**NOTES**

**Machine Learning-Driven Initial Guess Prediction for SCF Procedures**

**Thank you everyone for coming to the presentation of my master’s thesis Machine Learning-Driven Initial Guess Prediction for SCF Procedures.**

**Motivation:**

2:

* 100 years QM (Heisenberg) published Sept. 1925
* Hartree developed method shortly thereafter
* DFT introduced a bit later

3:

Time is (still) gold. Age where you can simulate molecules on handhelds / PC. Simulation time is long for larger systems. Computational time in general expensive & backbone of many industries:

* Automated drug discovery
* Material discovery
* Reverse material engineering

**Outline**

4:

Background -> Intro to SCF method & Neural Networks / ML

Fock based guessing -> a stepstone for further experimentation

Graph Neural Networks -> custom model development

Application of GNNs -> can we beat established guessing strategies

**Background – SCF**

5:

Born-Oppenheimer approximation: e- on different time scales as nucleus -> solve the time-independent Schrödinger equation of the electron. Minimize energy of WF 🡪 but how do we do that. We can’t simulate the full n-elec interaction. So basic idea is to hold all but one electron fixed and use the Fock operator which expresses the mean field acting on one electron given by all other electrons.

6:

From this the Hartree-Fock eigenvalue equation follow, which we need to discretize in order to treat them in a computer simulation. 🡪 Gaussian-type Orbitals are commonly used to approximate spherical harmonics to model the AOs eventually forming MOs. Equation is self-dependent 🡪 Iterative process and initial guess needed.

7:

Density Functional Theory in its Kohn-Sham form uses the fact of the Hohenberg-Kohn theorems to state that the ground state energy of a many-electron system is a functional of the electron density. And the density which minimizes the energy gives the ground state.

B3LYP is Rung 4 functional in DFT – This means it uses methods from the first 3 rungs: GGA (General Gradient approximation) + mixes in some HF energy. Furthermore, it adds LYP corrections.

Generalized Wolfsberg-Helmholtz takes the diagonal part of Core Hamiltonian (explain: Kinetic part + potential part w.r.t. nucleus) and constructs off-diagonals by weighing using the overlap matrix (1-e integrals).

SAD / MINAO guesses: Calculate on minimal basis AO -> then build a superposition of these and project to given basis

Extra on LYP: Recasts Colle-Salvetti formula for electron pair correlation into DF expression – by expressing correlation energy density in terms of the local electron density and its gradients. Express molecular correlation energies

**Background – Machine Learning**

8:

Machine Learning is a subfield of Artificial Intelligence 🡪 paradigms: supervised, unsupervised, Reinforcement.

We have a supervised regression problem. Explain labels non labels regression vs. classification

Broad application from medicine over industry to science (LLMs etc.)

We treat (most) ML models as black-boxes (some research is done to let us understand them better).

9:

Perceptron – input weights + bias + output

10:

Use reward objective to train a ML model -> Minimize Loss-function

14: MLP

40k matrix elements to fit! 🡪 Model has to be gigantic if you do it one shot on whole matrix. 🡪 GWH reconstruction only 284 diagonal elements