Exercise Sheet 1

Due: 30.04.24 14h00

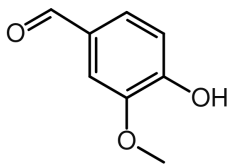
Fred Hamprecht
Gerrit Gerhartz, Christof Gehrig

General Regulations.

- Please hand in your solutions in groups of up to two people.
- Your solutions to theoretical exercises can be either handwritten notes (scanned), or typeset using LaTeX. In case you hand in handwritten notes, please make sure that they are legible and not too blurred or low resolution.
- For the practical exercises, always provide the (commented) code as well as the output, and don't forget to explain/interpret the latter. Please hand in both the notebook (`.ipynb`), as well as an exported pdf.
- Submit all your files in the Übungsgruppenverwaltung, only once for your group.

1 SMILES strings

(a) Write down the SMILES string for the following molecule. (2 pts)



(b) Explain why a molecule can have two valid SMILES strings. (3 pts)

2 PySCF calculations

In this exercise, you will do more PySCF calculations as you already did in the tutorials.

(a) Use the SMILES string from the first exercise and the RDKit library to calculate the positions of the atoms in the molecule and visualize them in 3d space. Run a restricted, closed-shell Kohn-Sham DFT calculation for this molecule.

(3 pts)

(b) Now, let's estimate how long these calculations take as the system gets bigger. Imagine a line of water molecules ($n = 10$ to 50) spaced 100 Å apart (You can do less if the computation takes too long on your machine). Run a restricted, closed-shell Kohn-Sham DFT calculation for each configuration with default parameters and record the runtime. We expect the runtime to increase proportionally to the number of electrons N raised to some power a, like so:

$$\text{Runtime} \propto \mathcal{O}(N^a).$$

Here, N represents the number of electrons in the molecule. Use the collected runtimes to estimate the value of a in this equation.

Hint: A function of the form $f(N) = aN^b$ can be helpful for fitting the data. (5 pts)

(c) Redo (b) using second-order Møller–Plesset perturbation theory (MP2) and using restricted configuration interaction with up to double excitations (CISD). You can decrease the number of atoms in a row, for the calculation to be completed in a reasonable time. Compare the exponents.

(5 pts)

(d) Use your run time approximation to estimate how long it would take to calculate the energy for one SARS-CoV-2-Virion at restricted, closed-shell Kohn-Sham accuracy on your machine (You can estimate the number of electrons in the virion very roughly).

(2 pts)