Georgia Tech

Implementation and Application of Density Functional Theory based Symmetry-Adapted Perturbation Theory for Dimers, Trimers and Molecular Crystals

Yi Xie

July 25, 2022

Noncovalent Interaction

- Phase transition, stability of crystal structure
- Drug binding, DNA/RNA/protein structure
- Many-body Expansion for the energy of complex system:

$$E = \sum_A E_A + \sum_{AB} E_{AB}^{\mathrm{int},2} + \sum_{ABC} E_{ABC}^{\mathrm{int},3} + \cdots$$

Intermolecular Energies

► Supermolecular approach

$$E_{AB}^{\text{int},2} = E_{AB} - E_A - E_B$$

$$E_{ABC}^{\text{int},3} = E_{ABC} - E_{AB} - E_{AC} - E_{BC} + E_A + E_B + E_C$$

► Symmetry-Adapted Perturbation Theory (SAPT)

$$E^{\text{int},2} = E_{\text{elst}}^{(1)} + E_{\text{exch}}^{(1)} + E_{\text{ind}}^{(2)} + E_{\text{exch-ind}}^{(2)} + E_{\text{disp}}^{(2)} + E_{\text{exch-disp}}^{(2)}$$

SAPT(DFT)

▶ Hamiltonian partitioning:

$$H = F_A + F_B + V_{AB} + W_A + W_B$$

$$H = K_A + K_B + V_{AB}$$

▶ "Uncoupled" sum-over-states approximation of $E_{\text{disp}}^{(2)}$ and in terms of frequency-dependent density susceptibility (FDDS):

$$
\begin{aligned}
E_{\text{disp,u}}^{(2)} &= -4 \sum_{ar \in A, bs \in B} \frac{|(ar|bs)|^2}{\epsilon_{ab}^{rs}} \\
&= -\frac{1}{2\pi} \int_0^\infty d\omega \int d\mathbf{r}_A d\mathbf{r}_A' d\mathbf{r}_B d\mathbf{r}_B' \\
&\quad \frac{1}{|\mathbf{r}_A - \mathbf{r}_B|} \frac{1}{|\mathbf{r}_A' - \mathbf{r}_B'|} \chi_0^A \left(\mathbf{r}_A, \mathbf{r}_A'|i\omega\right) \chi_0^B \left(\mathbf{r}_B, \mathbf{r}_B'|i\omega\right)
\end{aligned}
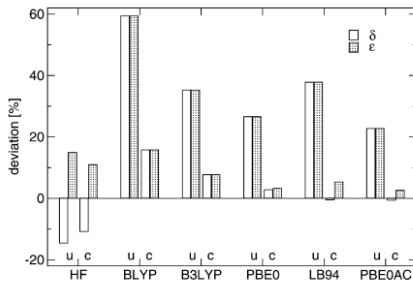$$

Uncoupled $E_{\text{disp}}^{(2)}$

Fig. 2. Mean (δ) and mean absolute (ϵ) percental deviations of the uncoupled (u) and coupled (c) second-order dispersion energies from the MP2 results.

Mean (δ) and mean absolute (ϵ) percentage deviations of uncoupled (u) and coupled (c) $E_{\text{disp}}^{(2)}$ from SAPT2+ results.

A. Heßelmann and G. Jansen, Chem. Phys. Lett. **367**, 778 (2003).

Coupled $E_{\mathrm{disp}}^{(2)}$

▶ Replacing uncoupled FDDS with coupled FDDS, solved from the coupled Kohn–Sham (CKS) TDDFT equation:

$$\boldsymbol{\chi} = \boldsymbol{\chi}_0 + \boldsymbol{\chi}_0 \mathbf{S}^{-1} \mathbf{W} \left(\mathbf{S} - \boldsymbol{\chi}_0 \mathbf{S}^{-1} \mathbf{W} \right)^{-1} \boldsymbol{\chi}_0$$

▶ Exchange-correlation kernel term in $\mathbf{W}$ approximated by adiabatic local-density approximation (ALDA) kernel:

$$
\begin{aligned}
W_{PQ} &= \left(P | r_{12}^{-1} | Q \right) + \left(P | f_{\mathrm{xc}} | Q \right) \\
&\approx \left(P | r_{12}^{-1} | Q \right) + \left(P | f_{\mathrm{xc}}^{\mathrm{ALDA}} | Q \right)
\end{aligned}
$$

M. Pitoňák and A. Hesselmann, J. Chem. Theory Comput. **6**, 168 (2010).

Hybrid Functional

- Local Hartree–Fock (LHF) approach
  - Computing LHF potential in each KS SCF iteration
  - $O(N^4)$ with very large constant factor
  - Different set of KS orbitals with smaller occupied–virtual gap
- Hybrid ALDA kernel
  - Mixing CHF and CKS equations to solve for FDDS
  - CKS involves integral of form $(ar|a'r')$, $O(N^4)$ with density fitting
  - CHF involves $(aa'|rr')$ and $(ar'|a'r)$, $O(N^5)$

Coupled $E^{(2)}_{\text{exch-disp}}$

▶ Scaling from scaling uncoupled $E^{(2)}_{\text{exch-disp}}$

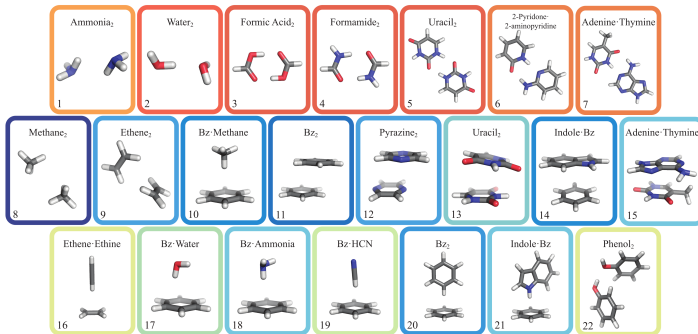$$\tilde{E}^{(2)}_{exch-disp,r} = E^{(2)}_{exch-disp,u} \cdot \frac{E^{(2)}_{disp,r}}{E^{(2)}_{disp,u}}$$

▶ Fixed scaling factor from fitting S22×5

$$\tilde{E}^{(2)}_{exch-disp,r} = \alpha \cdot E^{(2)}_{exch-disp,u} (\alpha = 0.686)$$

▶ Value of α above fitted from $E^{(2)}_{exch-disp,u}$ *with LHF orbitals*

A. Heßelmann and T. Korona, J. Chem. Phys. **141**, 094107 (2014).

S22 Dimer Set

Georgia Tech

Coupled FDDS with Hybrid Kernel

▶ Recall coupled FDDS for pure ALDA kernel:

$$\boldsymbol{\chi} = \boldsymbol{\chi}_0 + \boldsymbol{\chi}_0 \mathbf{S}^{-1} \mathbf{W} \left(\mathbf{S} - \boldsymbol{\chi}_0 \mathbf{S}^{-1} \mathbf{W} \right)^{-1} \boldsymbol{\chi}_0$$

▶ Coupled FDDS for hybird ALDA kernel, with $(aa'|rr')$ and $(ar'|a'r)$ contributions in $\boldsymbol{\chi}_0'$ and $\mathbf{K}'$:

$$\boldsymbol{\chi} = \boldsymbol{\chi}_0' + \left(\boldsymbol{\chi}_0' \mathbf{S}^{-1} \mathbf{W} + \mathbf{K}' \right) \left[\mathbf{S} - \left(\boldsymbol{\chi}_0' \mathbf{S}^{-1} \mathbf{W} + \mathbf{K}' \right) \right]^{-1} \boldsymbol{\chi}_0'$$

▶ Dispersion energy from integration over ω:

$$E_{\mathrm{disp,r}}^{(2)} = -\frac{1}{2\pi} \int_0^\infty d\omega \mathrm{Tr} \left(\mathbf{S}^{-1} \boldsymbol{\chi}^A \mathbf{S}^{-1} \boldsymbol{\chi} \right)$$

Refitting $E^{(2)}_{\text{exch}-\text{disp,r}}$ for non-LHF Orbitals

▶ LHF vs. non-LHF orbitals: Only affects uncoupled second-order terms like $E^{(2)}_{\text{disp,u}}$ and $E^{(2)}_{\text{exch}-\text{disp,u}}$

▶ Similar $E^{(2)}_{\text{disp,r}}$ for LHF + pure ALDA vs. non-LHF + hybrid ALDA, expect the same for $E^{(2)}_{\text{exch}-\text{disp,r}}$

▶ Can fit non-LHF $E^{(2)}_{\text{disp,u}}$ to LHF + pure ALDA $E^{(2)}_{\text{disp,r}}$:

$$
\begin{aligned}
E^{(2)}_{\text{disp,r}}(hybrid) &\approx E^{(2)}_{\text{disp,r}}(LHF) \\
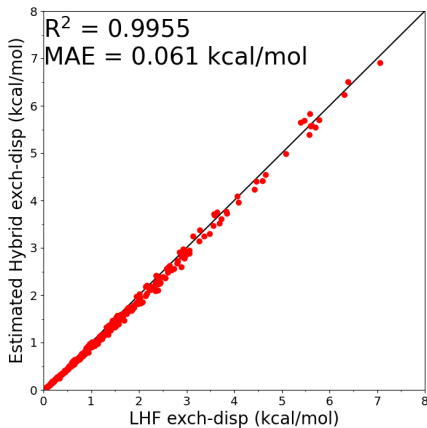&\approx \alpha \cdot E^{(2)}_{exch-disp,u}(non-LHF)
\end{aligned}
$$

▶ Fit for α using S22×5 and test with S66×8

SAPT Terms: LHF vs non-LHF Orbitals
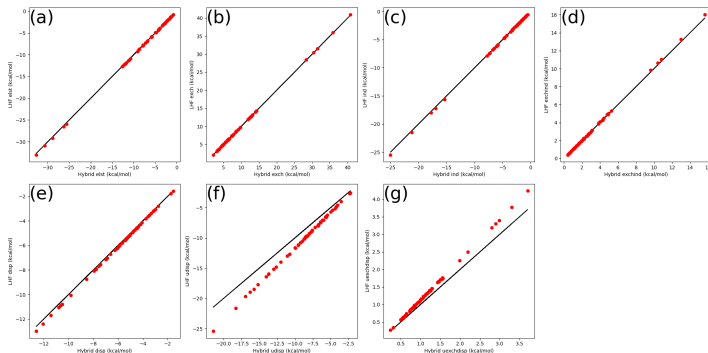
Hybrid vs. LHF values in kcal/mol for each term for S66 data set: (a) $E_{elst}^{(1)}$, (b) $E_{exch}^{(1)}$, (c) $E_{ind}^{(2)}$, (d) $E_{exch-ind}^{(2)}$, (e) $E_{disp,r}^{(2)}$, (f) $E_{disp,u}^{(2)}$, (g) $E_{exch-disp,u}^{(2)}$

Fitting: S22×5
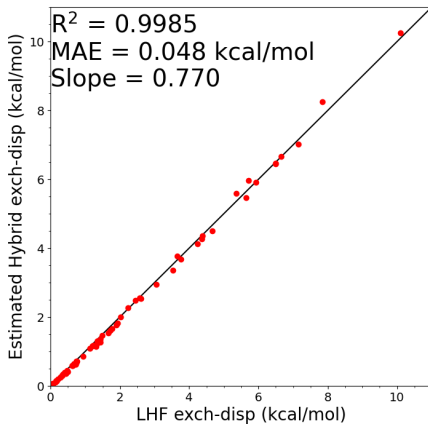
Testing: S66×8

Computational Approach

- ▶ Benchmarking with S66 dimer set
- ▶ Using SAPT2+3(CCD)δMP2/aug-cc-pVTZ as reference
- ▶ Compare electrostatics, exchange, induction, dispersion terms and total IE
- ▶ Using aug-cc-pVDZ for SAPT0 and SAPT2+, aug-cc-pVTZ for other methods

S66 Results

Method	Total	HB	MX	DD	Error Distribution						
					4	OB	1	0	1	UB	4
Electrostatics											
SAPT(DFT) hybrid	0.374	0.556	0.177	0.311							
SAPT(DFT) LHF	0.423	0.666	0.196	0.319							
SAPT0	0.613	1.034	0.439	0.297							
SAPT2+	0.236	0.270	0.136	0.263							
SAPT2+(3)δMP2	0.000	0.000	0.000	0.000							
Exchange											
SAPT(DFT) hybrid	0.886	1.127	0.426	0.926							
SAPT(DFT) LHF	0.886	1.121	0.431	0.928							
SAPT0	0.675	0.942	0.263	0.658							
SAPT2+	0.337	0.467	0.222	0.277							
SAPT2+(3)δMP2	0.000	0.000	0.000	0.000							
Induction											
SAPT(DFT) hybrid	0.211	0.201	0.212	0.220							
SAPT(DFT) LHF	0.224	0.223	0.223	0.225							
SAPT0	0.241	0.200	0.261	0.271							
SAPT2+	0.327	0.384	0.250	0.318							
SAPT2+(3)δMP2	0.152	0.179	0.121	0.145							

S66 Results (cont.)

Method	Total	HB	MX	DD	Error Distribution						
					4	OB	1	0	1	UB	4
Dispersion											
SAPT(DFT) hybrid	0.370	0.260	0.219	0.573							
SAPT(DFT) LHF	0.308	0.200	0.173	0.499							
SAPT(DFT) non-hybrid	0.635	0.581	0.419	0.822							
SAPT0	0.443	0.862	0.162	0.195							
SAPT2+	0.235	0.397	0.169	0.115							
SAPT2+(3)δMP2	0.093	0.129	0.056	0.080							
Total											
SAPT(DFT) hybrid	0.334	0.588	0.107	0.217							
SAPT(DFT) LHF	0.234	0.382	0.046	0.199							
SAPT(DFT) non-hybrid	0.604	0.955	0.389	0.385							
SAPT0	0.990	1.197	0.692	0.965							
SAPT2+	0.230	0.235	0.138	0.280							
SAPT2+(3)δMP2	0.105	0.056	0.082	0.169							

Timing Systems
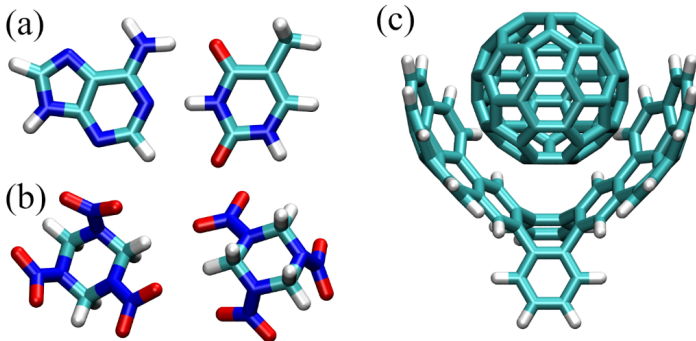
Dimer systems for timing: (a) Watson-Crick adenine-thymine complex, (b) RDX dimer, (c) C_{60}–buckycather complex.

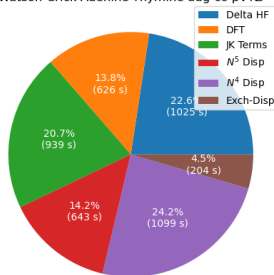# Watson-Crick Adenine-Thymine

(a) Watson-Crick Adenine-Thymine aug-cc-pVDZ
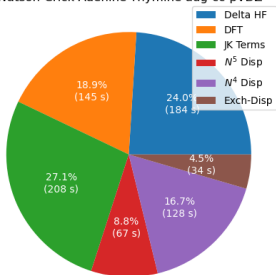
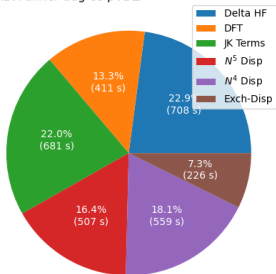(b) Watson-Crick Adenine-Thymine aug-cc-pVTZ

$N_{bf} = 536$

$N_{bf} = 1127$
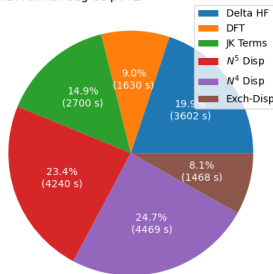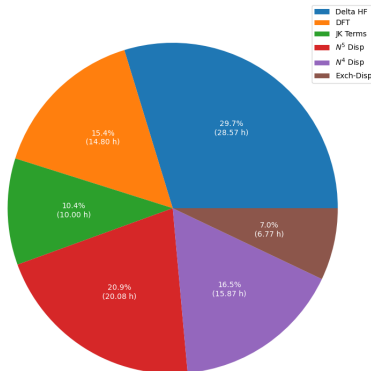
RDX Dimer

(c) RDX dimer aug-cc-pVDZ

(d) RDX dimer aug-cc-pVTZ

$N_{bf} = 798$

$N_{bf} = 1656$

C_{60}–Buckycatcher Complex

$$N_{bf} = 3012$$

Conclusions

▶ Implemented SAPT(DFT) using hybrid xc kernel, applicable to up to 3000 basis functions

▶ Scaling factor of $E^{(2)}_{\mathrm{exch-disp}}$ determined to be 0.770

▶ Accuracy of SAPT(DFT) comparable to SAPT2+ which scales as $O(N^7)$

▶ Iterative $O(N^4)$ terms not negligible; even dominating for smaller systems

Three-Body Interaction Energies

- Cheapest conventional methods to model $E^{\mathrm{int},3}$ scale as $O(N^6)$ [MP2.5, SCS(MI)-CCSD]
- MP2 lacks three-body dispersion interactions
- Combine supermolecular MP2 $E^{\mathrm{int},3}$ with three-body dispersion correction term
- Reduces scaling of computing three-body interaction energies to $O(N^5)$

CKS FDDS Dispersion

▶ Two-body FDDS dispersion term in SAPT(DFT):

$$E_{\text{disp,r}}^{(2)} = -\frac{1}{2\pi} \int_0^\infty d\omega \text{Tr} \left(\mathbf{S}^{-1} \boldsymbol{\chi}^A \mathbf{S}^{-1} \boldsymbol{\chi} \right)$$

▶ Generalize to three-body case:

$$E_{\text{disp,r}}^{(3)} = \int_0^\infty d\omega \, \text{Tr} \left(\mathbf{S}^{-1} \boldsymbol{\chi}^A \mathbf{S}^{-1} \boldsymbol{\chi}^B \mathbf{S}^{-1} \boldsymbol{\chi}^C \right)$$

Axilrod–Teller–Muto Dispersion

► Dispersion energy of atom triplet:

$$E_{\mathrm{ATM}}^{abc} = C_9^{abc} \frac{1 + 3\cos\theta_{\mathrm{a}}\cos\theta_{\mathrm{b}}\cos\theta_{\mathrm{c}}}{(R_{ab}R_{bc}R_{ca})^3}$$

$$C_9^{abc} \approx \sqrt{C_6^{ab}C_6^{bc}C_6^{ca}}$$

► Sum over atom triplets:

$$E_{\mathrm{ATM}}^{ABC} = \sum_{a \in A}\sum_{b \in B}, \sum_{c \in C} f_9^{abc}(\beta) E_{ATM}^{abc}$$

Georgia Tech

3B-69 Benchmark Set

▶ Benchmark set for 3-body interaction energies[0]

▶ 69 trimers extracted from 23 different molecular crystal structures (3 each)

▶ Used focal point approach to obtain CCSD(T) (and other wavefunction method) energies

$$E = E^{\mathrm{HF}}(\mathrm{aQZ}) + \Delta \mathrm{E}^{\mathrm{MP2}}(\mathrm{aTZ/aQZ}) + \Delta \mathrm{E}^{\mathrm{CCSD(T)}}(\mathrm{aDZ})$$

▶ Assessing accuracy for various wavefunction and DFT methods

▶ Authors recommended MP2.5 and SCS-MI-CCSD, both $O(N^6)$

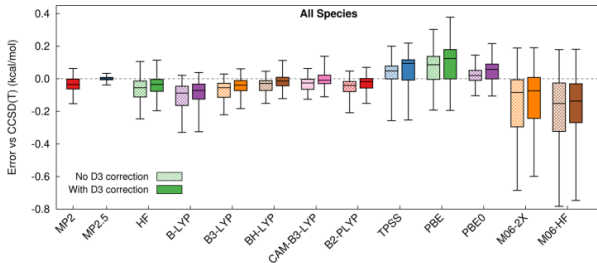▶ We will extend this work to assess the performance of MP2+FDDS (dispersion) for 3B-69 systems

---

[0]J. Řezáč et al., J. Chem. Theory Comput. **11**, 3065 (2015).

3B-69 Wavefunctional Methods
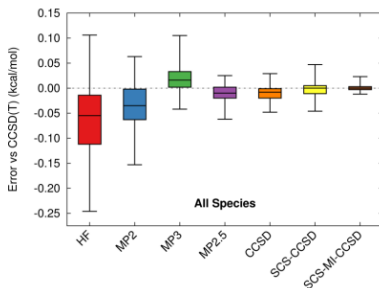
- ▶ MP2.5 and SCS-MI-CCSD exhibits best cost performance, as in the two-body case
- ▶ MP2.5, MP3: Non-iterative $O(N^6)$
- ▶ CCSD and variants: Iterative $O(N^6)$

3B-69 DFT Methods

▶ DFT-D3 accuracies comparable to MP2 at the best, in contrast to the two-body case where DFT-D3 models significantly outperform MP2

▶ Delocalization error leads to errors in many-body polarization and exchange

Georgia Tech

Three-Body FDDS Dispersion

▶ Three-body dispersion energy in terms of FDDS, analogous to the two-body dispersion:

$$
\begin{aligned}
E_{disp,r}^{(3)} &= -\frac{1}{\pi} \int_0^\infty d\omega \int d\mathbf{r}_A d\mathbf{r}_A' d\mathbf{r}_B d\mathbf{r}_B' d\mathbf{r}_C d\mathbf{r}_C' \\
&\quad \frac{1}{|\mathbf{r}_A - \mathbf{r}_B|} \frac{1}{|\mathbf{r}_A' - \mathbf{r}_C|} \frac{1}{|\mathbf{r}_B' - \mathbf{r}_C'|} \\
&\quad \chi^A\left(\mathbf{r}_A, \mathbf{r}_A'|i\omega\right) \chi^B\left(\mathbf{r}_B, \mathbf{r}_B'|i\omega\right) \chi^C\left(\mathbf{r}_B, \mathbf{r}_B'|i\omega\right)
\end{aligned}
$$

▶ Transform from position space into density-fitting auxiliary basis space:

$$
E_{disp,r}^{(3)} = \int_0^\infty d\omega \, \mathrm{Tr}\left(\mathbf{S}^{-1}\boldsymbol{\chi}^A \mathbf{S}^{-1}\boldsymbol{\chi}^B \mathbf{S}^{-1}\boldsymbol{\chi}^C\right)
$$

# New Set from X23

- ▶ Aiming to construct a "three-body version" of S22×5/S66×8 to investigate three-body interaction for trimers with different intermolecular distances and alignments
- ▶ Sampling trimer geometries from X23 crystal structures
- ▶ Distance: Geometric mean and mininum of 3 pairwise closest contact distances
- ▶ Alignment: Angles of the COM triangle; mainly looking at the greatest angle

Research Plan

- ▶ Check if MP2+FDDS (dispersion) is a good model for three-body non-additive interaction energy

- ▶ If not, compare FDDS dispersion with estimated three-body dispersion energy from $E^{\mathrm{CCSD(T)}} - E^{\mathrm{MP2}}$

- ▶ Investigate the dependence of three-body dispersion energies on intermolecular distances and alignments, and the difference between FDDS dispersion and $E^{\mathrm{CCSD(T)}} - E^{\mathrm{MP2}}$ for different trimer geometries

- ▶ Choosing dispersion dominated systems (such as benzene) in X23 to avoid zero dispersion energies at longer distance.
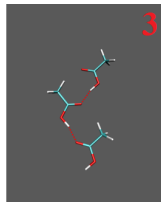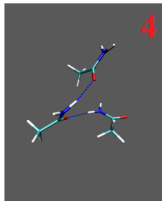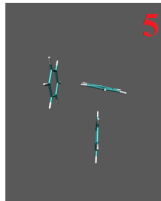
3B-69 Initial Test: Systems
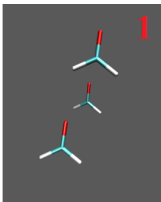
formaldehyde b
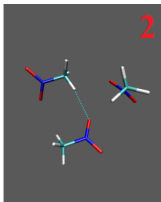
nitromethane a

acetic acid b

acetamide a

benzene c

3B-69 Initial Test: Results

▶ Total three-body non-additive interaction energies in kcal/mol

▶ CCSD(T) and MP2.5 interaction energies from focus point approach

| System	CCSD(T)	MP2+ FDDS/aDZ	MP2+ FDDS/aTZ	MP2	MP2.5
1	0.181	0.207	0.210	0.161	0.179
2	-0.122	-0.069	-0.065	-0.178	-0.143
3	-0.922	-0.905	-0.904	-0.937	-0.913
4	-0.089	-0.003	-0.003	-0.239	-0.151
5	-0.027	0.002	0.003	-0.061	-0.023

3B-69 Initial Test: Results

- Comparison of estimated 3-body dispersion energies
- ΔCCSD(T) corresponds to dispersion energy estimated by $E^{\mathrm{CCSD(T)}} - E^{\mathrm{MP2}}$

System	ΔCCSD(T)	FDDS/aDZ	FDDS/aTZ
1	0.020	0.046	0.049
2	0.056	0.109	0.113
3	0.015	0.032	0.033
4	0.150	0.236	0.242
5	0.034	0.063	0.064

3B-69 Three-Body Dispersion: $\triangle$MP2.5
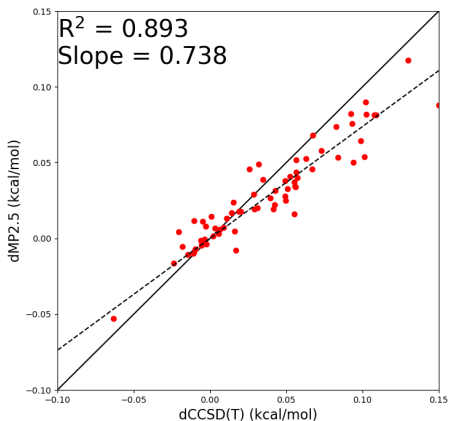
3B-69 Three-Body Dispersion: FDDS

$R^2 = 0.859$
Slope = 1.796
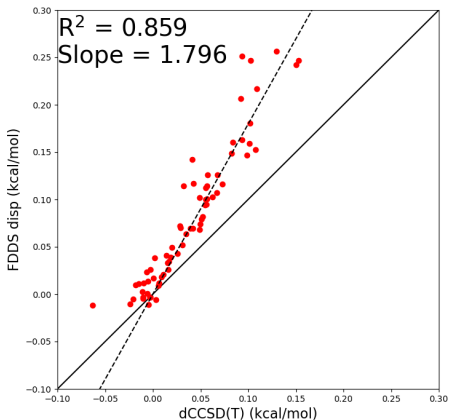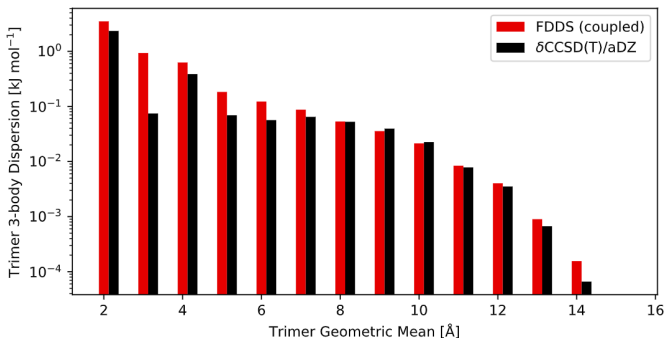
Crystalline Benzene 3-Body Dispersion
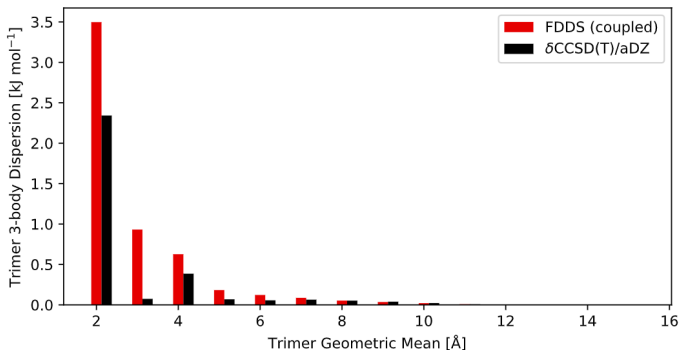
Crystalline Benzene 3-Body Dispersion

Acknowledgment

Thesis advisor:

       Dr. C. David Sherrill

Thesis committee members:

       Dr. Jesse G. McDaniel

       Dr. Thomas M. Orlando

       Dr. Edmond Chow

       Dr. Joshua S. Kretchmer

Special thanks for implementing 3-body FDDS dispersion and providing data for crystalline benzene:

       Zachary L. Glick