Data Review

Yi Xie

April 14, 2022

Table of contents

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 inlude 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

Table of contents

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: $k, l \geq 2$, $O(N^7)$ 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 them

Dispersion Term

- ALDA kernel good for pure GGA functional, but not for hybrid functional
- Exact exchange in v_{xc} $\rightarrow$ increased ϵ_{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}

$$f_{xc} = \alpha f_{xc}^{HF} + (1-\alpha) f_{xc}^{ALDA}$$

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

- Computing coupled FDDS:

$$\chi = \chi_0' + \left(\chi_0' \mathbf{S}^{-1} \mathbf{W} + \mathbf{K}\right) \left[\mathbf{S} - \left(\chi_0' \mathbf{S}^{-1} \mathbf{W} + \mathbf{K}\right)\right]^{-1} \chi_0'$$

$$\chi_0' = \chi_0 - \alpha \mathbf{K}_2 (\lambda)$$

$$\mathbf{K} = \left[-\alpha \mathbf{K}_1 (\lambda d) - \alpha \mathbf{K}_2 (\lambda d) + \alpha^2 \mathbf{K}_{21} (\lambda)\right] \left(\mathbf{R}^t\right)^{-1} \mathbf{S}$$

- $O(N^5)$ scaling is limited to forming $\mathbf{K}_1$, $\mathbf{K}_2$ and $\mathbf{K}_{21}$
- Separates nontrivial and trivial parts of the code, also highly reduces need of disk I/O operations
- $E_{disp}^{(2)}$ from coupled FDDS:

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

Exchange-Dispersion Term

- Coupled exchange-dispersion requires storing 4-index tensors on disk
- Estimate from scaling uncoupled exchange-dispersion; Scale with ratio in dispersion term or with pre-fitted (with S22×5) fixed factor

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

$$\tilde{E}^{(2)}_{exch-disp,r} = \alpha \cdot E^{(2)}_{exch-disp,u} (\alpha = 0.686361)$$

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

LHF vs non-LHF orbitals

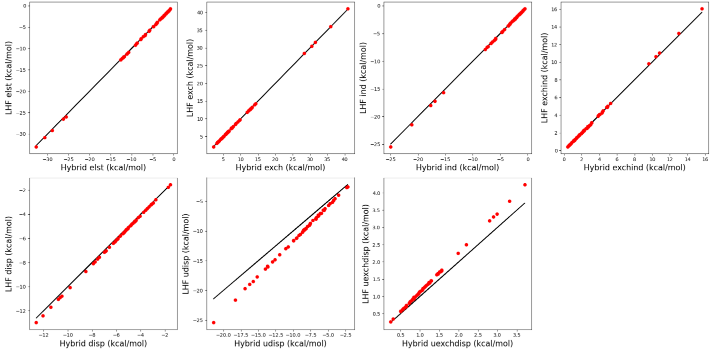


Figure: Hybrid vs. LHF values for each term for S66 data set ($E^{(1)}_{elst}$, $E^{(1)}_{exch}$, $E^{(2)}_{ind}$, $E^{(2)}_{exch-ind}$, $E^{(2)}_{disp,r}$, $E^{(2)}_{disp,u}$, $E^{(2)}_{exch-disp,u}$)

GRAC

- (Hybrid-)GGA functionals does not have correct long-range behavior $v_{xc}(r) \to -1/r + (I_p + \epsilon_{\text{HOMO}})$
- 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 $f[g(\boldsymbol{r})]$:

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

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

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

GRAC & Long-Range Behavior

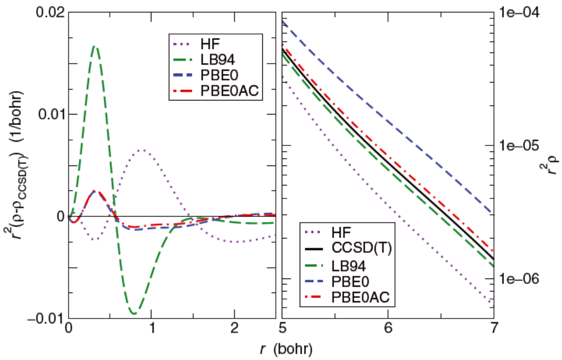


Figure: Radial densities $r^2\rho(r)$ of Ne atom (right) and errors compared to CCSD(T) density (left) for various xc potentials.[1]

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

Table of contents

Exchange-Dispersion Refitting

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

S22×5 Fitting Results

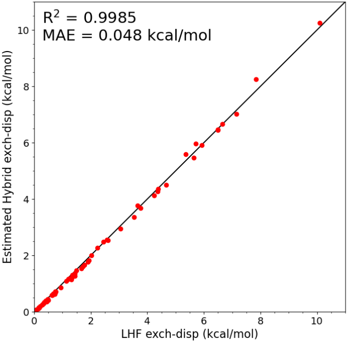
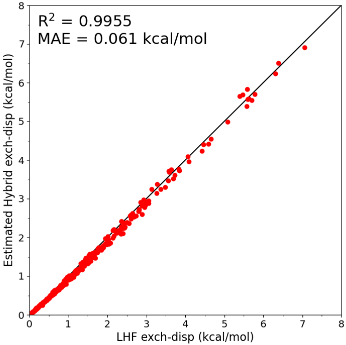


Figure: Scaling Factor = 0.770

S66×8 Validating Results



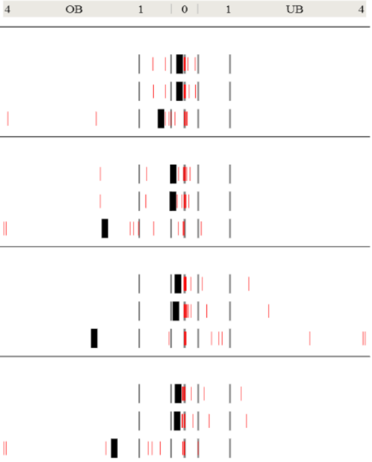
Termwise results

- Compared the SAPT(DFT)/aug-cc-pVTZ results of our code to SAPT(CCSD)/aug-cc-pVTZ results from Korona S2[2], which consists of 14 small dimers
- Also comparing the results for S66. SAPT(CCSD) results are not available, used SAPT2+3(CCD)δMP2/aug-cc-pVTZ as reference. Also added SAPT0/aug-cc-pVDZ, SAPT2+/aug-cc-pVDZ and SAPT2+(3)δMP2/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)

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

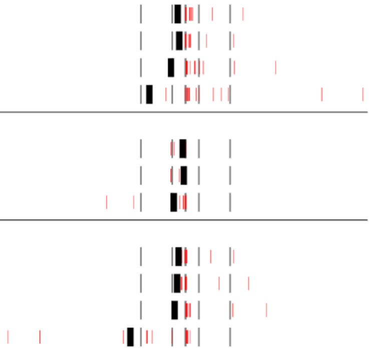
Korona S2 Results

Method[a]	MAE	MURE	Error Distribution[b]
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	



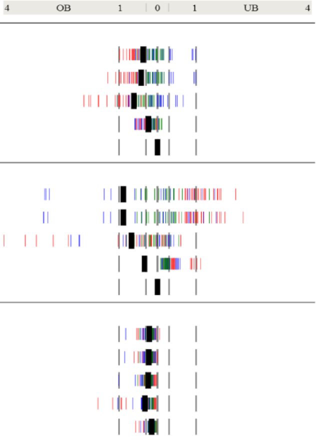
Korona S2 Results

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) hon-hybrid	0.244	10.64
SAPT0	1.237	19.63



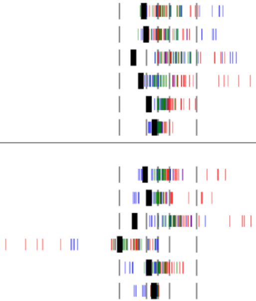
S66 Results

Method[a]	Total	HB	MX	DD	Error Distribution[b]
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

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 Performance

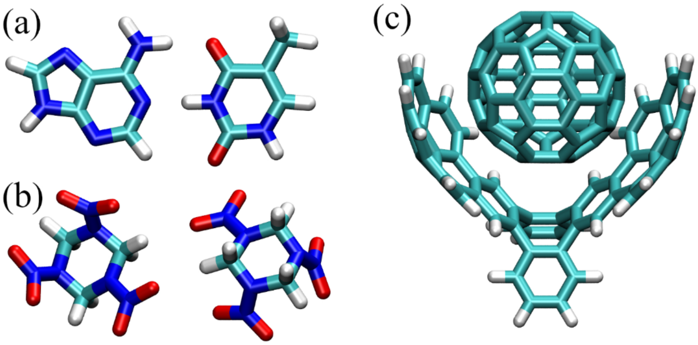
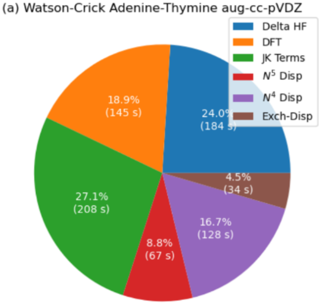
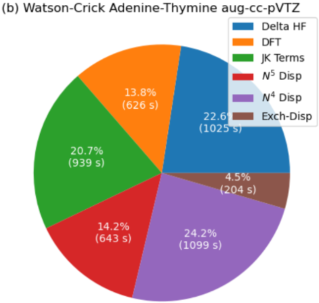


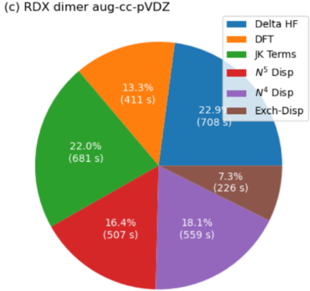
Figure: Dimer systems for timing: (a) Watson-Crick adenine-thymine complex, (b) RDX dimer, (c) C_{60}–buckycather complex.

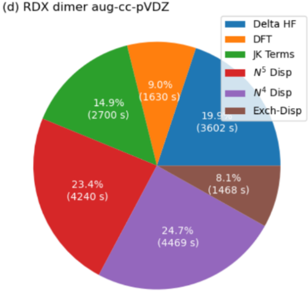
Watson-Crick Adenine-Thymine





RDX Dimer


(c) RDX dimer aug-cc-pVDZ
Delta HF
DFT
JK Terms
N^5 Disp
N^4 Disp
Exch-Disp
13.3%
(411 s)
22.9
(708 s)
22.0%
(681 s)
7.3%
(226 s)
16.4%
(507 s)
18.1%
(559 s)


(d) RDX dimer aug-cc-pVDZ
Delta HF
DFT
JK Terms
N^5 Disp
N^4 Disp
Exch-Disp
9.0%
(1630 s)
14.9%
(2700 s)
19.9
(3602 s)
23.4%
(4240 s)
8.1%
(1468 s)
24.7%
(4469 s)

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
FDDS 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 δHF 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 $E^{(2)}_{disp}$; 20.1 hours for the $O(N^5)$ subroutines
- Cost of DFT and $E^{(2)}_{ind}$ still somewhat significant

Table of contents

Table of contents

3B-69 Data Set

TODO