A simple fragment-based method for van der Waals corrections over density functional theory

Prasanta Bandyopadhyay, Priya, and Mainak Sadhukhan*

Department of Chemistry, Indian Institute of Technology, Kanpur, 208016, India

E-mail: mainaks@iitk.ac.in

Abstract

Modelling intermolecular noncovalent interactions between large molecules remains a challenge for electron structure theory community. This is due to the high cost of calculating electron correlation energy. Fragment-based methods usually fare well in reducing the cost of computations in such systems while quantum Drude oscillators turn out to be a good model for van der Waals interactions. In this paper, we have developed a simple yet effective method based on oscillator methods for calculating van der Waals interactions between molecular fragments as a correction to low-cost DFT functional PBE. We have tested our method on S66X8 with significant success.

1 Introduction

van der Waals (vdW) is a crucial ingredient in understanding the non-covalent interactions between molecular fragments.\textsuperscript{1-4} It originates from instantaneous and correlated quantum fluctuation of electron density electronic systems. As a result, they are ubiquitous and essential to compute with high accuracy. Being long-range part of electron correlation energy, an accurate estimation of the vdW interactions can be obtained from post-Hartree-Fock methods. These methods including coupled-cluster (CC) and configuration interaction (CI)
calculations provide a systematic and accurate results. However, these methods are computationally too expensive to be used for large molecules particularly for molecular dynamics simulations. On the contrary, density functional theory-based approximations (DFTA) usually fail to describe the vdW interactions properly due to inadequate description of electron correlations. As a result, dispersion corrections are usually employed on top of DFT methods. In this paper, we aim to describe one such correction scheme which is an amalgamation of quantum oscillator based models for quantum fluctuations\(^2\) and fragment-based models (see\(^7\) for a comprehensive review).

The usual vdW corrections on top of DFT-based methods can be obtained by several strategies. From the perturbation theory, the dispersion energy can be written as a sum of pair-wise terms as\(^,6,8\)

\[ E_{\text{disp}}^{(2)} = -\sum_{i<j} \left( \frac{C_{6,ij}}{R_{ij}^6} + \frac{C_{8,ij}}{R_{ij}^8} + \frac{C_{10,ij}}{R_{ij}^{10}} + \cdots \right) \]  

(1)

The i, j are pair of atoms/systems, \(R_{ij}\) is the distance between i and j, \(C_n\) are dispersion coefficients arising from dipole-dipole (n=6), dipole-quadruple (n=8), quadruple-quadruple (n=10) interactions. In general, the dispersion energy is not pair-wise additive because of the many-body effects. Widely utilized methods for dispersion corrections as Grimme’s -D2\(^9\), -D3\(^10\) uses semi-empirical parameters. On the other hand, Tkatchenko-Scheffler method\(^11\) captures the electronic many-body effects to some extent by defining the electron density based Hirshfeld screening of polarizabilities. Both of these methods has been employed successfully for large molecular systems\(^12,13\). The other way of introducing the vdW correction is through quantum Drude oscillator (QDO) model which is a quantum harmonic oscillators with two equal and opposite charges tagged at two ends. As a result, two QDOs interact between themselves through Coulomb interactions. This model has been successfully employed to capture long-range interaction between atoms\(^6,14–17\). In particular, dispersion interaction in anionic water-clusters,\(^18,19\) vibrational spectroscopy in charges water
clusters,\textsuperscript{20} as a precursor in force-field development,\textsuperscript{21–23} simulating polarizable molecular ionic-liquid\textsuperscript{24} and water-lipid interface\textsuperscript{25} have been investigated using QDO. Many-Body Dispersion (MBD) method\textsuperscript{5} has been another direction where QDO finds its use. In this method, the long-range quantum fluctuations of each atom is modeled by a QDO. The oscillators’ polarizabilities have been set to the screened polarizabilities. The dipole-coupled interaction energy between these oscillators therefore provides the interaction energy which is equivalent to random phase approximation (RPA) of correlation energy.\textsuperscript{26,27} Since this model allows for an exact diagonalization of the interaction matrix, it captures many-body effects of all orders.\textsuperscript{28,29}

On the other hand, molecular fragmentation techniques have been developed to reduce computational costs for large molecules. Upon fragmentation, a large composite system can be divided into small sub-systems. Thereafter, the results of the high-level calculations on small fragment can be assembled to the desired properties of the large system.\textsuperscript{7,30} Conventional fragment-based methods to calculate the energy of the system can be categorized broadly as overlapping or top-down approach (based inclusion-exclusion principle) and non-overlapping or bottom-up approach (based on many-body expansion).\textsuperscript{7} MFCC (Molecular Fragmentation with Conjugated Caps),\textsuperscript{31,32} MTA (Molecular Tailoring Approach), GEBF (Generalized Energy Based Fragmentation)\textsuperscript{34,35} are examples as overlapping fragmentation methods. In these methods, the overcounted fragments are cancelled out to calculate the energy of the derivative subsystems. MFCC utilizes this methodology implicitly. Bottom-up approaches, based upon many-body interactions, first create small non-overlapping fragments to calculate the interactions between them. FMO (Fragment Molecular Orbital),\textsuperscript{30,36,37} EE-MB (Electrostatically Embedded Many-Body Method),\textsuperscript{38,39} KEM (Kernel Energy Method) methods\textsuperscript{40–42} follow this methodology. MOBE (Many Overlapping Body expansion)\textsuperscript{43} and GMBE (Generalized Many-Body Expansion)\textsuperscript{44} use both types of top-down and bottom-up approach. In these studies, several fragmentation strategies based on inter-molecular distances, system-size have been employed. Several fragmentation schemes require inter-
fragment covalent bonds to be cut and be capped that vacancy by atoms (mostly H) or a simple group of atoms. This capping makes the charge and multiplicity of the fragments and valence of the adjacent atom intact. Apart from the above mentioned classifications, Richard and Herbert also suggested classification of fragmentation based on intersecting nuclei and disjoint nuclei methods.\textsuperscript{44}

In this article, we used an amalgamation of the two above-mentioned methods where after fragmenting the molecule based on simple chemical intuitions, we calculated the long-range vdW interaction energies between them via dipole-coupled quantum Drude oscillators. Such coarse-graining of dispersion interaction helps reduce the computational costs for large molecules. As a result, it is less expensive than that of Dn methods of Grimme as well as traditional oscillator-based methods. We then tested the efficacy of our method on S66x8 dataset.\textsuperscript{45} We found that even this simple procedure provides very good results for most of the systems.

This article is organized as follows. We discussed the theoretical motivations behind using QDO model as well as computational methodologies in section 2. Next, we analysed the fragmentation procedure carefully in section 3.1 followed by the performance of our method for S66X8 database in section 3.2. Then we discussed the performance of other exchange-correlation functionals other than PBE in section 3.3. Finally, we conclude the paper in section 4.

2 Theory and Computational Methods

2.1 QDO model for dispersion energy

We start with the definition of electron correlation energy derived from adiabatic connection fluctuation dissipation theorem (ACFDT).\textsuperscript{5} According to this formulation, the electron
correlation can be written in terms of density response function \( \chi \) and the interaction \( v \).\(^{26}\)

\[
E_c = -\frac{1}{2\pi} \int_0^\infty dw \int_0^1 d\lambda \, \text{Tr}[\chi_\lambda v - \chi_0 v] \tag{2}
\]

Since we are interested in long range correlation energy only, RPA\(^{46,47}\) can be employed to compute the vdW interactions. In RPA interacting response function \( \chi_\lambda \) is approximated as:

\[
\chi_{\lambda}^{\text{RPA}} = \frac{\chi_0}{1 - \chi_0 \lambda v} \tag{3}
\]

In earlier studies,\(^{26}\) it has been shown that the molecules can be divided into effective atomic fragments. Atomic response functions then can be represented by a set of quantum harmonic oscillators.\(^{48}\) Therefore in Eq.(2) can be expressed as

\[
E_{c,\text{RPA}} = \frac{1}{2\pi} \int_0^\infty dw \, \text{Tr}[\ln(1 - AT)]
\]

\[
= -\frac{1}{2\pi} \int_0^\infty dw \sum_n \text{Tr} \left[ \frac{(AT)^n}{n} \right]. \tag{4}
\]

where \( A \) is a \( 3N \times 3N \) matrix (\( N \) is number of atoms). For isotropic QHO \( A_{lm} = -\delta_{lm} \alpha_l (iw) \) and \( T \) is the dipole-dipole interaction tensor.

We can, however, use a unitary operator \( U \) such that

\[
A' = UAU^\dagger \quad \text{and} \quad T' = U^\dagger T U^\dagger \tag{5}
\]

Consequently, \( A' \) is no longer a diagonal matrix, although

\[
AT = U^\dagger A'T'U \tag{6}
\]

leading to

\[
\text{Tr}[(AT)^n] = \text{Tr}[(A'T')^\dagger]. \tag{7}
\]
If $A'$ can be constructed in a block-diagonal form, these blocks can be interpreted as fragment polarizabilities and therefore Eq.(4) can be written, in principle, as

$$E_{c,RPA} = \frac{1}{2\pi} \int_0^\infty dw \sum_n \text{Tr} \left[ \frac{(A'T')^n}{n} \right]$$

(8)

where $A'$ is the fragment polarizability matrix. However, in this paper we have chosen fragments out of simple chemical intuitions. As a result, our fragments are insensitive to chemical environments. Note that $A'$ defined through Eq.(5) depends on chemical ambiance. More rigorous way of defining polarizability fragments based on unitary transformation will be taken up in future. Here we also approximate $T'$ by a simple dipole-dipole interaction matrix between the center-of-masses of two fragments. We modeled the quantum fluctuations of the molecular fragments by anisotropic oscillators.$^{14,26,49}$ The dispersion energies by perturbation theory given by second-order energy.$^{50}$

$$U_{Disp} = - \sum_{m_A \neq 0, n_B \neq 0} \frac{\langle 0_A 0_B | H' | m_A n_B \rangle \langle m_A n_B | H' | 0_A 0_B \rangle}{W_{m0}^A + W_{n0}^B}.$$  

(9)

Here $m_i = (m_x, m_y, m_z)$ is an excited state of $i$ oscillator. This expression gives the dispersion energy of electronic fragments A and B. $H'$ is interaction energy operator (see SI) and $W_{m0}^i = W_{m0} - W_{00}^i$ is the excitation energy of molecule $i$ with $m$ quanta.

The detailed derivations and expressions for the dispersion energy using anisotropic oscillators are given in supplementary information (SI).

### 2.2 Computational Methodologies

To benchmark our method, we used S66X8 dataset for its diversity of interaction types. The interaction energy curves from this data-set also help analyzing the effects of fragmentation as well as many-body effects. We have computed all single-point energies using def2-QZVPPD$^{51}$ basis set. The underlying DF interaction energies are calculated as the difference between
dimer and monomer energies. In this study, several exchange-correlation (XC) functionals including PBE,\textsuperscript{52} PBE0,\textsuperscript{53} B3LYP\textsuperscript{54,55} and BHLYP\textsuperscript{54,55} are used. These XC functionals provide good approximations for short-range interactions. The mid-range and long-range part of correlation energy which constitutes vdW interaction can be computed as,

\[ E_{vdW} = \Delta E_{CCSD(T)/CBS} - \Delta E_{DF} \]  \hspace{1cm} (10)

Here \( \Delta E_{CCSD(T)/CBS} \) and \( \Delta E_{DF} \) are interaction energies computed by CCSD(T) and approximated density functionals, respectively. Here \( \Delta E_{DF} \) contains no dispersion correction. After performing single point energy calculation of the systems at all inter-monomer distances by DF methods, we have spliced the larger monomers into smaller fragments. Detailed fragmentation procedure is mentioned in results and discussion section.

To compute the vdW correction over DF energies, each fragment is replaced by an anisotropic QDO. The dipole-coupled dispersion energies between these oscillators (\( E_{QDO}^{ij}(r_{ij}) \) between \( i^{th} \) and \( j^{th} \) fragments) have been used as the vdW correction energy. The total dispersion energy is obtained by Eq.(11),

\[ E_{vdW} = \sum_{i \in A, j \in B} E_{QDO}^{ij}(r_{ij}) f_{damp}(r_{ij}). \]  \hspace{1cm} (11)

A Fermi-like damping function,\textsuperscript{56} 

\[ f_{damp}(r_{ij}) = \frac{1}{1 + \exp \left[ -B \left( \frac{r_{ij}}{R_{ij}} - 1 \right) \right]} \]  \hspace{1cm} (12)

is used to avoid the near-singularity at very small distance. Parameter \( B \) is taken as 6 for optimal performance of the method. \( r_{ij} \) is the distance between the centre of mass of the fragments \( i \) and \( j \). \( R_{ij} \) is the sum of vdW radii of the fragments \( R_{ij} = R_i + R_j \) obtained via\textsuperscript{57}

\[ R_i(\alpha) = 2.54\alpha^{1/7} \]  \hspace{1cm} (13)
Here $\alpha = \text{Tr} (\alpha_i)/3$, where $\alpha_i$ is the polarizability tensor of the $i^{th}$ fragment. Energies are computed with Turbomole v7.5 and static polarizability tensor at CCSD are computed with Psi4.

3 Results and discussion

3.1 Fragmentation of Monomers

We fragmented the molecular moieties purely based on chemical intuitions (e.g. along the functional groups). Apart from ethene, ethyne and water, all monomers are subjected to fragmentation. As an example, the partitioning of AcNH$_2$ results in two fragments, -CH$_3$ and -CONH$_2$ (Fig.1(a)). The fragments are then capped at the vacant site with H atoms to keep the charge and spin multiplicity unaltered. The polrizability tensor is computed for these fragments subsequently. Note that a balance between number of fragments and accuracy is imperative from the computational point-of-view. Later in this section, effect of fragmentation and the associated errors are discussed in details.

![Figure 1: Fragmentation of AcNH$_2$ and Benzene molecule. Green, white, blue and red represents C, H, N and O atoms, respectively. The cyan dashed lines indicates the line along which fragmentation are performed.](image)

8
The monomer fragmentation is done as follows.

- Non-cyclic or non-branched systems are cut and masked with H atom along the mid-points of bonds (see SI). Fragmenting these molecules in such a manner helps in parametrizing the associated QDOs and therefore provide a scope for transferibility of oscillator parameters across molecules. On the flip side, such fragmentations do not reflect the chemical environment for different molecules. The fragmentation of AcNH₂ is shown in Fig.1(a).

- For cyclic molecules, atom-centric fragmentation are performed. As an example, benzene is fragmented along the alternate C-H bond (as shown in Fig.1(b)), thus obtaining three systems of propene. Note two fragments share a bridging atom, thereby making our fragmentation scheme less accurate. However, since the bonds in these systems are electron-rich, the bond-centric fragmentation turns out to be even less accurate. The effects of double-counting atoms will be analyzed and corrected in future works.

3.2 Performance in S66x8

The S66x8 dataset can be categorized into three broad sections, electrostatically bound complexes (named Hydrogen bond by Rezac et al\textsuperscript{45}), dispersion-bound and other different interactions. In the following sections, we discuss the role of dispersion in the complexes and how effective our model is to capture this interaction.

3.2.1 Electrostatically bound complexes

Electrostatically bound complexes can be characterized by two properties, inter-monomer distance and very low dispersion/electrostatic ratio (hereafter $D/E$) as computed from SAPT analysis. In most cases, the $D/E < 0.5$\textsuperscript{45} Therefore the qualitative pictures for the interaction energies are essentially obtained from DF methods(see Fig.3). The -D2, -D3 and QDO correction on the DF interaction energies provide more accurate PEC. To analyze further,
the deviation of interaction energy by PBE-D2 ($\Delta E_{PBE-D2}$) and PBE-QDO ($\Delta E_{PBE-QDO}$) from benchmark coupled-cluster energies are plotted in Fig.2(a)) as obtained via

$$\Delta E_{PBE-QDO} = \Delta E_{PBE} + E_{QDO} - \Delta E_{(CCSD(T)/CBS)}$$

$$\Delta E_{PBE-D2} = \Delta E_{PBE-D2} - \Delta E_{(CCSD(T)/CBS)}$$

(14)

Except for the Uracil dimer base-pair interactions, deviation from the CCSD(T)/CBS is well within 1 kcal/mol. Out of 23 electrostatically bound complexes, seven complexes show error $\geq$ 0.5 kcal/mol. Ten systems show slightly overbinding (blue bars) and for other systems we observe underbinding with respect to CCSD(T). On the contrary, PBE-D2 method (red bars) always shows overbinding. The contribution from dispersion is negligible for majority of these complexes (i.e. < 20%) (Fig.2(b)). However in Peptide$\cdots$MeOH, Peptide$\cdots$MeNH$_2$ and Peptide$\cdots$Peptide complexes $D/E$ lies within 0.5 and 0.56 mirroring elevated QDO dispersion ($> 20\%$). Furthermore, for MeNH$_2$-dimer, MeNH$_2$$\cdots$Peptide and MeNH$_2$$\cdots$MeOH the dispersion contribution from QDO rise to $\approx 20 - 30\%$ (2 kcal/mol) while the D/E ratio lies within 0.71 and 0.79. Therefore our results are in excellent agreement with the benchmark results.

A close inspection of Fig.2 can reveal the nature and origin of the errors resulted in PBE-QDO method. For most of the complexes, our method predicts under-binding. This may occure due to the absence of dipole-quadrupole and higher order terms. Since 3-body ATM$^{60,61}$ correction is also important in such systems, inclusion of it may further correct our results. The overall deviation of our method along with pure PBE and PBE-D2 from CCSD(T) is tabulated in Table-1. QDO correction lowers the error of pure PBE significantly. Also for this subset, the QDO correction fares better than that of PBE-D2 method.

The interaction energy curve of MeOH dimer, as a representative of these class of complexes (fig.3) shows the comparison of interaction energy between CCSD(T), QDO and D2 corrected PBE. Throughout the distance range the PBE-QDO results show better agreement with CCSD(T) values compared to PBE-D2. The energy from the QDO interaction
Figure 2: a) $\Delta E_{PBE-QDO}$ (blue bars) and $\Delta E_{PBE-D2}$ (red bars) in kcal/mol with respect to CCSD(T)/CBS (represented by 0.0 value for abscissa) and b) percentage contributions from PBE and QDO dispersion towards PBE+QDO corrected interaction energy at equilibrium distance for electrostatically stabilized complexes.

Table 1: Statistical analysis of error at equilibrium distance with respect to benchmark results for PBE-QDO, PBE and PBE-D2 for electrostatically-bound complexes.

<table>
<thead>
<tr>
<th></th>
<th>PBE-QDO</th>
<th>PBE</th>
<th>PBE-D2</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAE$^a$</td>
<td>0.376785</td>
<td>0.961688</td>
<td>0.684686</td>
</tr>
<tr>
<td>MRE$^b$</td>
<td>0.044944</td>
<td>0.124844</td>
<td>0.089156</td>
</tr>
<tr>
<td>MRAE$^c$</td>
<td>0.002022</td>
<td>0.005149</td>
<td>0.003348</td>
</tr>
</tbody>
</table>

$^a$ Mean Absolute Error; $^b$ Mean Relative Error; $^c$ Mean Relative Absolute Error
is notable near equilibrium distance $\approx 1$ kcal/mol. At the long range, QDO contribution is negligible. PBE-D2 is found to be marginally overestimating the energy with respect to CCSD(T)/CBS.

![Figure 3: Interaction energy curve (in kcal/mol) for MeOH dimer plotted against the scaled distance between fragments.](image)

**3.2.2 Dispersion complexes**

Complexes with large D/E ratio ($1.3 < D/E < 5.42$) are categorized as dispersion complexes. These complexes include stacked dimers. Indeed Fig.5(b) shows relatively large QDO correction compared to DF interaction energies. It is evident that PBE functionals, by themselves do not predict any binding for these complexes. As expected for such cases, QDO correction is the most significant source of binding energies. In fact, for most of the complexes QDO contribution is very prominent.

As a representative of these complexes, we are exhibiting the interaction energy curve of cyclopentane dimer in fig.4. We can clearly see that the QDO contribution is very prominent near equilibrium distance. Moreover, our method consistently agree well with benchmark interaction energies at all distance. It is interesting to note that while D2 correction fares well in long-range, near equilibrium our method is far superior than it. In dispersion bound
complexes, we obtain both under- and overestimated interaction energies which are plotted in Fig.5(a). In most of the pi-pi stacking complexes interaction energy is over-estimated. Apart from these, uracil-cyclopentane and uracil-pentane also show overestimation of interaction energy to some extent. One of the possible causes may include non-optimal choice for the fragmentation method used in the present work. The results were deviated from benchmark numbers most for Benzene-dimer, Benzene-Pyridine and Pyridine-dimer complexes where the monomers are stacked (or parallel-displaced) over each other. To estimate the error associated with this fragmentation procedure, we have calculated the polarizability tensor of all the fragments and monomers at CCSD/def2-TZVPPD. According to Eq.(5), the $\text{Tr}(A') = \text{Tr}(A)$. We have been motivated by this to find out the difference between the sum of diagonal elements of $n$ fragment polarizability tensors (i.e. $\alpha_{\mu}$ for $\mu^{th}$ fragment) and trace of monomer polarizability tensor ($\alpha_{\text{monomer}}$) as

$$\Delta \alpha = \sum_{\mu=1}^{n} \text{Tr}(\alpha_{\mu}) - \text{Tr}(\alpha_{\text{monomer}})$$

13
and their ratio

\[ r_\alpha = \frac{\sum_{\mu=1}^{n} \text{Tr}(\alpha_\mu)}{\text{Tr}(\alpha_{\text{monomer}})} \] (16)

Clearly, an optimum fragmentation is expected to provide the \( r_\alpha \) that is close to unity. We can see from the table 2 that the systems with large deviations from the benchmark numbers are also associated with \( r_\alpha \gg 1 \). Therefore for stacking complexes the error in QDO interaction energies can be attributed to errors in fragmentations.

Table 2: Estimates of errors in fragmentation of monomers. All legends are explained in text (see Eq.(15) and Eq.(16)).

<table>
<thead>
<tr>
<th>Monomers</th>
<th>( n^a )</th>
<th>( \text{Tr}(\alpha_{\text{monomer}}) )</th>
<th>( \Delta\alpha )</th>
<th>( r_\alpha )</th>
</tr>
</thead>
<tbody>
<tr>
<td>AcNH(_2)</td>
<td>2</td>
<td>117.96</td>
<td>15.3</td>
<td>1.13</td>
</tr>
<tr>
<td>AcOH</td>
<td>2</td>
<td>102.67</td>
<td>14.92</td>
<td>1.14</td>
</tr>
<tr>
<td>MeNH(_2)</td>
<td>2</td>
<td>75.54</td>
<td>15.37</td>
<td>1.20</td>
</tr>
<tr>
<td>MeOH</td>
<td>2</td>
<td>62.62</td>
<td>15.6</td>
<td>1.25</td>
</tr>
<tr>
<td>Benzene</td>
<td>3</td>
<td>204.35</td>
<td>148.58</td>
<td>1.72</td>
</tr>
<tr>
<td>Cyclopentane</td>
<td>3</td>
<td>174.39</td>
<td>147.45</td>
<td>1.84</td>
</tr>
<tr>
<td>Pyridine</td>
<td>3</td>
<td>188.87</td>
<td>156.64</td>
<td>1.83</td>
</tr>
<tr>
<td>Uracil</td>
<td>3</td>
<td>211.26</td>
<td>131.03</td>
<td>1.62</td>
</tr>
<tr>
<td>Pentane</td>
<td>4</td>
<td>191.36</td>
<td>143.43</td>
<td>1.76</td>
</tr>
<tr>
<td>Peptide</td>
<td>4</td>
<td>152.88</td>
<td>41.12</td>
<td>1.27</td>
</tr>
<tr>
<td>Neopentane</td>
<td>4</td>
<td>190.52</td>
<td>145.03</td>
<td>1.76</td>
</tr>
</tbody>
</table>

\(^a\) Number of fragments.

For rest of the complexes, small under-estimatation of vdW interaction energy is observed. Notably, the interaction energy for pentane dimer and benzene-ethene complexes are underestimated by 1 kcal/mol. Overall, these large deviations in dispersion interactions can be attributed to the lack of Axilrod-Teller-Muto (ATM)\(^{60,61}\) correction and higher order attractive correction terms. Note that most of our corrections has opposite signs than that of PBE-D2 in contrast to electrostatically bound cases discussed before. While it is difficult to pinpoint the reasons for this observation, we can guess that the inclusion of higher order terms may eventually correct for this discrepancy. From statistical viewpoint, both -D2 and QDO dispersion correction reduces the error of PBE (Table -3), while -D2 method is slightly better than QDO corrections.
Table 3: Statistical analysis of error at equilibrium distance with respect to benchmark results for PBE-QDO, PBE and PBE-D2 for dispersion-bound complexes.

<table>
<thead>
<tr>
<th></th>
<th>PBE-QDO</th>
<th>PBE</th>
<th>PBE-D2</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAE&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.601501</td>
<td>3.665022</td>
<td>0.502804</td>
</tr>
<tr>
<td>MRE&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.207462</td>
<td>1.056156</td>
<td>0.158808</td>
</tr>
<tr>
<td>MRAE&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.008470</td>
<td>0.042752</td>
<td>0.006559</td>
</tr>
</tbody>
</table>

<sup>a</sup> Mean Absolute Error; <sup>b</sup> Mean Relative Error; <sup>c</sup> Mean Relative Absolute Error

Figure 5: a) $\Delta E_{PBE-QDO}$ (blue bars) and $\Delta E_{PBE-D2}$ (red bars) in kcal/mol with respect to CCSD(T)/CBS (represented by 0.0 value for abscissa) and b) percentage contributions from PBE and QDO dispersion towards PBE+QDO corrected interaction energy at equilibrium distance for dispersion stabilized complexes.
3.2.3 Other interactions

Complexes within this category interact mostly via π-electrons. Here we show the interaction energy curve of Benzene-water dimer as our representative system (fig.7). As expected, the magnitude of QDO correction here is in between those from previous two sections. While the error due to D2 correction is not much, our correction is significantly better than D2 here. However, here the D2 approximates the long-range correction much better compared to our method. This may be due to the change of polarizability of fragments at different intermonomer distances which is not considered in our method. In future, therefore, such information should be incorporated in our method. Similar to previously mentioned systems, interaction energy by DF approximations gets better with QDO correction. The deviations of PBE-QDO and PBE-D2 from the CCSD(T) are presented in Fig.6(a). Our results deviate from CCSD(T) by more than 1 kcal/mol only for ethyne-AcOH and peptide-ethene. For majority of complexes, PBE-QDO predicts under-binding. This feature has also been observed for dispersion-dominated complexes. As expected, for systems with large dispersion energy component, the QDO correction becomes more significant (Fig.6(b)). We find that for most of these systems dispersion correction constitutes < 40% of total binding energies. Statistically, (Table-4) we find that both D2 and QDO improves PBE binding energies. However, our method appears to work slightly better than D2.

Table 4: Statistical analysis of error at equilibrium distance with respect to benchmark results for PBE-QDO, PBE and PBE-D2 for mixed complexes.

<table>
<thead>
<tr>
<th></th>
<th>PBE-QDO</th>
<th>PBE</th>
<th>PBE-D2</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAE(^a)</td>
<td>0.459231</td>
<td>1.76766</td>
<td>0.630254</td>
</tr>
<tr>
<td>MRE(^b)</td>
<td>0.134138</td>
<td>0.49693</td>
<td>0.175168</td>
</tr>
<tr>
<td>MRAE(^c)</td>
<td>0.007765</td>
<td>0.026944</td>
<td>0.010483</td>
</tr>
</tbody>
</table>

\(^a\) Mean Absolute Error; \(^b\) Mean Relative Error; \(^c\) Mean Relative Absolute Error
Figure 6: a) $\Delta E_{PBE-QDO}$ (blue bars) and $\Delta E_{PBE-D2}$ (red bars) in kcal/mol with respect to CCSD(T)/CBS (represented by 0.0 value for abscissa) and b) percentage contributions from PBE and QDO dispersion towards PBE+QDO corrected interaction energy at equilibrium distance for other type complexes.

Figure 7: Interaction energy curve (in kcal/mol) for Benzene···Water OH−π complex plotted against the scaled distance between fragments.
3.2.4 Overall performance

The overall error in terms of MAE, MRE and MRAE for all the complexes are tabulated in Table-5. In terms of accuracy, the interaction energy of PBE-QDO is rather accurate. In electrostatic and mixed interaction type complexes, the accuracy is even surpassed than that of the the -D2 correction. Although the accuracy was decreased in dispersion complexes, mostly due to the overestimation of stacking complexes, it is still comparable to -D2. In 66 complexes, only 8 complexes show deviations from CCSD(T) energy by more than 1 kcal/mol.

Table 5: Statistical analysis of error at equilibrium distance with respect to benchmark results for PBE-QDO, PBE, PBE-D2, BHLYP-QDO, B3LYP-QDO and PBE0-QDO for all complexes.

<table>
<thead>
<tr>
<th>MAE</th>
<th>PBE-QDO</th>
<th>PBE</th>
<th>PBE-D2</th>
<th>BHLYP-QDO</th>
<th>B3LYP-QDO</th>
<th>PBE0-QDO</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.480079</td>
<td>2.147993</td>
<td>0.604808</td>
<td>0.628023</td>
<td>1.118247</td>
<td>0.516513</td>
<td></td>
</tr>
<tr>
<td>0.128608</td>
<td>0.562147</td>
<td>0.139492</td>
<td>0.182652</td>
<td>0.323134</td>
<td>0.133437</td>
<td></td>
</tr>
<tr>
<td>0.001399</td>
<td>0.005983</td>
<td>0.001617</td>
<td>0.001687</td>
<td>0.002810</td>
<td>0.001691</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\) Mean Absolute Error; \(^b\) Mean Relative Error; \(^c\) Mean Relative Absolute Error

3.3 Comparison of data with different density functionals

To analyse the role of different energy components we have applied QDO correction on top of other XC functionals as well. For this we used B3LYP and BHLYP as well as PBE0 functionals. In agreement with Brauer et al\(^62\) we find even with QDO correction B3LYP provides very inaccurate binding energies. Furthermore, we tested QDO model with hybrid functional BHLYP which shows better results than B3LYP but worse than PBE. For all of the functionals the inclusion of QDO correction however improves the binding energy. This suggests the exchange energy (since B3LYP and BHLYP both contain LYP correlation) is less important in computing interaction energies around equilibrium geometries. Interaction energy variation with intermonomer distance of all the complexes for these functionals are provided in supporting information. The effect of damping function for different functionals

18
remains however an open question which will be taken up in future.

4 Conclusion

In summary, we have developed a method to calculate intermolecular dispersion interaction based on a fragmentation technique. We represented the quantum fluctuations of electron densities for these fragments by anisotropic quantum Drude oscillators. It turns out that the method’s success heavily depends on the type of fragmentation used. This aspect has been tested with comparing the monomer polarizability with fragment polarizability. We applied our methodology on S66X8 database to understand its efficacy as well as drawbacks. We find that the method works less accurately only for stacked dimers where the electron density is spreadout and therefore highly polarizable. In general, however, this method provides a simple yet effective approximation of dispersion interactions in large molecular systems with accuracy at least at par with popular D2 method. Particularly, we find that except 8 systems, QDO-corrected interaction energy for all other complexes lies within 1 kcal./mol of CCSD(T) energy. We also find that our correction due to QDO is maximum for dispersion-bound complexes. For such systems, the QDO energy almost provides 100% binding energy. The mean absolute error also shows that our method works $\approx 20\%$ better than PBE-D2 methods. Another important aspect of the present work is the use of functional group-based fragmentation scheme. This fragmentation scheme while insensitive to different chemical environments in different systems, is more amenable to application in large systems such as proteins. We also compared the efficacy of different exchange-correlation energies to be used in conjunction to our correction. It turns out that our result, can be used for multiple exchange correlation functionals. However, for B3LYP the correction does not improve the binding energy which is at par with existing literature. The major limitations of our method are twofold: (1) the ad-hoc fragmentation procedure and (2) the lack of screening effects due to local environment. In future, we will investigate ways to include such information
effectively in our scheme.

Supporting Information Available

The calculation for the expression of dispersion energy, fragmentation scheme applied for monomers and interaction energy curves for all 66 complexes have been provided.

Acknowledgement

PB thanks IIT Kanpur for an Institute Post-Doctoral Fellowship. Priya thanks IIT Kanpur for a junior research fellowship. MS thanks SERB grant No. SRG/2019/000369 for computational support and Debashree Manna for numerous stimulating and useful discussions.
References


(9) Grimme, S. Semiempirical GGA-type density functional constructed with a long-range

(10) 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

(11) Tkatchenko, A.; Scheffler, M. Accurate Molecular Van Der Waals Interactions from


with dispersion corrections for supramolecular structures, aggregates, and complexes of

(14) Voora, V. K.; Ding, J.; Sommerfeld, T.; Jordan, K. D. A self-consistent polarization
potential model for describing excess electrons interacting with water clusters. *Journal

(15) Sadhukhan, M.; Manby, F. R. Quantum mechanics of Drude oscillators with full

(16) Odbadrakh, T.; Voora, V.; Jordan, K. Application of electronic structure methods to


(34) Li, W.; Li, S.; Jiang, Y. Generalized Energy-Based Fragmentation Approach for Com-


