

Data Review

Yi Xie

May 16, 2022

Table of contents

- 1 Overview
- 2 SAPT(DFT) Implementation
 - Theory
 - Results
- 3 Three-Body FDDS Dispersion
 - Background
 - Results

Intermolecular Energies

► Supermolecular approach

$$E_{int} = E_{AB} - E_A - E_B$$

- Straightforward, but cannot separate different types of interactions
 - Can adopt to different electronic structure methods
 - DFT-D3 with proper functional can be both cheap and accurate
- Symmetry-Adapted Perturbation Theory
- Can give details about different types of interactions; important in understanding their nature
 - Not as cheap as DFT-D3
 - SAPT0 is somewhat cheap, but does not include intramonomer correlation

SAPT(DFT)

- ▶ Attempt to include intramonomer correlation in a cheap way
- ▶ Replaces HF orbitals with KS orbitals
- ▶ Needs to consider orbital response for dispersion terms
- ▶ Exchange-dispersion term needs to be estimated from scaling
- ▶ Investigate the accuracy and efficiency of SAPT(DFT)

Three-Body Interaction

- ▶ Crucial in computing lattice energies
- ▶ DFT-D3 does not perform well for three-body interaction
- ▶ MP2.5 scales as $O(N^6)$, MP2 is $O(N^5)$ but lacks three-body dispersion
- ▶ Three-body dispersion can be implemented with SAPT(DFT) in $O(N^5)$
- ▶ Combine MP2 with SAPT(DFT) dispersion to model three-body interaction

Table of contents

- 1 Overview
- 2 SAPT(DFT) Implementation
 - Theory
 - Results
- 3 Three-Body FDDS Dispersion
 - Background
 - Results

Table of contents

- 1 Overview
- 2 SAPT(DFT) Implementation
 - Theory
 - Results
- 3 Three-Body FDDS Dispersion
 - Background
 - Results

Idea of SAPT(DFT)

- ▶ SAPT energy in orders of interaction and fluctuation potentials; n denotes order in V and k, l for W_A, W_B

$$H = F_A + F_B + V + W_A + W_B$$

$$E_{int} = \sum_{n=1}^{\infty} \sum_{k=0}^{\infty} \sum_{l=0}^{\infty} \left(E_{pol}^{(nkl)} + E_{exch}^{(nkl)} \right)$$

- ▶ SAPT0: $n = 2, k = l = 0$, no intramonomer correlation, $O(N^5)$ cost
- ▶ Many-body SAPT: $k + l \geq 2$, $O(N^6)$ or higher cost
- ▶ SAPT(DFT): Use Kohn-Sham operator $K_{A,B}$ instead of Fock operator $F_{A,B}$, $O(N^5)$ cost
- ▶ Primitive SAPT(DFT) works well on 1st-order terms, but not 2nd-order terms (especially dispersion). Needs orbital response for these terms

Coupled Dispersion Energy

- Uncoupled dispersion energy in terms of frequency-dependent density susceptibility (FDDS):

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

- Kohn-Sham DFT constructs a fictitious system of non-interacting particles, which reproduces the density and energy of the real electronic system
- Kohn-Sham FDDS does not reflect the correct response properties of the electronic system



Dispersion Term

- ▶ Coupled FDDS from solving TDDFT equations:

$$\chi = \chi_0 + \chi_0 \mathbf{S}^{-1} \mathbf{W} (\mathbf{S} - \chi_0 \mathbf{S}^{-1} \mathbf{W})^{-1} \chi_0$$

- ▶ \mathbf{S} and \mathbf{W} corresponds to Coulomb metric and xc kernel
- ▶ Pure ALDA kernel good for pure GGA functional, but not for hybrid functional
- ▶ Exact exchange in $v_{xc} \rightarrow$ increased $\epsilon_{ij}^{ab} \rightarrow$ decreased $E_{disp}^{(2)}$

$$E_{disp,u}^{(2)} = -4 \sum_{ia \in A, jb \in B} \frac{|(ia|jb)|^2}{\epsilon_{ij}^{ab}}$$

- ▶ Hybrid ALDA kernel to compensate, or localized HF (LHF) exchange to avoid increase in ϵ_{ij}^{ab}

Coupled FDDS with hybrid kernel

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

$$\mathbf{K}' = [-\xi \mathbf{K}_1(\lambda d) - \xi \mathbf{K}_2(\lambda d) + \xi^2 \mathbf{K}_{21}(\lambda)] (\mathbf{R}^T)^{-1} \mathbf{S}$$

$$[\mathbf{K}_1(\lambda d)]_{PQ} = (P|ar)\lambda_{ar}d_{ar}[(aa'|rr') + (ar'|a'r)](a'r'|\mathbf{Q}|Q)$$

$$[\mathbf{K}_2(\lambda d)]_{PQ} = (P|ar)\lambda_{ar}d_{ar}[(aa'|rr') - (ar'|a'r)](a'r'|\mathbf{Q}|Q)$$

$$\begin{aligned} [\mathbf{K}_{21}(\lambda)]_{PQ} = & (P|ar)\lambda_{ar}[(aa''|rr'') - (ar''|a''r)] \\ & [(a'a''|r'r'') - (a'r''|a''r')](a'r'|\mathbf{Q}|Q) \end{aligned}$$

$$[\mathbf{K}'_2(\lambda)]_{PQ} = (P|ar)\lambda_{ar}[(aa'|rr') - (ar'|a'r)](a'r'|\mathbf{Q})$$

$$\chi'_0 = \chi_0 - \xi \mathbf{K}_2(\lambda)$$

$$(ar|Q) = (ar|\mathbf{Q}|P)(P|\mathbf{R}|Q)$$

Equation for $E_{disp}^{(2)}$

- ▶ Coupled Kohn-Sham (CKS) FDDS reflects correct response properties of electrons
- ▶ $O(N^5)$ scaling is limited to forming \mathbf{K}_1 , \mathbf{K}_2 , \mathbf{K}_{21} and \mathbf{K}'_2
- ▶ Separates “nontrivial” and “trivial” parts of the code, also highly reduces need of disk I/O operations
- ▶ $E_{disp}^{(2)}$ from coupled FDDS:

$$E_{disp,u}^{(2)} = -\frac{1}{2\pi} \int_0^\infty d\omega \int d\mathbf{r}_A d\mathbf{r}'_A d\mathbf{r}_B d\mathbf{r}'_B \frac{1}{|\mathbf{r}_A - \mathbf{r}_B|} \frac{1}{|\mathbf{r}'_A - \mathbf{r}'_B|} \chi^A(\mathbf{r}_A, \mathbf{r}'_A | i\omega) \chi^B(\mathbf{r}_B, \mathbf{r}'_B | i\omega)$$



Exchange-Dispersion Term

- ▶ Explicit coupled exchange-dispersion not trivial to implement; currently working on this
- ▶ Estimate from scaling uncoupled exchange-dispersion; Scale with ratio in dispersion term or with pre-fitted (with S22×5) fixed factor

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

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

- ▶ The value above is fitted from $E_{exch-disp,u}^{(2)}$ with LHF orbitals
- ▶ Non-LHF orbitals have greater o-v gaps and smaller $E_{disp,u}^{(2)}$, needs to re-fit with non-LHF results

LHF vs non-LHF orbitals

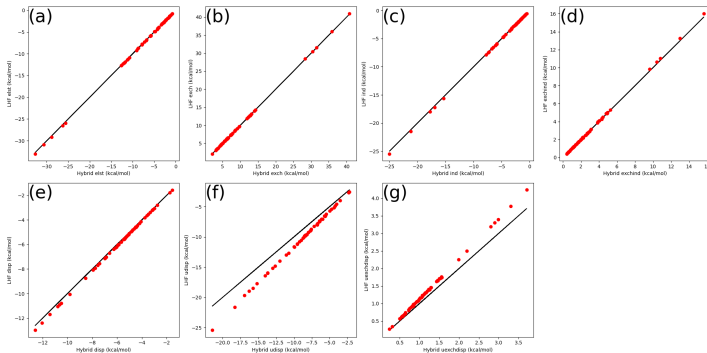


Figure: Hybrid vs. LHF values in kcal/mol for each term for S66 data set: (a) $E_{elst}^{(1)}$, (b) $E_{exch}^{(1)}$, (c) $E_{ind}^{(2)}$, (d) $E_{exch-ind}^{(2)}$, (e) $E_{disp,r}^{(2)}$, (f) $E_{disp,u}^{(2)}$, (g) $E_{exch-disp,u}^{(2)}$



GRAC

- ▶ (Hybrid-)GGA functionals does not have correct long-range behavior
 $v_{xc}(r) \rightarrow -1/r + (I_p + \epsilon_{\text{HOMO}})$
- ▶ Underestimates o-v gap as a consequence
- ▶ Functionals like LB94 have correct asymptotic behavior, but poor in bulk region
- ▶ Using gradient-regulated asymptotic correction (GRAC) scheme to connect PBE0 and LB94 with the switching function $f[g(\mathbf{r})]$:

$$v_{xc}^{\text{GRAC}} = \{1 - f[g(\mathbf{r})]\} v_{xc}^{\text{PBE0}} + f[g(\mathbf{r})] v_{xc}^{\text{LB94}}$$

$$f[g(\mathbf{r})] = \left(1 + e^{-\alpha[g(\mathbf{r}) - \beta]}\right)^{-1}$$

$$g(\mathbf{r}) = \frac{|\nabla \rho(\mathbf{r})|}{\rho^{4/3}(\mathbf{r})}$$

GRAC & Long-Range Behavior

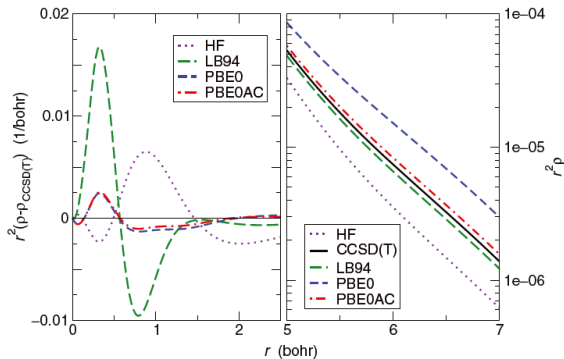


Figure: Radial densities $r^2\rho(r)$ of Ne atom (right) and errors compared to CCSD(T) density (left) for various xc potentials.¹

¹G. Jansen, WIREs Comput. Mol. Sci. **4**, 127 (2014).

Table of contents

- 1 Overview
- 2 SAPT(DFT) Implementation
 - Theory
 - Results
- 3 Three-Body FDDS Dispersion
 - Background
 - Results

Exchange-Dispersion Refitting

- ▶ Need to fit the uncoupled $E_{exch-disp,r}^{(2)}$ with non-LHF orbitals on the coupled LHF orbital values (implemented in Molpro)
- ▶ Assuming coupled LHF and non-LHF orbital $E_{exch-disp,r}^{(2)}$ from the behavior of $E_{disp,r}^{(2)}$
- ▶ Exchange-related components depend heavily on distance between monomers, sets like S22×5 and S66×8 would be preferred
- ▶ Determine the scaling factor with S22×5, validate with S66×8



S22 \times 5 Fitting Results

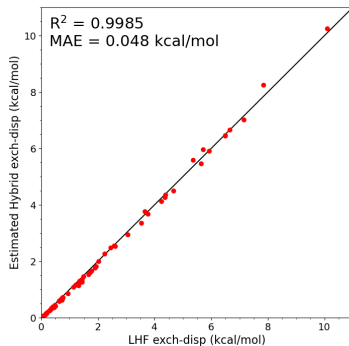
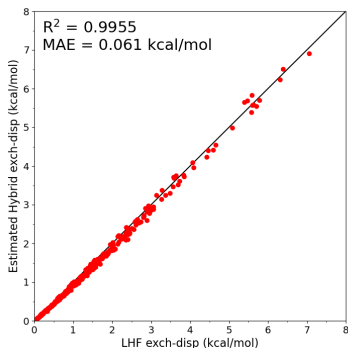


Figure: Scaling Factor = 0.770

S66×8 Validating Results





Termwise results

- ▶ Compared the SAPT(DFT)/aug-cc-pVTZ results of our code to SAPT(CCSD)/aug-cc-pVTZ results from Korona S2², which consists of 14 small dimers
- ▶ Also comparing the results for S66. SAPT(CCSD) results are not available, used SAPT2+3(CCD) δ MP2/aug-cc-pVTZ as reference. Also added SAPT0/aug-cc-pVDZ, SAPT2+/aug-cc-pVDZ and SAPT2+(3) δ MP2/aug-cc-pVTZ into comparison as side-reference.
- ▶ Errors of each system with respect to reference shown as vertical lines
- ▶ Mean absolute error (MAE) and mean unsigned relative error (MURE) listed for S2
- ▶ Color scheme for S66 systems: Hydrogen-bonded (HB, red), mixed-influence (MX, green), dispersion-dominated (DD, blue)

²T. Korona, Mol. Phys. **111**, 3705 (2013).

Korona S2 Results

Method ^a	MAE	MURE	Error Distribution ^b					
			4	OB	1	0	1	UB
Electrostatics								
SAPT(DFT) hybrid	0.112	2.39					█	
SAPT(DFT) LHF	0.114	3.68					█	
SAPT0	0.520	8.61					█	
Exchange								
SAPT(DFT) hybrid	0.251	3.38					█	
SAPT(DFT) LHF	0.258	3.09					█	
SAPT0	1.757	12.88	█	█				
Induction								
SAPT(DFT) hybrid	0.148	2.79					█	
SAPT(DFT) LHF	0.192	2.97					█	
SAPT0	1.993	16.83		█				
Exchange-Induction								
SAPT(DFT) hybrid	0.144	4.03					█	
SAPT(DFT) LHF	0.165	4.76					█	
SAPT0	1.551	26.80	█	█				

Korona S2 Results

Dispersion			
SAPT(DFT) hybrid	0.175	3.68	
SAPT(DFT) LHF	0.141	2.77	
SAPT(DFT) non-hybrid	0.326	9.58	
SAPT0	0.811	24.86	
Exchange-Dispersion			
SAPT(DFT) hybrid	0.062	12.47	
SAPT(DFT) LHF	0.039	3.25	
SAPT0	0.265	36.11	
Total			
SAPT(DFT) hybrid	0.155	4.98	
SAPT(DFT) LHF	0.189	4.17	
SAPT(DFT) non-hybrid	0.244	10.64	
SAPT0	1.237	19.63	

S66 Results

Method ^a	Total	HB	MX	DD	Error Distribution ^b					
					4	OB	1	0	1	UB
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) δ MP2	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) δ MP2	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) δ MP2	0.152	0.179	0.121	0.145						

S66 Results

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) δ MP2	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) δ MP2	0.105	0.056	0.082	0.169	██████████

Timing Performance

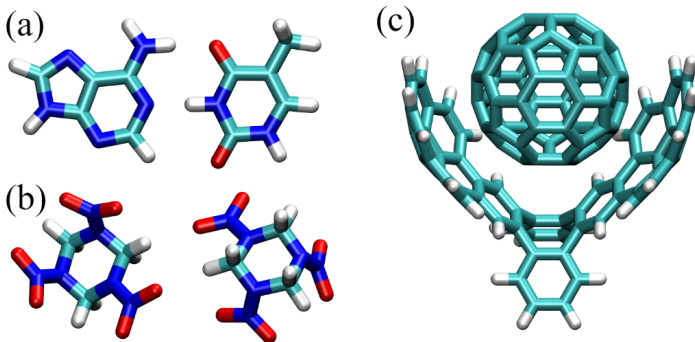
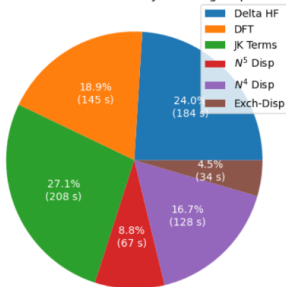


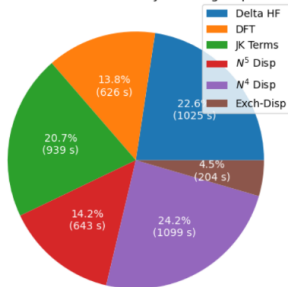
Figure: Dimer systems for timing: (a) Watson-Crick adenine-thymine complex, (b) RDX dimer, (c) C₆₀-buckycatcher complex.

Watson-Crick Adenine-Thymine

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

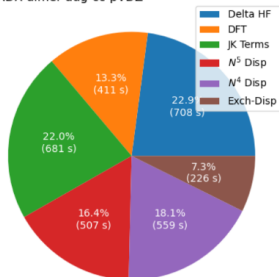


(b) Watson-Crick Adenine-Thymine aug-cc-pVTZ

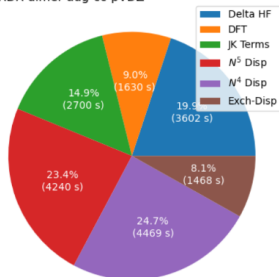


RDX Dimer

(c) RDX dimer aug-cc-pVDZ



(d) RDX dimer aug-cc-pVDZ



Comparison with LHF Approach

TABLE IV. Wall times (in hours) for SAPT(DFT) computations of RDX dimer/aug-cc-pVTZ with hybrid and LHF algorithm.

Subroutine	Hybrid	LHF
Delta HF	0.96	N/A ^a
DFT	0.45	2.29
xc kernel	0.08	4.17
FDOS object ^b	2.35	N/A
Disp time integration	0.37	3.59
Exch-disp	0.41	1.99
Total	5.03	12.80

^a The δ HF correction, recommended for SAPT(DFT) computations of polar molecules, is performed by default in Psi4 but not in Molpro.

^b Including integral transformation, form X/form Y (the $O(N^5)$ part) and QR factorization. In Molpro, the integral transformation is integrated with other terms, and the other steps are not relevant for LHF.



C_{60} –Buckycatcher Complex

- ▶ $N_{bf} = 3012, N_{aux} = 9284$ with aug-cc-pVDZ basis set
- ▶ Using Intel i9-10980XE processor with 18 cores, completed entire calculation in 4.03 days
- ▶ 42.7 hours for $E_{disp}^{(2)}$; 20.1 hours for the $O(N^5)$ subroutines
- ▶ Cost of DFT and $E_{ind}^{(2)}$ still somewhat significant

Table of contents

- 1 Overview
- 2 SAPT(DFT) Implementation
 - Theory
 - Results
- 3 Three-Body FDDS Dispersion
 - Background
 - Results

Table of contents

- 1 Overview
- 2 SAPT(DFT) Implementation
 - Theory
 - Results
- 3 Three-Body FDDS Dispersion
 - Background
 - Results



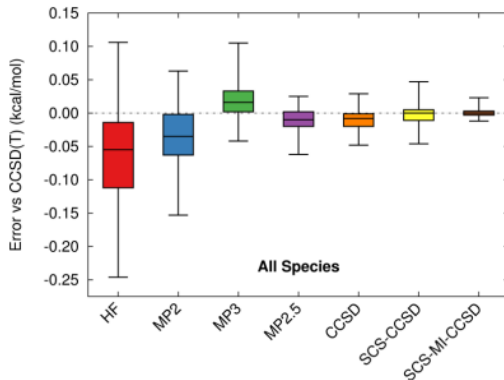
3B-69 Benchmark Set

- ▶ Benchmark set for 3-body interaction energies
- ▶ 69 trimers extracted from 23 different molecular crystal structures (3 each)
- ▶ Used focal point approach to obtain CCSD(T) (and other wavefunction method) energies

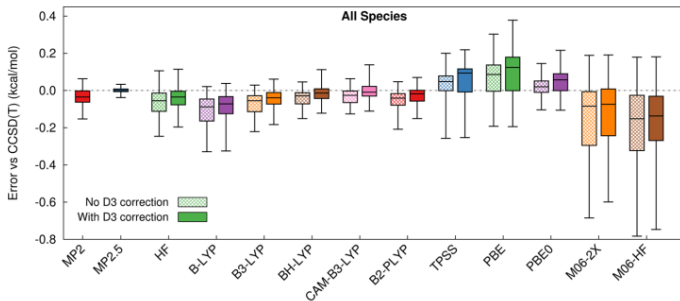
$$E = E^{HF}(aQZ) + \Delta E^{MP2}(aTZ/aQZ) + \Delta E^{CCSD(T)}(aDZ)$$

- ▶ Assessing accuracy for various wavefunction and DFT methods
- ▶ Recommended MP2.5 and SCS-MI-CCSD, both $O(N^6)$
- ▶ Trying to assess performance of MP2 + FDDS dispersion for 3B-69 systems

3B-69 Wavefunctional Methods



3B-69 DFT Methods



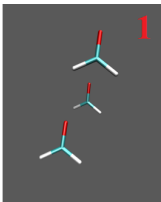
New Set from X23

- ▶ Sampling trimer geometries from X23 crystal structures
- ▶ Trying to include trimer with different intermolecular distances and alignment
- ▶ Aiming to serve as a "three-body version" of $S22 \times 5 / S66 \times 8$
- ▶ Distance: Geometry mean of 3 pairwise closest contact distance
- ▶ Alignment: Angles of the COM triangle; mainly looking at the greatest angle

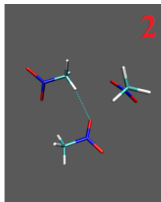
Table of contents

- 1 Overview
- 2 SAPT(DFT) Implementation
 - Theory
 - Results
- 3 Three-Body FDDS Dispersion
 - Background
 - Results

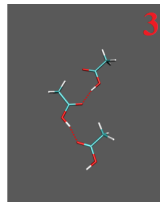
3B-69 Systems



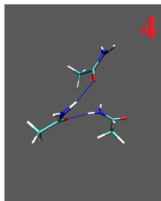
formaldehyde b



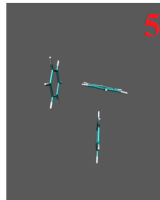
nitromethane a



acetic acid b



acetamide a



benzene c

3B-69 Results

- ▶ Total interaction energies in kcal/mol
- ▶ FDDS(aDZ) and FDDS(aTZ) correspond to estimated total IE with MP2 + FDDS

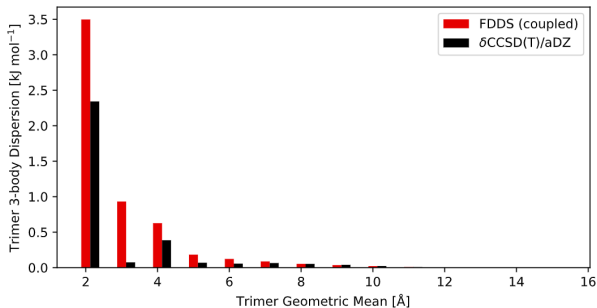
System	CCSD(T)	FDDS(aDZ)	FDDS(aTZ)	MP2	MP2.5
1	0.181	0.207	0.210	0.161	0.179
2	-0.122	-0.069	-0.065	-0.178	-0.143
3	-0.922	-0.905	-0.904	-0.937	-0.913
4	-0.089	-0.003	-0.003	-0.239	-0.151
5	-0.027	0.002	0.003	-0.061	-0.023

3B-69 Results

- ▶ Estimated 3-body dispersion energies comparison
- ▶ CCSD(T) corresponds to dispersion energy estimated by $E^{CCSD(T)} - E^{MP2}$

System	CCSD(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

Distance Dependence



Distance Dependence

