A Heuristics-Based Approach to the Prediction of FODs in Fermi-Lowdin Orbital Self Interaction Correction (FLOSIC) Theory

Angel-Emilio Villegas Sanchez

A thesis presented for the degree of

Master's of Science

Supervised by:

Dr. Juan Ernesto Peralta

Dr. Koblar Alan Jackson

Central Michigan University

February 2024

Table of Contents

Introduction

Theoretical Background

Density-Functional Theory (DFT)

Density Functional Theory's principal tenet holds that any property of a system of interacting particles can be seen as a fucntional of the ground state density (Martin 2004). Precurse theory includes Thomas-Fermi theory which looked at the density of systems to calculate its energy. However, it was not until 1965 that Pierre Hohenberg and Walter Kohn formalized the theory.

Hohenberg-Kohn Theorems

Kohn-Sham Formulation

Kohn and Sham replace the Interacting system of N electrons, with a non-interacting system (Martin 2004).

$$\begin{aligned} E^{KS} = & \sum_{i\sigma} f_{i\sigma} \left\langle \psi_{i\sigma} \left| -\frac{\nabla^2}{2} \right| \psi_{i\sigma} \right\rangle + \int d^3r \rho(\vec{r}) v_{ext}(\vec{r}) \\ & + \frac{1}{2} \iint d^3r d^3r' \frac{\rho(\vec{r})\rho\left(\vec{r}'\right)}{|\vec{r} - \vec{r}'|} + E_{xc}\left[\rho_\uparrow, \rho_\downarrow\right]. \end{aligned} \tag{1}$$

Limitations of DFT

The Self-Interaction Error (SIC)

In contrast to Hartree-Fock theory, where the exchange integral cancels out exactly the Hartree energy, KS-DFT does not fully remove the self-interaction correction.

Perdew-Zunger SIC

The exchange-correlation should cancel the Hartree energy in a non self-interacting system

$$U[n_{\alpha\sigma}] + E_{xc}[n_{\alpha\sigma}, 0] = 0 \tag{2}$$

However, Equation 2 does not hold true in KS-DFT. On an orbital-by-orbital basis, we can remove the difference generated by the self-interaction.

$$E^{DFA} = E^{DFA}[\rho] + \sum_i U^{SIC}[\rho_i]$$

The FLOSIC Method

The Fermi-Lowdin Orbital (FLOs)

The canonical KS orbitals can be transformed into Fermi orbitals (FOs):

$$F_{i\sigma}(\vec{r}) = \frac{\sum_\alpha \psi^*_{\alpha\sigma}(\vec{a}_{i\sigma})\,\psi_{\alpha\sigma}(\vec{r})}{\sqrt{\sum_\alpha |\psi_{\alpha\sigma}(\vec{a}_{i\sigma})|^2}}. \tag{3}$$

The transformation can be rewritten more simply as

$$F_{i\alpha} = \frac{\psi^*_\alpha(\vec{a}_i)}{\sqrt{\rho(\vec{a}_i)}} \tag{4}$$

The position $\vec{a}_i$ is called the Fermi-orbital descriptor (FOD). It parametrizes each individual electron wavefunction in the system. These orbitals are not always orthogonal, so they are orthogonalized using the Lowdin method of orthogonalization.

The NRLMOL Code

- Describe the code and status of usability

Overview of Code

- Big-picture description of how the code works.

FOD Determination and Practical Considerations

- Describe methods used to input FODs into the FLOSIC program.
- Prime motivation behind this project: an automated and faster process to begin calculations, informed from "empirical" data, capable to extending to larger systems.

FOD Geometric Motifs

- Describe how FODs arrange themselves similar to LDQ Theory. This (emergent?) behavior makes it

As noted in paper X, optimized FODs get arranged in geometries similar to Thompson structures. This regular structure is most apparent in monoatomic systems, but the reagular FOD arrangement distorts in the presence of additional atoms.

SP3 Arrangements

SP3D Arrangement

The shape formed by 9 electrons in the Thompson problem is the polyhedron known as an *augmented triangular prism*, or a *tricapped trigonal prism*. The same geometry was proposed by King (1996) as the coordination polyhedra with idealized maximum symmetry. FOD arrangement is different from this idealized geometry, but qualitatively is the same on visual inspection.

Dominance Schema

- Sometimes FODs are closer to atoms that are of lesser atomic number Z than the neighbouring heavier atoms. This could be attributed to shielding caused by the filled shells of periods 3 and above. The behavior is described in Figure 1. This will play a special part in the prediction of bonding FODs.

Figure 1: dominance

The FOD-Lego Program

- Note: Different name needed probably

General Overview

- Give a **Big-Picture** overview as seen in image Figure 2

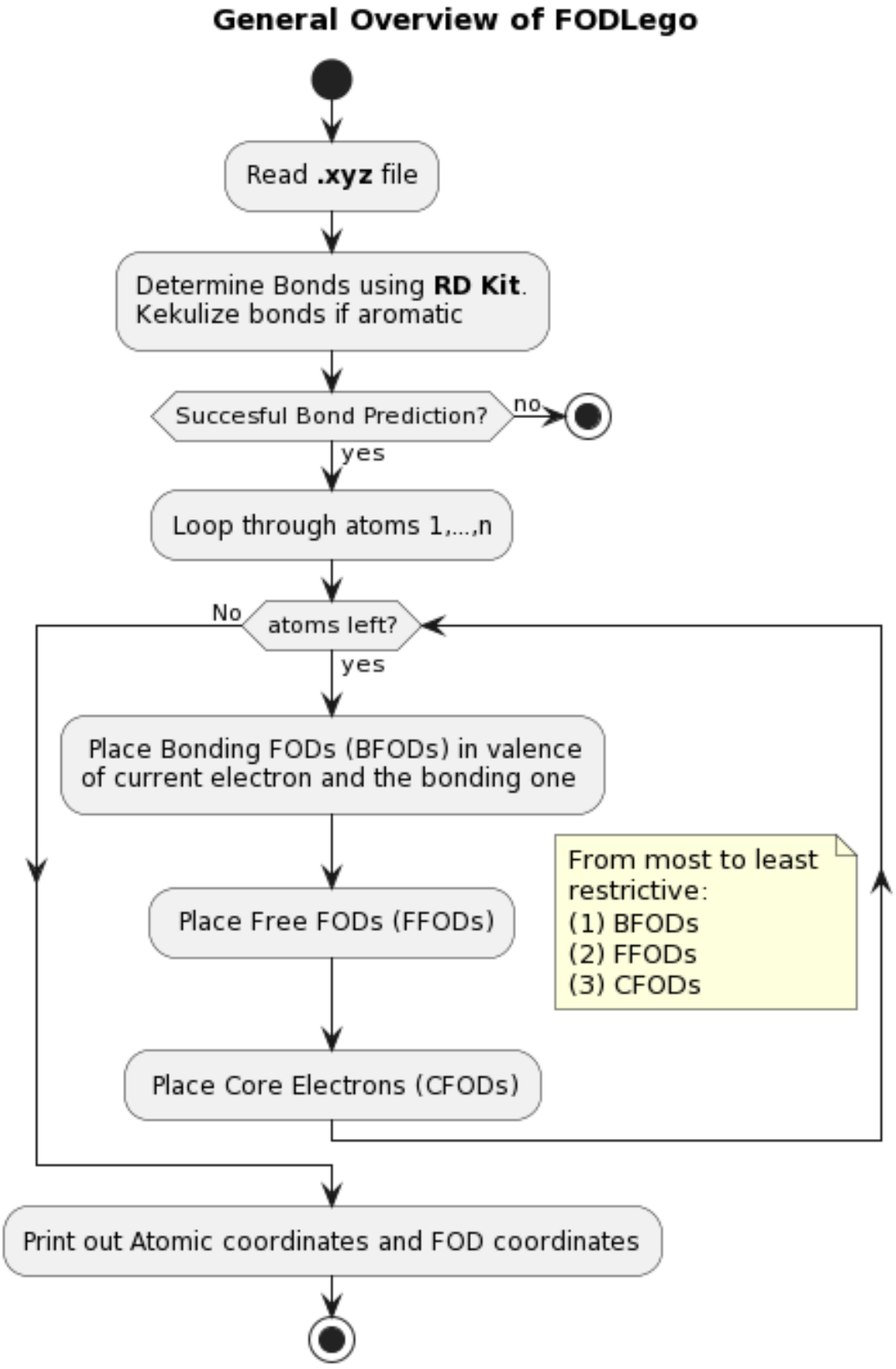

Figure 2: overview

A Heuristic-Based Approach to FOD Placement

- How can Heuristics help in computational problems?

Monoatomic Calculations

- Talk about *Dynamic Programming* and how monoatomic geometries could be pieced together in modular fashion (Lego-like).

Reparametrizing the FOD geometries

- sp3 shells stop being perfect tetrahedra and become irregular in the presence of other atoms.
- Describe how to parametrize FODs in irregular tetrahedra for the sp3 shells.

Implementation Details

- Describe code structure and use of object-oriented design
- Describe fundamental as seen in Figure 3

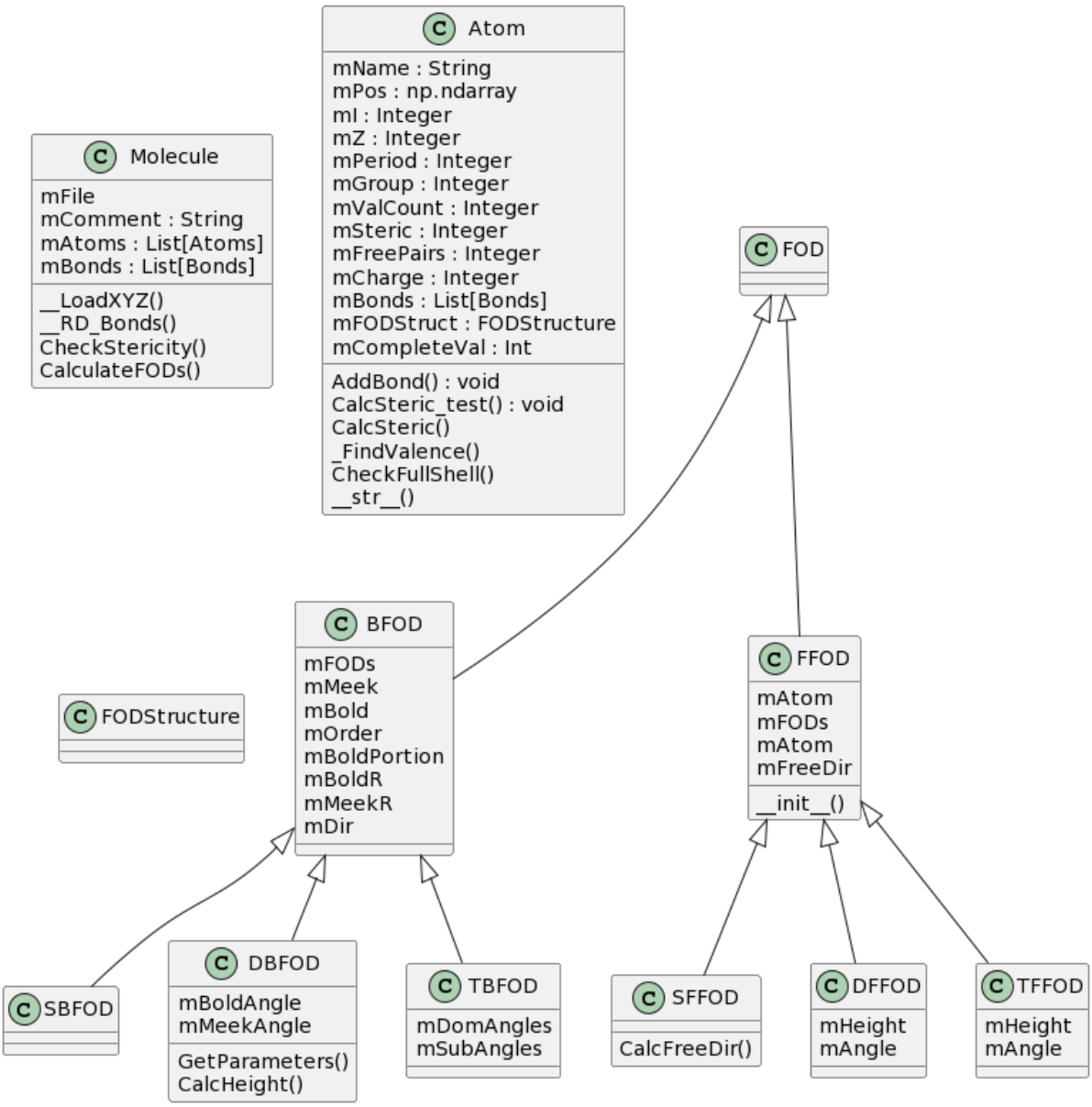

Figure 3: classes

Determination of FODs

- Describe the different determination heuristics for different types of FODs

Bonding FODs (BFOD)

Core FODs (CFOD)

Free FODs (FFOD)

- Sample Diagram and calculation for Double-FFOD (DFFOD) Heuristic in Figure 4. The edge distance is applied to as a restriction.

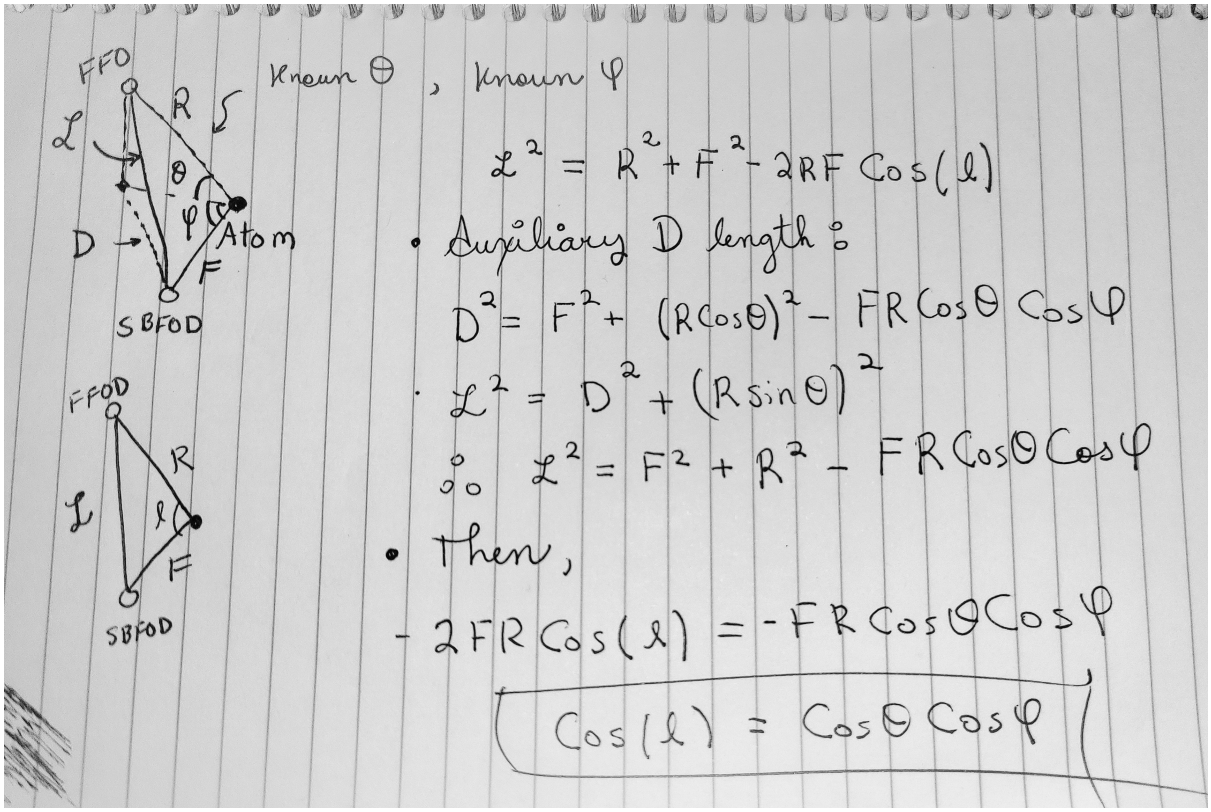

Figure 4: draft DFFOD

A Data-driven Approach to analyzing FODs

In several, and some recent papers the FODs have usually been characterized via their distances from nucleus. This is merely qualitative.

Reverse Determination of Parameters

This is the following implementation used for the reverse determination of parameters

Algorithm 1: Reverse Parameter Determination Basic Sequence

Data: $\{T_n \mid \text{Target FODs}\}$, $\{P_n \mid \text{Predicted FODs}\}$

for P_n *in Predicted FODs* **do**

  Save closest T_i to FOD as T_{close};

  Initialize a new FOD with T_{close} N_n;

Reverse Determine parameters for all N_n;

Free FODs (FFODs). The free-direction, described in section XYZ, cannot be determined exactly. In order to best compare Free FODs in this parameter-based scheme, I propose using only an average direction formed from the average of all FFOD vectors and use an evaluation function. A cosine distance is too flat near its maximum, so a custom function can offer a more strict metric of our prediction

$$f(\theta) = 1 - \frac{6}{\pi}\theta \tag{5}$$

Equation 5 yields unity when there at 0 degrees between the vectors, and 0 when there are 30 degrees between the vectors.

A database of compiled FOD Data

- No current database of FODs. Seeking to create a database
- **TODO!**
  - Design structure of database
  - Using FODLego, then FLOSIC, create optimized structures for a selection of molecules (scope?)

Validity of FOD Prediction

Explore the predictions done by FODLego

Comparison of Heuristic FODs to FLOSIC-Optimized FODs

- **TODO:**
  - Create scatter plots for different parameters described in the scheme of previous section.

Comparison to FOD-MC

Biasing Predictions after Sampling (?)

- – Assuming data follows certain trends, we can further modify the predictions used in FODLego. For example, the tightening of closed shells in heavier atoms in the presence of the next shell closed.

Conclusions/Findings

Discussion

Limitations

Assumptions

- Explore assumptions such as First neighbors are mainly considered in this scheme.

Automated Bonding Prediction

For Future Implementation

Extending modularity to metal FODs?

- Talk about possible parametrization of the *Triaugmented Triangular Prism* composed of 9 FODs.

Spin-Unpolarized Prediction

King, R.B., 1996. The shapes of coordination polyhedra. *Journal of Chemical Education*, 73(10), p.993. Available at: https://doi.org/10.1021/ed073p993.

Martin, R.M., 2004. *Electronic structure: Basic theory and practical methods*, Cambridge University Press.