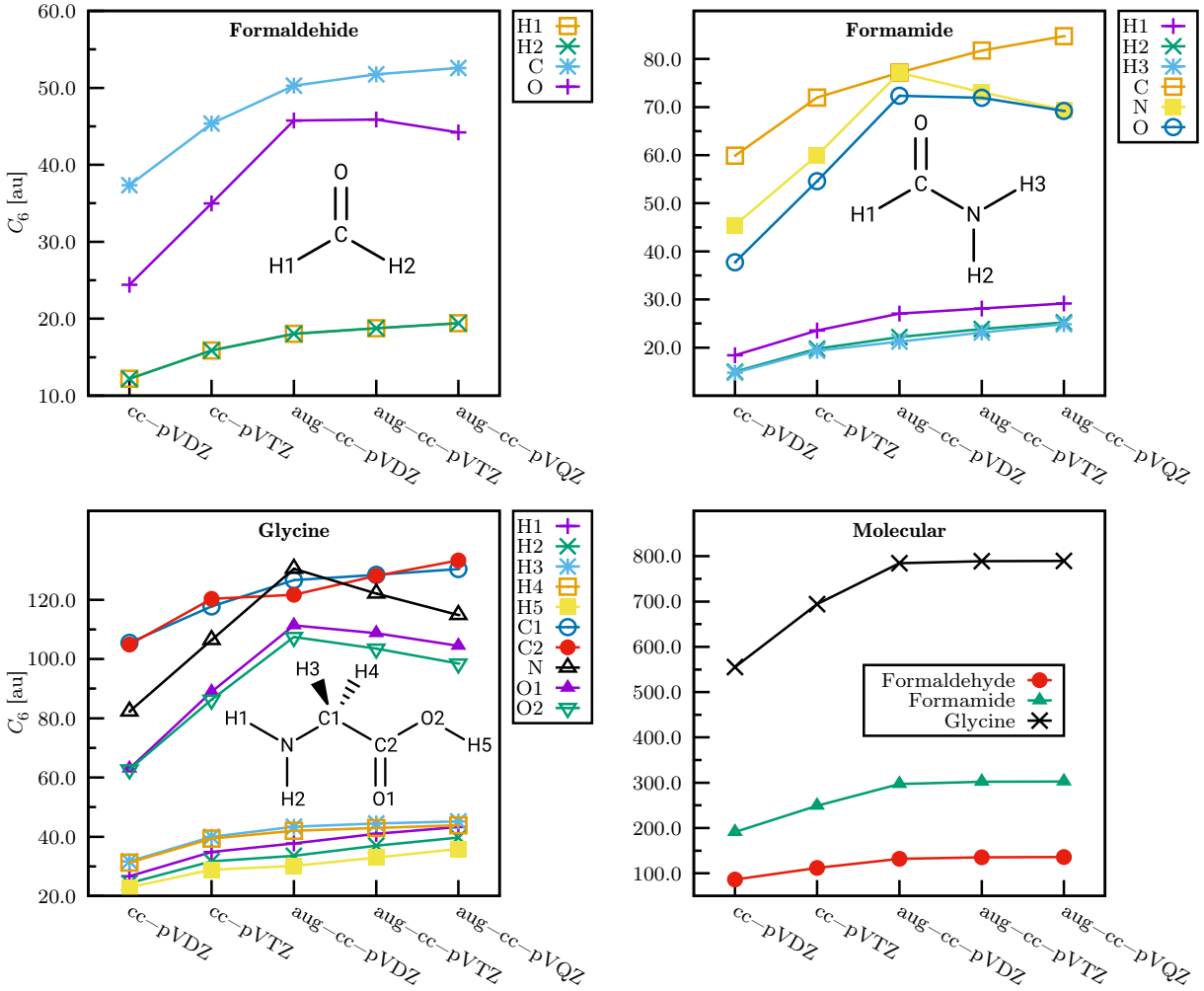


Figure 2: Basis set convergence of LoProp decomposed atomic and total molecular C_6 coefficients

polarizabilities are reduced in importance.

4.2 Diatomic systems H_2, N_2, and CO

We investigate four particular relative orientations for dimers H_2-H_2, N_2-N_2, and CO-CO, as shown in Figure 3. In each orientation, the first molecule is placed along the x-axis with the bond midpoint in the origin. The second molecule is then defined by a rotation and a translation relative to the first molecule. The perpendicular orientation is defined by a 90

degree rotation about the z-axis and a translation along the y-axis. In the parallel dimer, one monomer molecule is translated along the y-axis relative to the reference molecule, without any rotation. The shifted dimer is a parallel dimer with a given separation along the y-axis, and one of the monomers is translated along the x-axis by keeping the other monomer fixed. The parallel rotated dimer is defined by a 90 degree rotation about the y-axis and a translation along this axis.

Results for dispersion interaction energies calculated with the anisotropic models along with the SAPT2+ reference values are presented in Figures 4, 5, and 6 for the different orientations of dimers H_2-H_2, N_2-N_2, and CO-CO, respectively. Although our dispersion models agree well with the reference SAPT2+ results for the H_2 dimer, the alignment is not as good either for the N_2 or the CO dimers. For the latter two dimers our anisotropic models seem to slightly underestimate the magnitude of the dispersion interaction energy at each relative orientation. This trend is clearly seen in the scatter plots for the diatomics in the Supplementary Material.

It can also be seen from Figures 5a and 6a that the LoProp anisotropic dispersion gives larger contributions to the dispersion energy at closer distances for the perpendicular configuration, as compared to the parallel configurations, where the molecular model fares best. This can be shown analytically based on the fact that the total dispersion energy of the LoProp model is simply the sum of the individual atom-atom terms. This sum for the parallel orientations will never be as high as the value calculated with the molecular anisotropic model because the interatomic distances weighted with their local polarizabilities will always collectively add up to a smaller value than the COM distances weighted with the total molecular polarizabilities as defined in the molecular model.

Considering the results for the parallel shifted orientations for N_2 and CO dimers in Figures 5b and 6b, the underestimation of the dispersion interaction energy can be clearly seen. The underestimation is the largest for the LoProp model at short intermolecular separations and small shifts of the molecules in the dimers. However, at larger separations

and shifts this model gives better overall agreement with the SAPT2+ results than the molecular model.

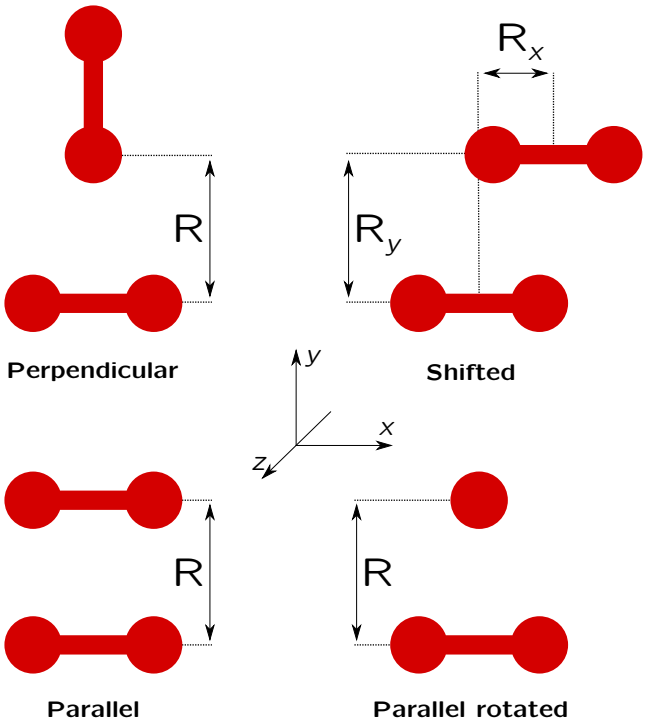
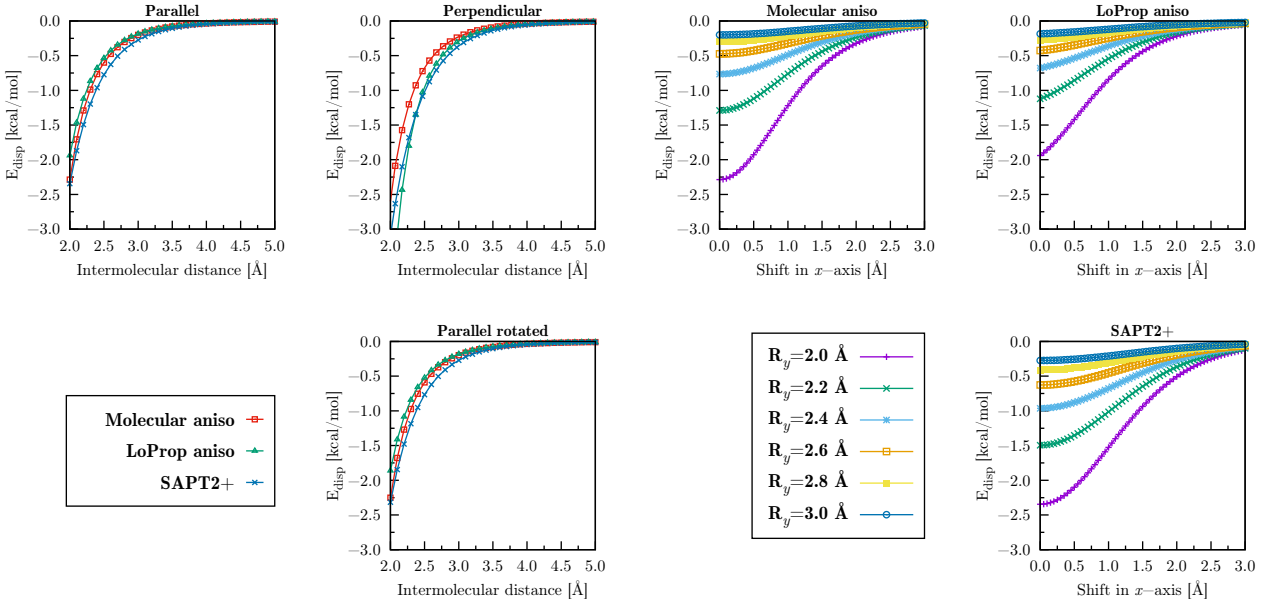


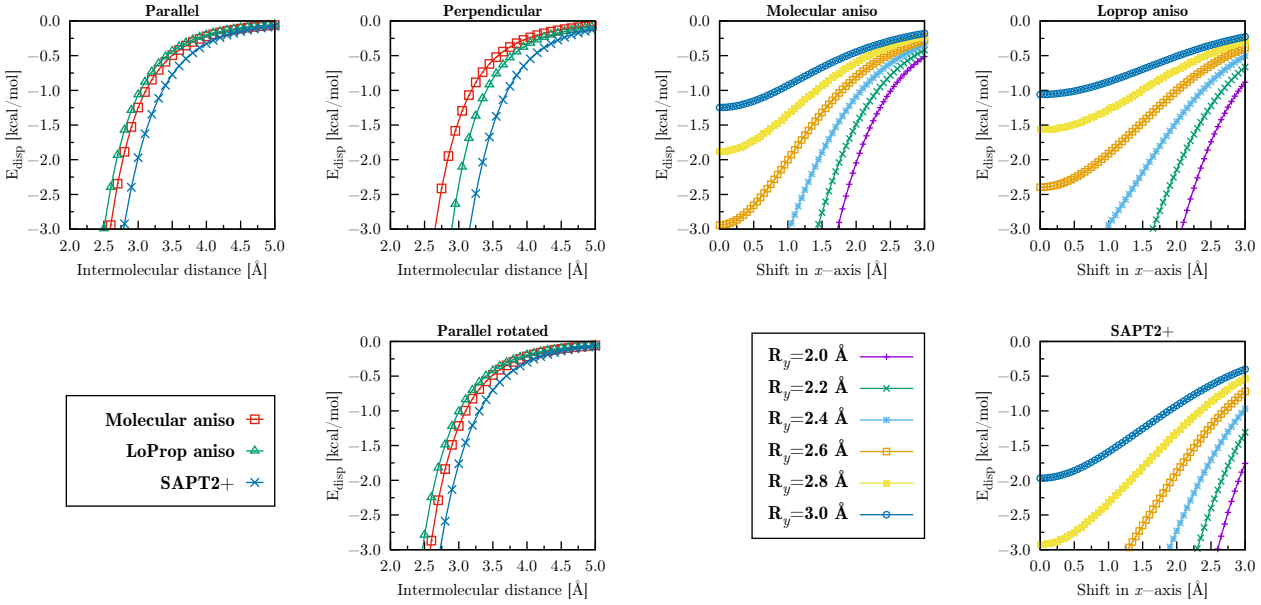
Figure 3: Schematic representation of the relative orientations for the investigated diatomic molecular dimer systems H_2–H_2, N_2–N_2, and CO–CO.



(a) Perpendicular, parallel, and parallel rotated orientations

(b) Parallel shifted orientations

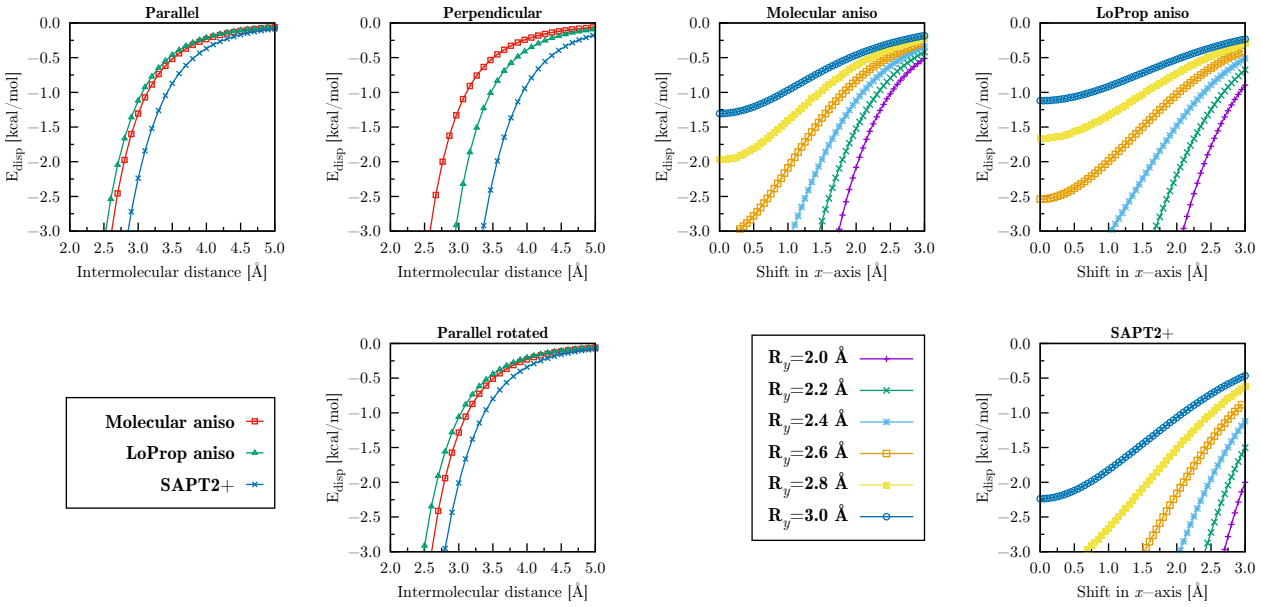
Figure 4: Comparison of SAPT2+ results with the anisotropic dispersion energy models computed for the hydrogen dimer. See Figure 3 for coordinate system definition.



(a) Perpendicular, parallel, and parallel rotated orientations

(b) Parallel shifted orientations

Figure 5: Comparison of SAPT2+ results with the anisotropic dispersion energy models computed for the nitrogen dimer. See Figure 3 for coordinate system definition.



(a) Perpendicular, parallel, and parallel rotated orientations

(b) Parallel shifted orientations

Figure 6: Comparison of SAPT2+ results with the anisotropic dispersion energy models computed for the CO dimer. See Figure 3 for coordinate system definition.

4.3 Benzene dimer

The C_6 molecular dispersion coefficient of benzene has previously been calculated[16] with the CPP method at the TDHF and TDDFT levels of theory. The TDHF calculation with the 6-31G** basis set resulted in 1737 a.u. In the TDDFT calculation the B3LYP functional together with the 6-31G** basis set was invoked and 1773 a.u. was obtained for the C_6 coefficient. It was concluded that the TDHF dispersion was better than the TDDFT one due to the better static polarizability obtained at the TDHF level. Using experimental oscillator strengths and molar refractivity with the DOSD[33] method, the isotropic coefficient for benzene was calculated to 1723 a.u. As compared to the above results, our calculated value of 1762 a.u. for the benzene C_6 coefficient obtained at the TDHF/aug-cc-pVTZ level is likely a minor overestimation of the actual real quantity. However, it serves as a good

qualitative measure and confirms the adequacy of the model proposed in this paper.

In Figure 7, dispersion energy is shown for three specific dimer systems of benzene, labeled as M1, S3, and S8 in Ref 34. Each orientation corresponds to a local minimum on the dimer potential energy surface, and reference dispersion energies were calculated with the SAPT(DFT) method.[34] For the tilted dimer structure M1, the potential is in quite good agreement with the molecular model. For the T-shaped configuration S3 in Figure 7, both the LoProp and molecular potential gives very low dispersion energies. For SAPT(DFT), already at intermolecular separation $r > 4.5$ Å, the energy is around -10 kcal/mol, whereas the isotropic C_6 dispersion predicts only -7.5, and -3.0 kcal/mol for the LoProp, and molecular model, respectively. By stacking the dimer to the S8 configuration, however, the energy is higher for the SAPT(DFT) method, and lies in between the LoProp and molecular models, but agrees better with the LoProp model at larger distances. It can be noted that for systems having horizontal delocalized plane–plane interactions, the anisotropic LoProp model gives a much larger contribution to the dispersion energy, as compared to the other models. This can be seen in Figure 8 for a planar dimer denoted arbitrarily as Z1, where one benzene molecule in the xz-plane with COM in the origin is shifted along the x-axis by keeping the other monomer fixed. There is thus merit in considering the LoProp dispersion model for systems such as DNA, in which a significant portion of the structure stabilization occurs by the analogous S8 and Z1 structures between the nucleic base pairs.

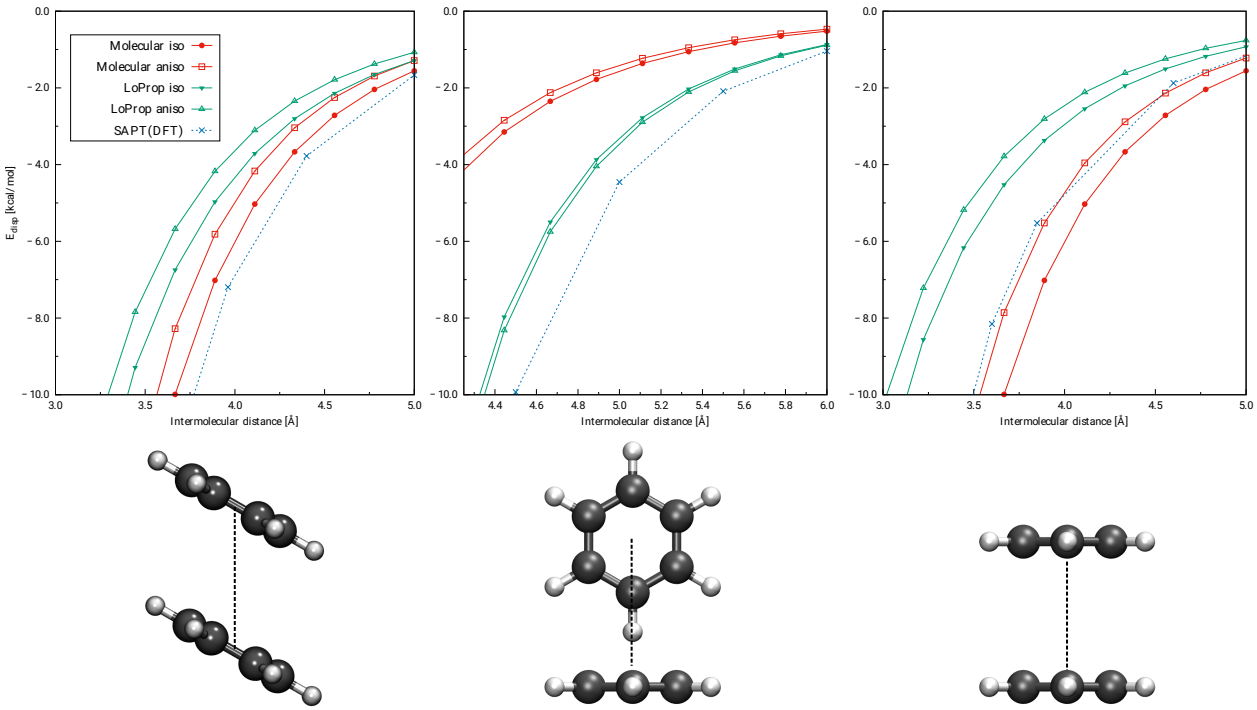


Figure 7: Dispersion energies for benzene dimers M1, S3, and S8, from left to right, respectively. SAPT(DFT) reference points were taken from Ref 34.

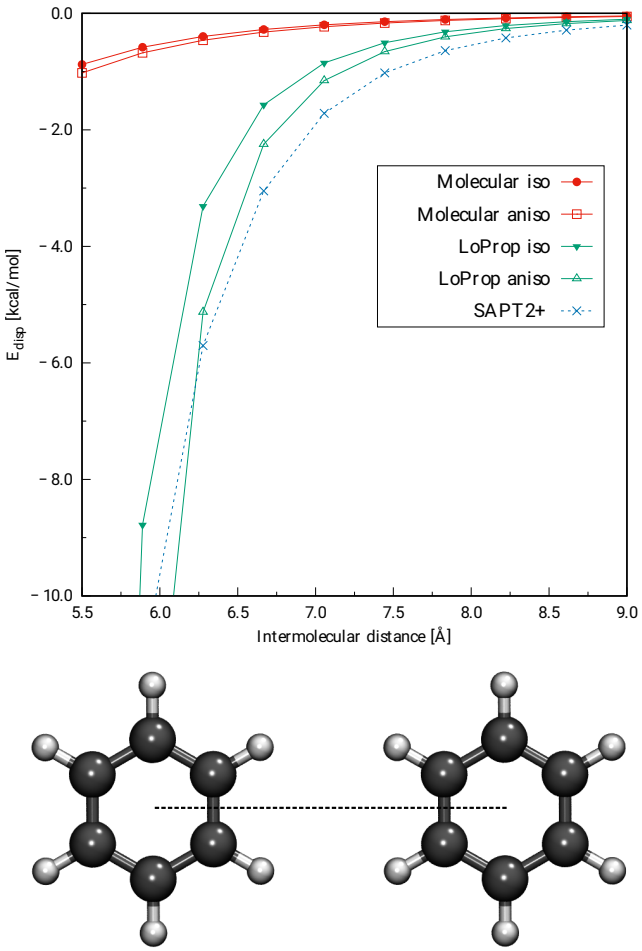


Figure 8: Dispersion energy for the planar (Z1) benzene dimer. Reference energies were calculated with the SAPT2+ method.

5 Methane and pyridine

In figures 9 and 10 and figures S7 and S8 in the Supporting Information we also compare our isotropic and anisotropic models for the methane and pyridine dimers against SAPT(DFT) reference data. The reference data set contained nearly 2500 random dimer structures and the corresponding SAPT(DFT) dispersion energies for methane, while nearly 500 structure–energy pairs for the pyridine-pyridine dimers. Molecular isotropic and anisotropic as well as LoPop isotropic and anisotropic dispersion interaction energies were calculated for these ran-

dom dimers. In figures 9 and 10 scatter plots of our results against the reference SAPT(DFT) dispersion interaction energies are shown. In figures S7 and S8 in the Supporting Information our calculated energies along with the reference data is plotted as a function of the monomer-monomer separation. Similar conclusions can be drawn as for diatomics and benzene. The qualitative agreement between our models and the reference SAPT(DFT) energies is, in general, reasonable. However, in case of the molecular models the systematic overshoot of the dispersion interaction can also be observed here, similarly to the diatomics. Based on figures S7 and S8 it is clearly seen that both the isotropic and the anisotropic molecular models fail to adequately describe the anisotropy of the dispersion interaction giving essentially the same energies. For the pyridine dimer, a slight improvement of the anisotropic molecular model can be seen as compared to the isotropic one. The LoProp models give more accurate, however, less precise results, and the anisotropy is much better in this case as compared to the molecular models. The difference between the isotropic and anisotropic LoProp models is slightly larger than that of the molecular models. The isotropic LoProp model seems to provide the best results as compared to the reference SAPT(DFT) results.

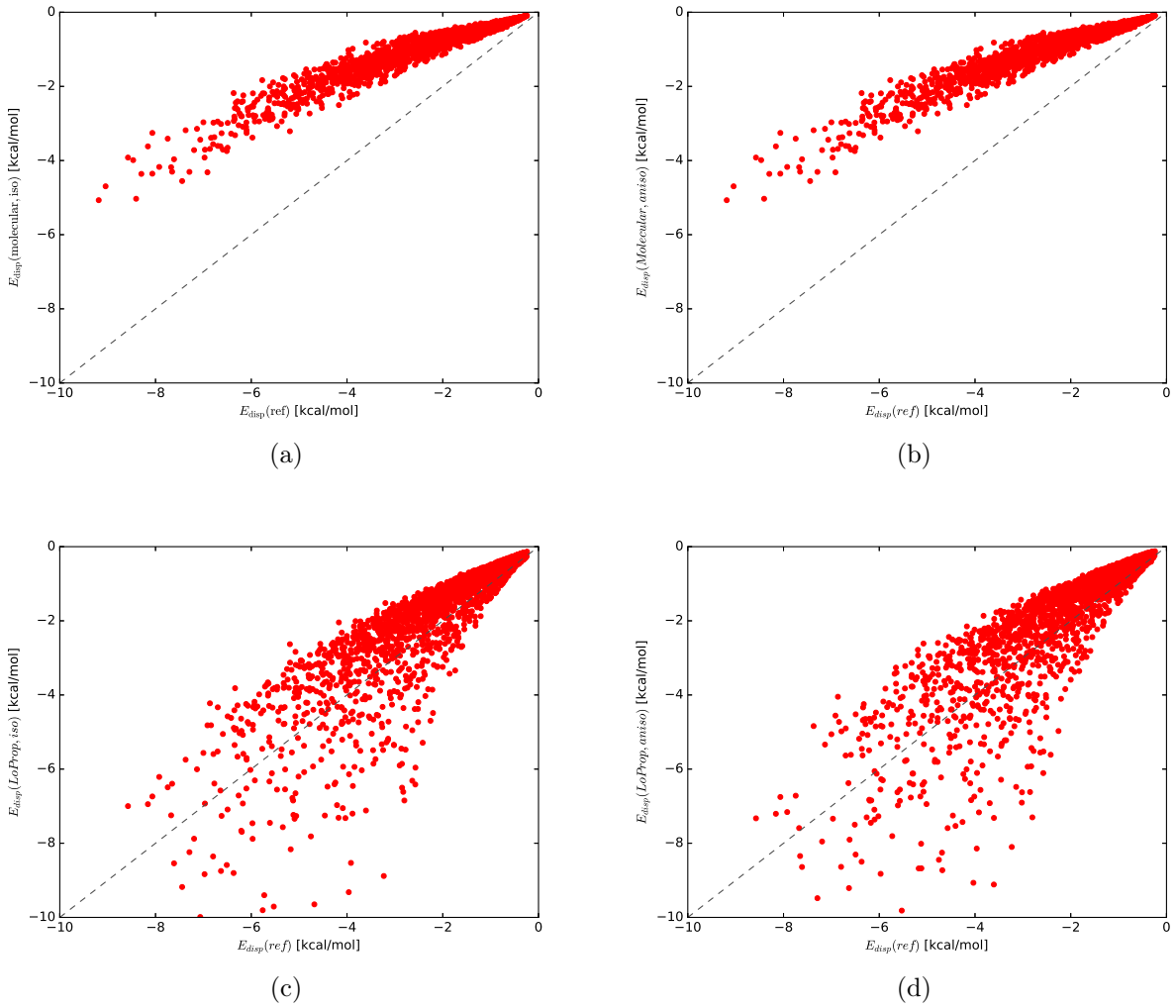


Figure 9: Performance of molecular isotropic (a), molecular anisotropic (b), LoProp isotropic (c), and LoProp anisotropic (d) dispersion interaction energy models for the CH_4 dimer against reference SAPT(DFT) energy.

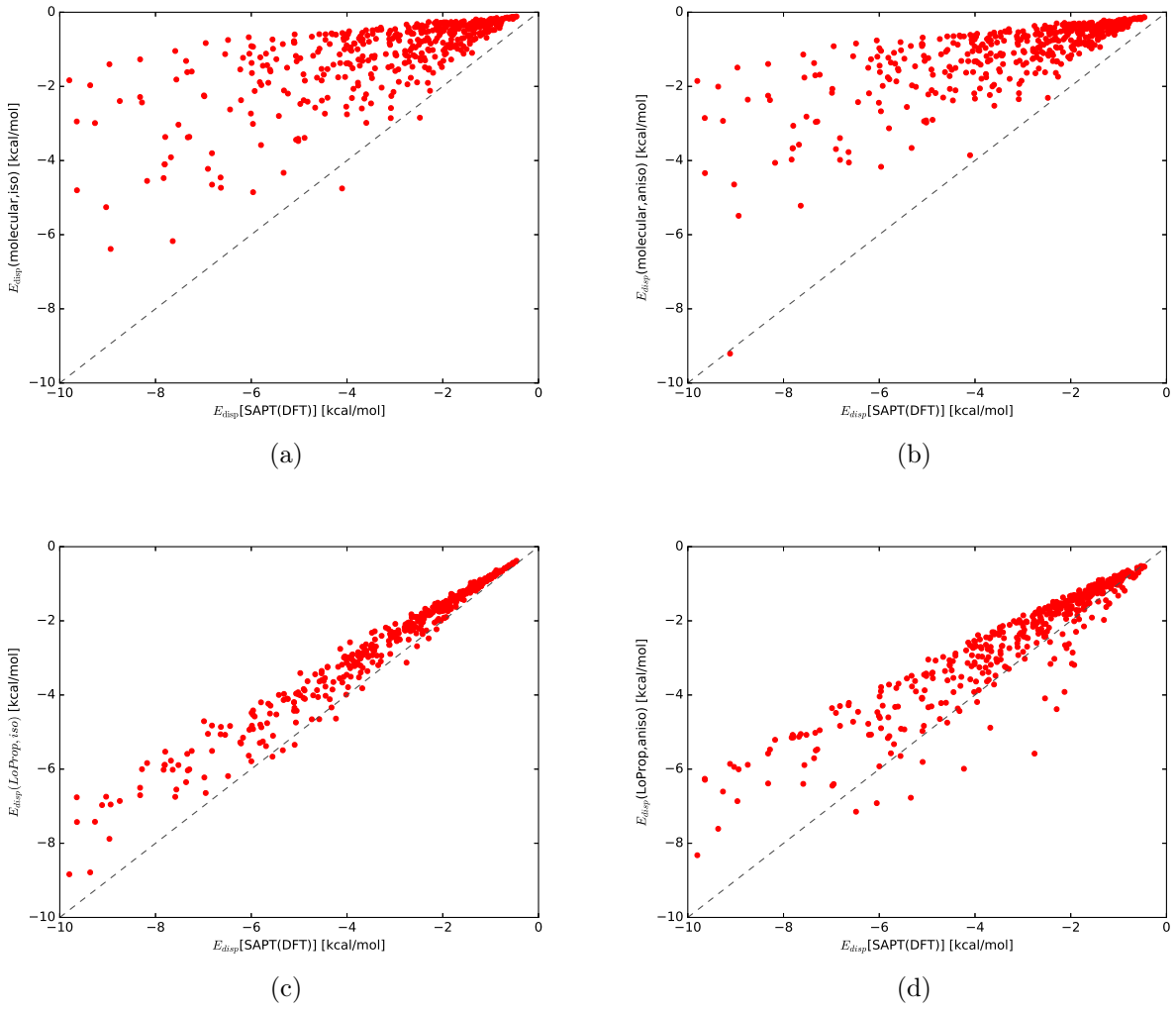


Figure 10: Performance of molecular isotropic (a), molecular anisotropic (b), LoProp isotropic (c), and LoProp anisotropic (d) dispersion interaction energy models for the pyridine dimer against reference SAPT(DFT) energy.

6 Conclusions

We have presented a novel technique to calculate the atomic contributions to the molecular real polarizability evaluated at imaginary frequencies. This was accomplished by using the LoProp transformation of the complex molecular polarizability calculated using the complex polarization propagator. This allows for the construction of the dispersion energy between all-atom sites for the calculation of the total dispersion energy between two molecules. We

have investigated this intermolecular dispersion energy for a wide range of relative orientations for the simple diatomic H_2/H_2 and N_2/N_2 dimer systems. We also computed the anisotropic, in addition to the more simpler isotropic dispersion energy for the benzene dimer, where high level (SAPT(DFT)) values were used for reference. It was found that the simple description of the intermolecular dispersion, using only dipole-dipole polarizabilities, is complete only for the smallest H_2 system. For the larger N_2 and benzene dimer systems, some specific relative orientations, such as the perpendicular configuration in N_2 and the T-shaped S3 configuration in benzene, are not well described by the first order dipole-dipole term, which leads to an underestimation of the dispersion energy. The LoProp decomposition was, however, found to give quite good improvement for the perpendicular N_2 dimer configuration. For benzene, where the largest discrepancy for the dispersion models was found in the T-shaped S3 configuration, the anisotropic LoProp model was found to increase the dispersion energy.

The presented exploration of the LoProp decomposition of the dispersion energy presents applications to novel *ab initio* parameterizations of force-fields. As the dispersion energy contribution to the two-body potential is significant for large systems, combined with the fact that the LoProp transformation gives transferable properties between similar groups between different systems, these results could prove to be useful for accurate fitting of new force-fields with unprecedented accuracy, and in particular to biological DNA/protein systems. Our decomposition scheme lends itself rather straightforwardly for a generalization that includes multipole interactions higher than the dipole-dipole interaction. This will be the topic for a following-up study.

Acknowledgement

The authors gratefully acknowledge Professor Alston J. Misquitta for providing the reference SAPT(DFT) data for the methane and pyridine dimers. The simulations were performed on

resources provided by the Swedish National Infrastructure for Computing (SNIC) at National Super-computing Center (NSC) and High Performance Computing Centre North (HPC2N), project "Design of Force Fields for Theoretical Spectroscopy", SNIC-2015–1–230.

Supporting Information Available

This material is available free of charge via the Internet at `http://pubs.acs.org/`.

References

(1) Mas, E. M.; Bukowski, R.; Szalewicz, K.; Groenenboom, G. C.; Wormer, P. E. S.; van der Avoird, A. Water pair potential of near spectroscopic accuracy. I. Analysis of potential surface and virial coefficients. *J. Chem. Phys.* **2000**, *113*, 6687–6701.

(2) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; Mackerell Jr., A. D. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **2010**, *31*, 671–690.

(3) Case, D.; Darden, T.; Cheatham, T.; III,; Simmerling, C.; Wang, J.; Duke, R.; Luo, R.; Walker, R.; Zhang, W.; Merz, K.; Roberts, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Swails, J.; Goetz, A.; Kolossv??ry, I.; Wong, K.; Paesani, F.; Vanicek, J.; Wolf, R.; Liu, J.; Wu, X.; Brozell, S.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M.-J.; Cui, G.; Roe, D.; Mathews, D.; Seetin, M.; Salomon-Ferrer, R.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. AMBER 12, University of California, San Francisco. **2012**,

(4) Jorgensen, W. L.; Tirado-Rives, J. The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.* **1988**, *110*, 1657–1666.

(5) Piquemal, J.-P.; Perera, L.; Cisneros, G. A.; Ren, P.; Pedersen, L. G.; Darden, T. A. Towards accurate solvation dynamics of divalent cations in water using the polarizable amoeba force field: From energetics to structure. *J. Chem. Phys.* **2006**, *125*, 054511.

(6) Casimir, H.; Polder, D. The influence of retardation on the London-van der Waals forces. *Phys. Rev.* **1948**, *73*, 360.

(7) Power, E.; Thirunamachandran, T. Casimir-Polder potential as an interaction between induced dipoles. *Phys. Rev. A* **1993**, *48*, 4761.

(8) Stone, A. J.; Misquitta, A. J. Atom–atom potentials from ab initio calculations. *Int. Rev. Phys. Chem.* **2007**, *26*, 193–222.

(9) Khaliullin, R. Z.; Cobar, E. A.; Lochan, R. C.; Bell, A. T.; Head-Gordon, M. Unravelling the origin of intermolecular interactions using absolutely localized molecular orbitals. *J. Phys. Chem. A* **2007**, *111*, 8753–8765.

(10) Azar, R. J.; Head-Gordon, M. An energy decomposition analysis for intermolecular interactions from an absolutely localized molecular orbital reference at the coupled-cluster singles and doubles level. *J. Chem. Phys.* **2012**, *136*, 024103.

(11) Misquitta, A. J.; Stone, A. J. Distributed polarizabilities obtained using a constrained density-fitting algorithm. *J. Chem. Phys.* **2006**, *124*, 024111.

(12) Misquitta, A. J.; Podeszwa, R.; Jeziorski, B.; Szalewicz, K. Intermolecular potentials based on symmetry-adapted perturbation theory with dispersion energies from time-dependent density-functional calculations. *J. Chem. Phys.* **2005**, *123*, 214103.

(13) Misquitta, A. J.; Stone, A. J. Dispersion energies for small organic molecules: first row atoms. *Mol. Phys.* **2008**, *106*, 1631–1643.

(14) Norman, P.; Bishop, D. M.; Jensen, H. J. Aa.; Oddershede, J. Near-resonant absorption

in the time-dependent self-consistent field and multiconfigurational self-consistent field approximations. *J. Chem. Phys.* **2001**, *115*, 10323–10334.

(15) Norman, P.; Bishop, D. M.; Jensen, H. J. Aa.; Oddershede, J. Nonlinear response theory with relaxation: The first-order hyperpolarizability. *J. Chem. Phys.* **2005**, *123*, 194103.

(16) Jiemchooroj, A.; Norman, P.; Sernelius, B. E. Complex polarization propagator method for calculation of dispersion coefficients of extended π-conjugated systems: the C_6 coefficients of polyacenes and C_{60}. *J. Chem. Phys.* **2005**, *123*, 124312.

(17) Jiemchooroj, A.; Norman, P.; Sernelius, B. E. Electric dipole polarizabilities and C_6 dipole-dipole dispersion coefficients for sodium clusters and C_{60}. *J. Chem. Phys.* **2006**, *125*, 124306.

(18) Norman, P.; Jiemchooroj, A.; Sernelius, B. E. Polarization propagator calculations of the polarizability tensor at imaginary frequencies and long-range interactions for the noble gases and n-alkanes. *J. Chem. Phys.* **2003**, *118*, 9167–9174.

(19) Magnasco, V.; Figari, G.; Costa, C. Understanding van der Waals bonding. *J. Mol. Struct.* (THEOCHEM) **1992**, *261*, 237–253.

(20) Salam, A. *Molecular Quantum Electrodynamics*; John Wiley & Sons, Inc., 2009; pp 378–380.

(21) Gagliardi, L.; Lindh, R.; Karlström, G. Local properties of quantum chemical systems: the LoProp approach. *J. Chem. Phys.* **2004**, *121*, 4494–500.

(22) Harczuk, I.; Vahtras, O.; Ågren, H. Frequency-dependent force fields for QMMM calculations. *Phys. Chem. Chem. Phys.* **2015**, *17*, 7800–7812.

(23) Harczuk, I.; Vahtras, O.; Ågren, H. Hyperpolarizabilities of extended molecular mechanical systems. *Phys. Chem. Chem. Phys.* **2016**, *18*, 8710–8722.

(24) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A.; Jr.,; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian 09, REvision C.1. Gaussian, inc., Wallingford CT 2009.

(25) Huber, K.; Herzberg, G. *Molecular Spectra and Molecular Structure IV. Constants of Diatomic Molecules*; Van Nostrand Reinhold Co., 1979.

(26) Lovas, F.; Tiemann, E.; Coursey, J.; Kotochigova, S.; Chang, J.; Olsen, K.; Dragoset, R. *NIST Diatomic Spectral Database* (version 2.1), [Online]. Available: http://physics.nist.gov/Diatomic (accessed November 2016). National Institute of Standards and Technology, Gaithersburg, MD., 2005.

(27) Kendall, R. A.; Dunning Jr., T. H.; Harrison, R. J. Electron Affinities of the First-Row Atoms Revisited. Systematic Basis Sets and Wave Functions. *J. Chem. Phys.* **1992**, *96*, 6796–6806.

(28) Aidas, K.; Angeli, C.; Bak, K. L.; Bakken, V.; Bast, R.; Boman, L.; Christiansen, O.; Cimiraglia, R.; Coriani, S.; Dahle, P.; Dalskov, E. K.; Ekström, U.; Enevoldsen, T.; Eriksen, J. J.; Ettenhuber, P.; Fernández, B.; Ferrighi, L.; Fliegl, H.; Frediani, L.;

Hald, K.; Halkier, A.; Hättig, C.; Heiberg, H.; Helgaker, T.; Hennum, A. C.; Hettema, H.; Hjertenæs, E.; Høst, S.; Høyvik, I.-M.; Iozzi, M. F.; Jansík, B.; Jensen, H. J. Aa.; Jonsson, D.; Jørgensen, P.; Kauczor, J.; Kirpekar, S.; Kjærgaard, T.; Klopper, W.; Knecht, S.; Kobayashi, R.; Koch, H.; Kongsted, J.; Krapp, A.; Kristensen, K.; Ligabue, A.; Lutnæs, O. B.; Melo, J. I.; Mikkelsen, K. V.; Myhre, R. H.; Neiss, C.; Nielsen, C. B.; Norman, P.; Olsen, J.; Olsen, J. M. H.; Osted, A.; Packer, M. J.; Pawlowski, F.; Pedersen, T. B.; Provasi, P. F.; Reine, S.; Rinkevicius, Z.; Ruden, T. A.; Ruud, K.; Rybkin, V. V.; Sałek, P.; Samson, C. C. M.; de Merás, A. S.; Saue, T.; Sauer, S. P. A.; Schimmelpfennig, B.; Sneskov, K.; Steindal, A. H.; Sylvester-Hvid, K. O.; Taylor, P. R.; Teale, A. M.; Tellgren, E. I.; Tew, D. P.; Thorvaldsen, A. J.; Thøgersen, L.; Vahtras, O.; Watson, M. A.; Wilson, D. J. D.; Ziolkowski, M.; Ågren, H. The Dalton quantum chemistry program system. *WIREs: Comput. Mol. Sci.* **2014**, *4*, 269–284.

(29) Vahtras, O. LoProp for Dalton. 2014; `http://dx.doi.org/10.5281/zenodo.13276`.

(30) Hohenstein, E. G.; Sherrill, C. D. Density fitting of intramonomer correlation effects in symmetry-adapted perturbation theory. *J. Chem. Phys.* **2010**, *133*, 014101.

(31) Turney, J. M.; Simmonett, A. C.; Parrish, R. M.; Hohenstein, E. G.; Evangelista, F. A.; Fermann, J. T.; Mintz, B. J.; Burns, L. A.; Wilke, J. J.; Abrams, M. L.; Russ, N. J.; Leininger, M. L.; Janssen, C. L.; Seidl, E. T.; Allen, W. D.; Schaefer, H. F.; King, R. A.; Valeev, E. F.; Sherrill, C. D.; Crawford, T. D. Psi4: an open-source ab initio electronic structure program. *WIREs: Comput. Mol. Sci.* **2012**, *2*, 556–565.

(32) Nagy, B.; Jensen, F. In *Reviews in Computational Chemistry*; Parrill, A. L., Lipkowitz, K. B., Eds.; John Wiley & Sons, Inc, 2017; Vol. 30; pp 93–149.

(33) Kumar, A.; Meath, W. J. Dipole oscillator strength properties and dispersion energies for acetylene and benzene. *Mol. Phys.* **1992**, *75*, 311–324.

(34) Podeszwa, R.; Bukowski, R.; Szalewicz, K. Potential energy surface for the benzene dimer and perturbational analysis of π-π interactions. *J. Phys. Chem. A* **2006**, *110*, 10345–10354.