Lab 4: Electron Correlation
CHE 525/PHY 567, Spring 2021, Stony Brook University

1 Introduction

Choosing an appropriate treatment of electron correlation is key to successfully applying quantum chemical methods to real chemical problems. In this exercise, you will explore different treatments of electron correlation to compute PESs for dissocation of three simple molecules/complexes. Your goal is to determine which methods provide an accurate desciption of dissociation and which do not. You will then explain these results based on concepts that we learned in class.

2 Procedure

A useful side effect of this exercise is that you will learn to submit individual jobs to the scheduler on Seawulf so that they will run asynchronously. This allows you to avoid requesting (and waiting for) an interactive session on a compute node, which is convenient for several reasons. You can set up and submit multiple calculations simultaneously, and rather than waiting for them to complete, you can log out of the machine and come back later. This is nice when you have many jobs to run, when the jobs will take a long time, and/or when queue wait times are long.

If you have not used the scheduler in this way before, you will want to read up. This web page has useful information on submitting to the SLURM scheduler on Seawulf: `https://it.stonybrook.edu/help/kb/using-the-slurm-workload-manager`

Now try it yourself using an example batch script we provide. You will find two files in the repository: `Lab4starter-HF.py` and `Lab4starter-HF.sbatch`. After logging into Seawulf, you will run `module load slurm` as you usually would, but do not request an interactive session on a node. Edit the file `Lab4starter-HF.sbatch` in the text editor of your choice. Read it from top to bottom. You fill find a note about how to modify it to make it work for your job. Make the necessarily change (it is only a single line) and save the file. Now, at the linux command prompt, run `sbatch Lab4starter-HF.sbatch`. This will submit the job to the queue. When an appropriate node or group of nodes becomes available, they will be assigned to your job, and then the `Lab4starter-HF.sbatch` script will run on that node. This may be right away, or it could be after a delay. You can check the status of the job by running `squeue -u <username>`, replacing <username> with your actually Seawulf username. For the remainder of this lab, you may run your jobs this way, or you may want to run them more interactively as was explained in Lab0, if you prefer.

Now, for the scientific portion of the exercise, we will run a series of potential energy scans to look at the dissociation of three systems: H_2, N_2, and Ar_2. Our goal is to investigate the behavior of several correlated electronic structure methods. In the case of Ar_2, we will also investigate the effect of basis set.

1. Above, you already ran `Lab4starter-HF.py`, which computed the PES of H_2 as a function of the bond distance at the HF/cc-pvdz level. Read through the script to see how it worked. You will find the results in Lab4starter-HF.csf. Take a look.

2. Now use modified copies of this starter script to perform similar scans at the Møller-Plessett second order perturbation theory (MP2) level instead of Hartree-Fock. Remember that the Psi4 keyword for MP2 is just `mp2`. If you want to submit your job to the queue as described above, you will also need to create an appropriate sbatch script by copying and modifying `Lab4starter-HF.sbatch`.

3. Now, you will run a similar calculation at the CASSCF level. We provide a script, `Lab4starter-CAS.py`, that will perform this calculation. You will have to create your own sbatch script if you wish to submit through the queue.

4. Now, modify the script you used above to perform a similar scan for N_2 at the HF/cc-pvdz level. Note that N_2 has a longer equilibrium bond length than H_2, so you will want to adjust the array of distances to provide useful information. Look up the experimentally known distance, and choose the distances for your scan such that you have many points (perhaps one every 0.05 Å) near the equilibrium bond distance, but fewer as you stretch the bond. Also note that the wave function can be difficult to converge at distances much shorter than the equilibrium distance.

5. Now, also perform a similar scan at the CASSCF level using the cc-pvdz basis, modifying `Lab4starter-CAS.py`. Note that you will have to define the active space. Include all valence electrons and all valence molecular orbitals (bonding or antibonding) in the active space by modifying the `active` and `restricted_docc` options. To do this, count up all core orbitals (those derived from the 1s shell) and set `restricted_docc` to this number. These orbitals will be doubly occupied in all determinants used in the CASSCF expansion. (Note that this was zero for H_2, because H_2 has no core orbitals.) Now count up all valence orbitals (derived from 2s and 2p shells), and set `active` to this number. These orbitals may be doubly occupied, singly occupied, or unoccupied in the determinants defining the CASSCF expansion. All higher orbitals will not be occupied in the CASSCF wave function. Scan to a maximum distance of 20 Å, because the CASSCF wave function becomes very difficult to converge at larger distances.

6. Now try Ar_2. Chemically, this is much different than system than N_2 or H_2, as it is not held together by a chemical bond. Again the equilibrium distance is much different, so look up the experimental value and use it to guide your choice of distances. Again, the calculations will not converge if you push the atoms too close together, but you do want to include distances both shorter and longer than the equilibrium distance. Scan the PES of Ar_2 at the following levels of theory: Hartree-Fock, MP2, and CCSD. Use the cc-pvdz basis. The Psi4 keyword for CCSD is just `ccsd`.

7. Repeat the Ar_2 scans for all three methods using a bigger basis set: cc-pvtz.

8. Density functional theory (DFT) is a very popular method for treating electron correlation that we have not yet discussed in lecture. Recalculate all three surfaces (H_2, N_2, and Ar_2) using DFT with the B3LYP functional (keyword `b3lyp` in place of `scf`), a very common density functional approximation.

3   Deliverables

- Plot your results using the plotting program of your choice. Be sure to use chemically meaningful units. Shift the energies so that the longest distance that was calculated for each method/basis has zero energy. Make several plots such that one can clearly see the effect of using different methods on each molecule. Label axes. Set the x and y ranges so that you can see the important features of the PES (e.g., the potential well at the equilibrium distance, if one is present, and any anomalies you might note in the shape of the surface).

- H_2 is a special molecule, because the dissociated state is exactly solvable: it is just two hydrogen atoms, which you have likely solved by hand in a prior quantum class. Which wave function method predicts the most accurate energy in the dissociated state? (Note that here we will refer to Hartree-Fock, MP2, CCSD, and CASSCF as "wave function methods." This term does not apply to B3LYP, which we will discuss seperately below.) What is wrong with the PESs computed using the other wave function methods? Why do these errors occur?

- We cannot exactly solve the Schödinger equation for dissociated N_2, but we can look up the experimental heat of formation of N atoms from N_2. Which wave function method provides the closest estimate? Why did the others fail? Why was a larger active space chosen for the CASSCF calculation of N_2 than for H_2?

- Compare and contrast the PESs for dissocation of Ar_2. What big difference do you see between Hartree-Fock and the correlated wave function methods (MP2, CCSD)? Why do you think this occurs? What effect does increasing the size of the basis set have on the dissociation energy of Ar_2? Why do you think the choice of basis set has such an effect? (Hint: The variational principle is not the correct answer for two reasons: a) the variational principle does not apply to energy differences, and b) MP2 and CCSD are not variational methods.)

- Now consider the B3LYP results for all three systems? How does it perform compared to the wave function methods? We have not yet discussed the theory underlying DFT, so we will not ask "why?" just yet, but take note of these results, as we will discuss DFT in detail starting in a few weeks.