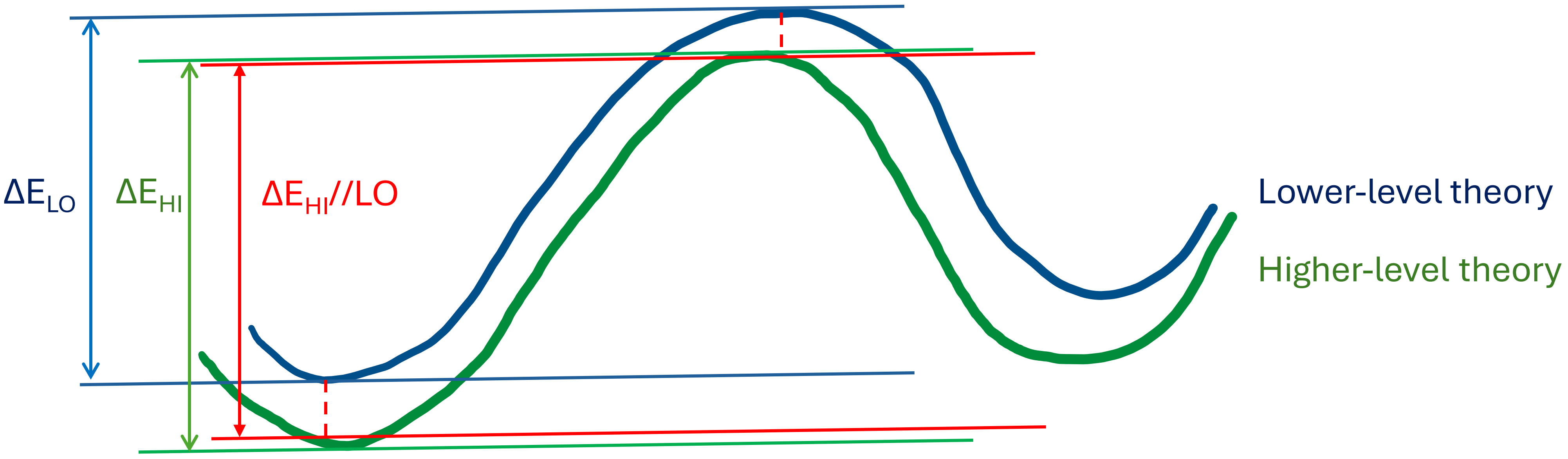
Single-Point Energies

(and the importance of knowing what geometry you are using)

C. David Sherrill

Georgia Institute of Technology

Single-Point Energies



- Barrier ΔE is slightly underestimated by lower-level theory
- Lower-level theory gives a good approximate location (geometry) for reactant and transition state
- Compute ΔE at higher level of theory, using lower-level geometries (b/c geometries expensive to optimize)
- Notation //LO means a higher level of theory was used at particular (single) geometry points obtained using a lower level of theory (LO)

Single-Point Energy Notation

- A//B means that a single point energy was computed using level of theory A at geometries obtained using level of theory B
- A “level of theory” (or “model chemistry”) means a combination of a theoretical method and a basis set, and can be denoted level-of-theory/basis-set (note single not double slash)
- Theoretical methods: like HF (Hartree-Fock), B3LYP, MP2, CCSD(T), etc.
- Basis set: like 6-31G*, aug-cc-pVDZ, etc.
- Thus, CCSD(T)/aug-cc-pVTZ // B3LYP/6-31G* means we used B3LYP in a 6-31G* basis to obtain the geometry, and then at that geometry we computed the energy with the higher-level method CCSD(T) and the aug-cc-pVTZ basis
- Usually both the method and basis set improve in A vs B, but not necessarily --- could improve only the method or only the basis

Cost Savings

- A gradient might take ~2x times as long to compute as an energy
- Let's say the energy of the high-level method might take ~10x as long to compute as the energy of the low-level method (the gap will depend on the methods and size of molecule and basis set)
- A geometry optimization may take on the order of 10 optimization steps (often more)
- Let Q be the cost of an energy with the low-level method
- The cost of a single-point high-level energy at the low-level geometry is 10Q + 20Q = 30Q
- The cost of the high-level energy at the high-level geometry is 200Q (geometry optimization gives us the energy as a byproduct)
- Cost savings of single-point procedure is 170Q, or (170Q/200Q) = 85%

Accuracy Benefits

- As in our graphical illustration, the geometry at the lower-level method is typically rather good
- Energies normally change fairly slowly with respect to geometry changes, near stationary points
- Thus, minor changes in geometry don’t normally lead to large changes in energy
- Thus, lower-level geometries are normally just fine for computing higher-level energies: the error in the energy due to the slight error in geometry is normally negligible
- DFT is surprisingly good at computing accurate geometries (B3LYP is popular). (Note: for van der Waals complexes, need –D corrections to get decent geometries, e.g., B3LYP-D3).
- DFT can often underestimate barrier heights (by ~1-5 kcal/mol), so it can pay to compute single-point energies at a higher level to reduce these errors
- DFT can also exhibit significant errors in reaction energies, conformational energies, excitation energies, etc.

Importance of Conveying the Geometry Used

- Because of the enormous computational cost benefits of computing high-level single-point energies at low-level geometries, this practice is very common
- This means it is very important to clarify whether you are computing high-level energies consistently at high-level geometries, or at lower-level geometries
- Whenever lower-level geometries are used, this should be clarified (often with the A//B notation); otherwise, other people will assume you got the geometries at the higher level of theory
- Saying you computed CCSD(T)/aug-cc-pVTZ barrier heights for a 30 atom system will sound very suspicious; saying you computed CCSD(T)/aug-cc-pVTZ // B3LYP/cc-pVDZ barrier heights for that system sounds much more plausible

Example: Single Points for van der Waals dimers

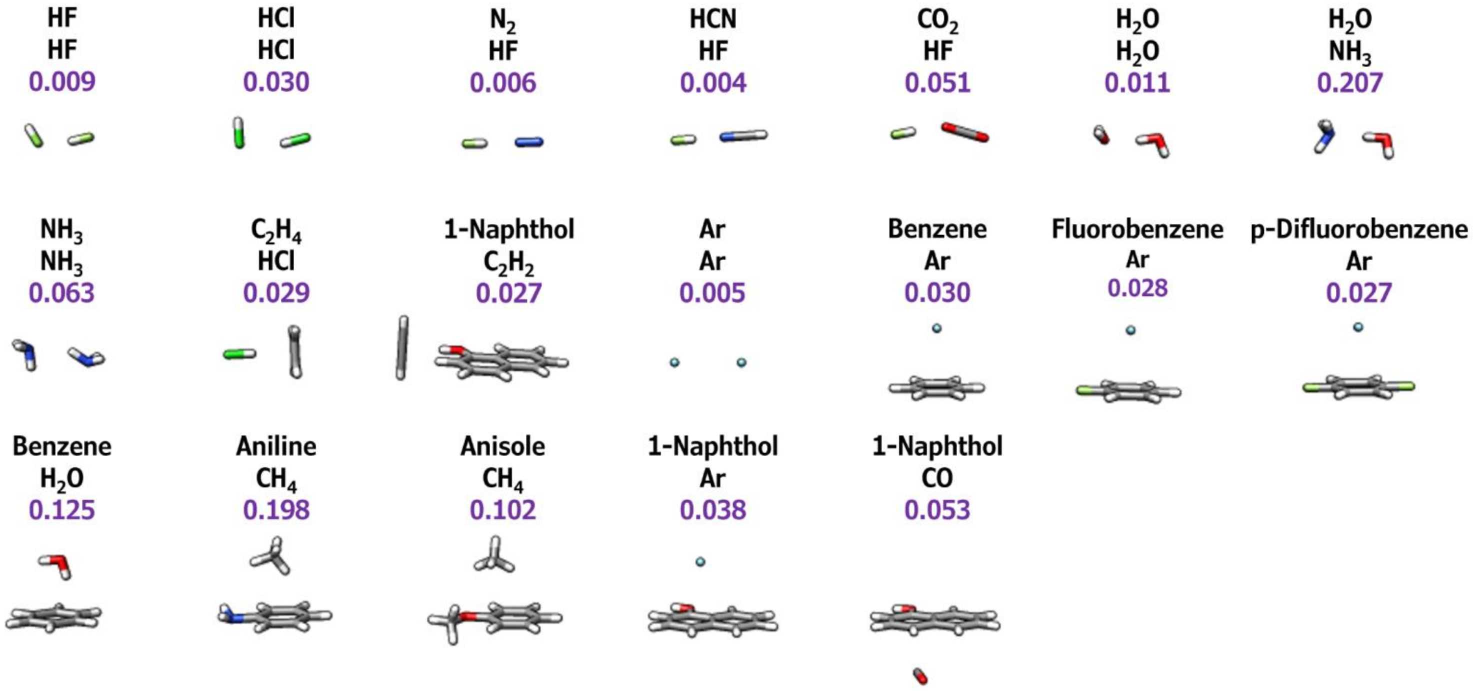
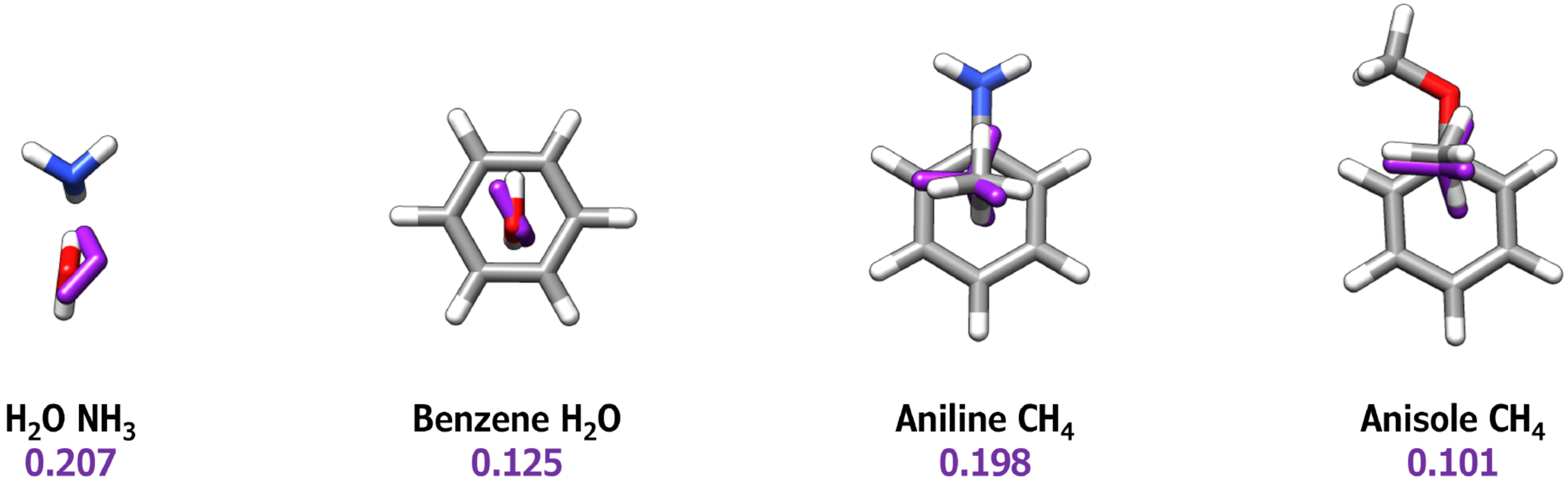


FIG. 1: CCSD(T)/CBS(a([TQ]+d)Z; δ:a(D+d)Z) focal-point geometries of all 19 van der Waals dimers. Values above each geometry are the computed RMSD (Å) against the MP2/aug-cc-pV(T+d)Z geometries.

L. A. Kurfman, J. Yon, P. M. Nelson, and C. D. Sherrill, in preparation



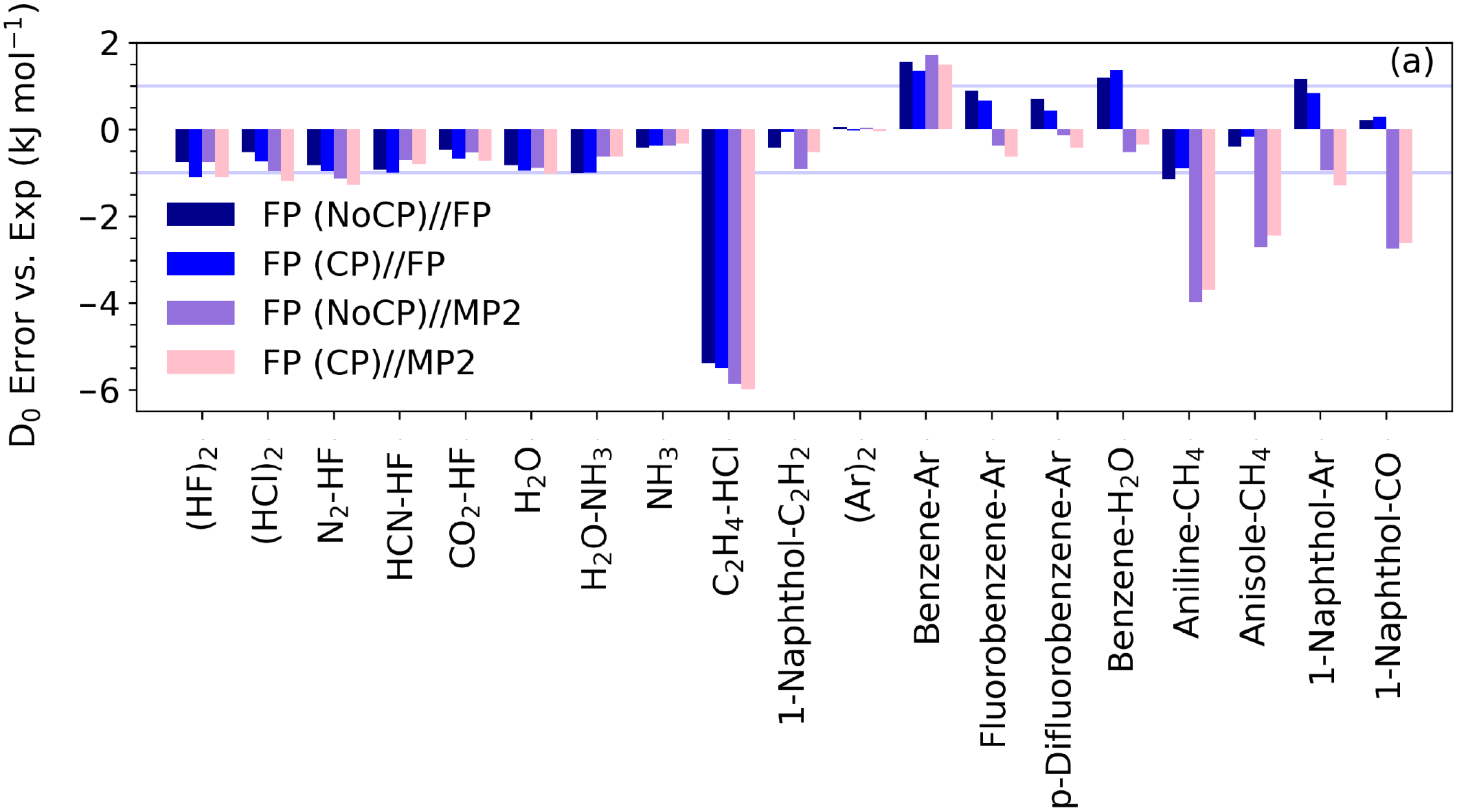
H_2O NH_3
0.207

Benzene H_2O
0.125

Aniline CH_4
0.198

Anisole CH_4
0.101

FIG. 2: All CCSD(T)/CBS(a([TQ]+d)Z; δ:a(D+d)Z) focal-point geometries with an RMSD ≥ 0.1 Å compared to their MP2/aug-cc-pV(T+d)Z counterparts. MP2 geometries are overlaid (purple), along with RMSD values (Å) in purple.

D_0 Error vs. Exp (kJ mol^{-1})
2
0
−2
−4
−6
(a)
FP (NoCP)//FP
FP (CP)//FP
FP (NoCP)//MP2
FP (CP)//MP2
$(HF)_2$
$(HCl)_2$
N_2-HF
HCN-HF
CO_2-HF
H_2O
H_2O-NH_3
NH_3
C_2H_4-HCl
1-Naphthol-C_2H_2
$(Ar)_2$
Benzene-Ar
Fluorobenzene-Ar
p-Difluorobenzene-Ar
Benzene-H_2O
Aniline-CH_4
Anisole-CH_4
1-Naphthol-Ar
1-Naphthol-CO