



DEPARTMENT OF PHYSICS

TFY4510 - PHYSICS, SPECIALIZATION PROJECT

---

# Simulations of Nickel Clusters using Density Functional Theory and a Machine Learning Guided Algorithm

---

*Author:*  
Christoffer Askvik Faugstad

*Supervisor:*  
Jaakko Akola

Date  
20th December 2022

---

## Abstract

The global energy transition away from fossil fuels is one of our most pressing issues. Fischer–Tropsch synthesis has the potential to be an important part of this transition. Fischer-Tropsch synthesis is a set of chemical reactions that convert hydrogen and carbon-monoxide gas into various hydrocarbons. For the reaction to occur, the presence of a catalyst is essential. One of the possible catalyst materials is nickel. In this work, the ground state geometries of nickel clusters is investigated as an initial step towards simulating Fischer–Tropsch synthesis. A global energy minimization for nickel clusters of 8, 12 and 30 atoms, is done using density functional theory. In the density functional theory calculations, the generalized gradient approximation of Pedrew-Bruke-Ernzerhof is used to approximate the exchange-correlation energy. A machine learning guided method is used to reduce the computation time of the global energy minimization. The method consists of an initial screening of structures using effective medium theory and a random geometry search. Promising geometries from the screening is locally optimized and used as training data for a Gaussian process regression model. This model acts as a surrogate potential. The potential guides a new search, where a combination of random candidates and rattle-generated candidates are used. The lowest energy geometry of the  $\text{Ni}_8$  and  $\text{Ni}_{12}$  cluster found by the method is in agreement with the literature. For the  $\text{Ni}_{30}$  cluster, the optimal structure has higher energy than the literature reference. For all relevant low energy geometries the cohesive energy, ionization potential, electron affinity, bond length and magnetization is found. The density of states and the charge distribution of the highest occupied and lowest unoccupied orbitals are also computed.

---

# Table of Contents

|  |           |
|--|-----------|
| <b>Abstract</b>  | <b>i</b>  |
| <b>List of Figures</b>                                     | <b>iv</b> |
| <b>List of Tables</b>                                      | <b>iv</b> |
| <b>1 Introduction</b>                                      | <b>1</b>  |
| <b>2 Theory</b>  | <b>2</b>  |
| 2.1 Chemistry . . . . .                                    | 2         |
| 2.1.1 Orbitals . . . . .                                   | 2         |
| 2.1.2 Cohesive Energy . . . . .                            | 3         |
| 2.1.3 Ionization Potential and Electron Affinity . . . . . | 3         |
| 2.1.4 Magnetization . . . . .                              | 3         |
| 2.2 Crystalline Solid State . . . . .                      | 3         |
| 2.2.1 Lattice . . . . .                                    | 4         |
| 2.2.2 Reciprocal Lattice and Brillouin Zone . . . . .      | 4         |
| 2.2.3 Density of States . . . . .                          | 4         |
| 2.3 Quantum Mechanics . . . . .                            | 4         |
| 2.3.1 Schrödinger Equation . . . . .                       | 4         |
| 2.3.2 Born-Oppenheimer Approximation . . . . .             | 5         |
| 2.3.3 Properties of the Wave Function . . . . .            | 5         |
| 2.3.4 Expectation Values . . . . .                         | 6         |
| 2.3.5 Variational Principle for the Ground State . . . . . | 6         |
| 2.3.6 Hartree-Fock Approximation . . . . .                 | 6         |
| 2.4 Density Functional Theory . . . . .                    | 7         |
| 2.4.1 Electron Density . . . . .                           | 8         |
| 2.4.2 Hohenberg-Kohn Theorems . . . . .                    | 8         |
| 2.4.3 Kohn Sham . . . . .                                  | 9         |
| 2.4.4 Unrestricted Kohn Sham . . . . .                     | 10        |
| 2.5 Exchange Correlation Functional . . . . .              | 10        |
| 2.5.1 Local Density Approximation . . . . .                | 10        |
| 2.5.2 Local Spin Density Approximation . . . . .           | 10        |
| 2.5.3 General Gradient Approximation . . . . .             | 11        |
| 2.6 Basis Sets and Pseudopotentials . . . . .              | 12        |

---

|                   |  |           |
|-------------------|--|-----------|
| 2.6.1             | Slater . . . . .   | 12        |
| 2.6.2             | Gaussian . . . . .   | 13        |
| 2.6.3             | Plane Wave . . . . .   | 13        |
| 2.6.4             | Augmented-plane-wave . . . . .   | 13        |
| 2.6.5             | Pseudopotential . . . . .  | 13        |
| 2.6.6             | Projector Augmented-Wave . . . . .   | 14        |
| 2.7               | Force Fields . . . . .   | 14        |
| 2.7.1             | Effective Medium Theory . . . . .  | 14        |
| <b>3</b>          | <b>Numerical Methods</b>   | <b>15</b> |
| 3.1               | Local Optimizers . . . . .   | 15        |
| 3.1.1             | Steepest Descent . . . . .   | 15        |
| 3.1.2             | Conjugate Gradient Algorithm . . . . .   | 15        |
| 3.1.3             | Newton Method . . . . .  | 16        |
| 3.1.4             | Broyden–Fletcher–Goldfarb–Shanno algorithm . . . . .                           | 16        |
| 3.2               | Density Functional Theory in the Vienna Ab-initio Simulation Package . . . . . | 16        |
| 3.2.1             | Periodic Boundary Conditions . . . . .   | 16        |
| 3.2.2             | Smearing . . . . .   | 17        |
| 3.2.3             | Parameter Tuning . . . . .   | 17        |
| 3.3               | Global Energy Minimization Method . . . . .                                    | 17        |
| 3.3.1             | Random Geometry Search . . . . .   | 17        |
| 3.3.2             | Oganov Fingerprint . . . . .   | 18        |
| 3.3.3             | Global Optimization with First-principles Energy Expressions . . . . .         | 18        |
| <b>4</b>          | <b>Results and Discussion</b>  | <b>20</b> |
| 4.1               | Parameter Search . . . . .   | 20        |
| 4.2               | $\text{Ni}_8$ cluster . . . . .  | 21        |
| 4.3               | $\text{Ni}_{12}$ cluster . . . . .   | 22        |
| 4.4               | $\text{Ni}_{30}$ cluster . . . . .   | 24        |
| 4.5               | Global Energy Minimization Method . . . . .                                    | 25        |
| <b>5</b>          | <b>Conclusion</b>  | <b>27</b> |
| <b>References</b> |  | <b>30</b> |
| <b>Appendix</b>   |  | <b>31</b> |
| <b>A</b>          | <b>Code - Random Geometry Search</b>   | <b>31</b> |

---

---

|          |   |           |
|----------|---|-----------|
| <b>B</b> | <b>Code - Global Optimization with First-principles Energy Expressions</b>                | <b>33</b> |
| <b>C</b> | <b>Density of States</b>  | <b>35</b> |
| <b>D</b> | <b>Orbitals</b>   | <b>38</b> |
| <b>E</b> | <b>Global Optimization with First-principles Energy Expressions Performance Show-case</b> | <b>41</b> |

## List of Figures

|    |   |    |
|----|---|----|
| 1  | Global Optimization with First-principles Energy Expressions Flowchart . . . . .  | 19 |
| 2  | Ni <sub>8</sub> Parameter Search . . . . .  | 20 |
| 3  | Ni <sub>8</sub> Cluster Bidispheoid Density of States . . . . .   | 22 |
| 4  | Ni <sub>12</sub> Cluster Density of States Comparison . . . . .   | 23 |
| 5  | Ni <sub>30</sub> Cluster Density of States Comparison . . . . .   | 25 |
| 6  | Ni <sub>30</sub> Cluster Global Optimization with First-principles Energy Expressions Candidate Energies . . . . .                          | 26 |
| 7  | Ni <sub>30</sub> Cluster Surrogate potential and Effective Medium Theory Comparison to Density Functional Theory . . . . .                  | 27 |
| 8  | Density of States for the Ni <sub>8</sub> Geometries . . . . .  | 35 |
| 9  | Density of States for the Ni <sub>12</sub> Geometries . . . . .   | 36 |
| 10 | Density of States for the Ni <sub>30</sub> Geometries . . . . .   | 37 |
| 11 | Charge Distribution of Highest Occupied and Lowest Unoccupied Molecular Orbitals of the Ni <sub>8</sub> Geometries . . . . .                | 38 |
| 12 | Charge Distribution of Highest Occupied and Lowest Unoccupied Molecular Orbitals of the Ni <sub>12</sub> Geometries . . . . .               | 39 |
| 13 | Charge Distribution of Highest Occupied and Lowest Unoccupied Molecular Orbitals of the Ni <sub>30</sub> Geometries . . . . .               | 40 |
| 14 | Ni <sub>8</sub> Cluster Global Optimization with First-principles Energy Expressions Candidates and Energy Prediction Comparison . . . . .  | 41 |
| 15 | Ni <sub>12</sub> Cluster Global Optimization with First-principles Energy Expressions Candidates and Energy Prediction Comparison . . . . . | 42 |
| 16 | Ni <sub>30</sub> Cluster Global Optimization with First-principles Energy Expressions Candidates and Energy Prediction Comparison . . . . . | 43 |

## List of Tables

|   |   |    |
|---|---|----|
| 1 | Calculation Parameters . . . . .                    | 20 |
| 2 | Properties of Ni <sub>8</sub> Geometries . . . . .  | 21 |
| 3 | Properties of Ni <sub>12</sub> Geometries . . . . . | 22 |

---

|   |   |    |
|---|---|----|
| 4 | Properties of Ni <sub>30</sub> Geometries . . . . . | 24 |
|---|---|----|

---

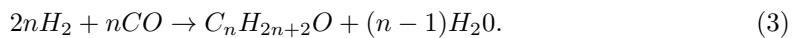
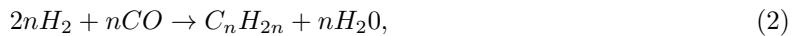
# 1 Introduction

As the world population continues to grow, the global energy demand is rapidly increasing. This causes more global warming and increases the demand on renewable energy sources. Around 84% of the world's energy consumption comes from fossil fuel [1]. While the transition to renewable electricity is one of the long term energy solutions, several sectors are hard to electrify. These sectors are among other aviation, transportation and mobility, chemical and material production. Fischer-Tropsch synthesis (FTS) could be a bridging solution to speed up the transition to renewable energy through the production of liquid fuels from renewable resources [2].

FTS is a set of chemical reaction that converts carbon-monoxide (CO) and hydrogen (H<sub>2</sub>) gas into liquid hydrocarbons. The process consists of polymerization (chain-growth) and hydrogenation (binding of H) of CO. The result is a mixture of different hydrocarbons, oxygenates (hydrocarbons including oxygen) and water. FTS is usually part of a three step process. The first is the production of the mixture of CO and H<sub>2</sub>. This mixture is called synthesis gas (syngas). The syngas is commonly made from coal, natural gas or bio-mass through a process called gasification. There also exist possibilities of using atmospheric CO<sub>2</sub> in a modified version of FTS [3]. Then the syngas is synthesised into the product mixture, this mixture is called syncrude. The syncrude needs to be filtered before further use. The products from FTS have several applications including, fertilisers, chemicals, plastics, natural gas substitute, diesel, petrol and aviation fuel.

FTS was first developed by Fisher and Tropsch in 1925. From FTSs first development and until the after the second world war, the main developments took place in Germany. The production and research were mainly due to strategic arguments, which exceeded the economic aspect at this time. In a ten year period after the war, there was a larger global interest in FTS. This was partly due to large coal reserves, pessimistic forecasts for oil reserves and an increasing demand for liquid fuels. However, this interest rapidly declined when several big oil fields were discovered. The new oil discovery meant that coal-based FTS would not be economically viable by comparison. The interest in FTS began to grow around the seventies and eighties. This was partly due to new pessimistic forecast for the global oil market. Several countries started to investigate both coal and natural gas based FTS [4]. This attention has continued to increase in recent years as the transition away from fossil fuels have accelerated due to global warming. There is currently much ongoing research into FTS.

The main reactions of FTS can be formulated as [5]



The reaction corresponds to the production of alkanes, alkenes and oxygenates respectively. The reactions happen because of a catalyst. The possible catalyst materials are Iron (Fe), Cobalt (Co), Ruthenium (Ru) and Nickel (Ni). These materials have certain common properties that are regarded as important for FTS. This includes that the materials are active for hydrogenation reactions and that they are capable of forming metal carbonyl at the operating conditions [4]. Metal carbonyls are structures where metal and CO molecules are bound together. The presence of the catalyst is essential to make the reactions happen. The catalysts lower the activation energy needed to go from the reactants to the products, which implies an increased reaction rate also called increased activity. The different catalyst has different characteristics and makes certain reaction steps have an increased rate. This has implications on the distribution of products, called the selectivity. Ru is the most active catalyst, but it is also the most costly and scares, thus not deemed viable for commercial use. Fe and Co are the ones considered commercially suitable at this point. Fe has a relatively low activity. It can operate under a wide range of conditions and is cheap. Co is more sensitive to the operating conditions and more expensive than Fe, but has a higher activity. Both Co and Fe have a selectivity towards longer chains. Ni catalyst has a high selectivity towards methane at higher temperatures [6]. As longer hydrocarbon-chains are preferred, Ni is not deemed fit for commercially in use.

---

Modeling FTS at the atomic level is a complicated process several step . The starting configuration or ground state of the catalyst need to be known. This should also preferably include the surface on which the catalyst is deposited on. Then one can introduce the singers into the system and map out binding sites for the molecules. This needs to be done both for reactants and products. Then the energy landscape between these states needs to be found to be able to determine the reaction mechanisms, pathways and rates.

The purpose of this work is to investigate the Ni catalyst. In this specialization project, the low energy geometries of three different sized Ni cluster will be explored. This is done as an introduction to the software and techniques used to model the Ni clusters. It is also a preliminary step in the simulations of the whole FTS reaction. The lowest energy geometries of the catalysts are the starting point for modelling the FTS reaction. These geometries are required to determine the state of the catalyst before the reaction starts. In the master thesis, this work will be continued to include the entire reaction. This will allow for looking into the effects of coverage, as the number of adsorbed molecules would affect the activity and selectivity. The structure is also possibly altered by an increased coverage.

In this work Ni clusters of size 8, 12 and 30 will be investigated. The clusters will be simulated using density functional theory (DFT). DFT is one of the most successful and widely used quantum mechanical tools [7]. DFT is based upon the Hohenberg-Kohn theorems. In essence, the first theorem states that the energy is a functional, i.e. a function of a function, of the electron density. The second one guarantees that the electron density minimizing this energy is the ground state. Since Ni is magnetic, the calculations are spin-polarized. The method is implemented in the Vienna Ab-initio Simulation Package (VASP). The use of DFT requires the choice of an exchange-correlation energy functional, which incorporate the energy that are not due to electrostatics and non-interacting kinetic energy. In this work the Perdew–Burke–Ernzerhof functional (PBE) is used [8].

DFT calculations are computationally expensive. Therefor one would like to limit the number of them. In this work, the effective medium theory (EMT) and a machine-learning algorithm is employed to limit the need for doing DFT calculations. EMT is a force field model which is computationally significantly cheaper than DFT. Therefore it is employed to do initial search when the accuracy of DFT is not needed. The machine learning algorithm used is called Global Optimization with First-principles Energy Expressions (GOFEE) and is invented by Bisbo and Hammer [9]. It is based upon a surrogate potent that replaces most DFT calculations by learning the potential energy landscape from previous data. The algorithm also uses clustering to guide the search.

## 2 Theory

This section will introduce the basic theory used for simulating the Ni clusters. This includes general concepts in chemistry, crystalline solid state physics and quantum mechanics. Furthermore, the fundamentals of DFT are outlined, together with common choices of exchange correlation functionals, basis sets and psuedopotentials. Lastly force fields and the effective medium theory are introduced.

### 2.1 Chemistry

The purpose of this section is to introduce the main aspects within chemistry relevant for this work. The presented theory is based on *Chemical Principles* by Zumdahl [10].

#### 2.1.1 Orbitals

An atomic orbital describes the spatial distribution of an electron bond to the atom. A molecular orbital is the spatial distribution of an electron in a molecule. These molecular orbitals are typically

---

modeled as a linear combination of the atomic orbitals of its constituents atoms. The orbitals form an ordered pair based on their energy. The highest occupied molecular orbit (HOMO) is as the name suggest, the molecular orbital with the highest energy that is occupied by an electron. Similarly the lowest unoccupied molecular orbit (LUMO) is the next orbital above the HOMO and therefore is the lowest unoccupied orbit. The difference between the HOMO and LUMO energy is called the HOMO-LUMO gap or the band gap for periodic systems. The Fermi energy is the energy level in the middle of the band gap.

### 2.1.2 Cohesive Energy

The cohesive energy is the energy difference between the bound structure and its free constituents per atom. For a structure of  $N$  identical atoms the cohesive energy is

$$E_{\text{Cohesiv}} = \frac{E_{\text{structure}} - NE_{\text{Free atom}}}{N}, \quad (4)$$

where  $E_{\text{structure}}$  is the energy of the cluster and  $E_{\text{Free atom}}$  is the energy of one free atom. If the cohesive energy is negative the bound structure is energetically favored over the free atoms.

### 2.1.3 Ionization Potential and Electron Affinity

The ionization potential (IP) it the energy required to remove an electron from an atom or molecule. The IP is therefore the energy required for the following reaction to happen for a molecule  $X$



Analogously the electron affinity(EA) is the energy required to add en electron to an atom or molecule. The EA is the energy required for the following reaction happen



The above reaction usually releases energy, meaning that the EA is negative, whereas the IP is usually positive since it requires energy to remove the electron and make an ion. An IP or EA calculation is either done vertical or adiabatic. Vertical calculations are when the original structure is unaltered for the addition or removal of an atom. When adiabatic calculations are done the structure is re-optimized with the new number of electrons. In this work the vertical IP and EA are reported.

### 2.1.4 Magnetization

Due to the spin of the electrons, the structure can have a net magnetization if there is a difference in the number of up and down orbitals. In particular, this is the case for transition metals that have partially filled d-orbitals. This is because each spatial-orbital is occupied by one electron spin, before placing two different spins in the same spatial-orbital. This is the case for Ni. The electron configuration is structured as an Argon equivalent core and 2 electrons occupying the 4s orbitals and 8 electrons occupying the 3d orbitals. Since there are five 3d orbitals, Ni has two partially filled d-orbitals. The other catalyst materials, Fe, Co and Ru, are also magnetic. Therefore it is important to take the magnetization into account when analysing the catalyst materials. The size of the magnetic moment of the electron spin is the Bohr-Magneton  $\mu_B$ .

## 2.2 Crystalline Solid State

This section will give a brief introduction to relevant crystalline solid state aspects and is based on *Introduction to Solid State Physics* by Kittel [11].

---

### 2.2.1 Lattice

A crystal is a periodic arrangement of identical groups of atoms. The group of atoms is called the basis and the points which the basis is attached to is called the lattice. A lattice is defined by the three translation vectors  $\mathbf{a}_{1,2,3}$  which have the property that for any position  $\mathbf{r}$  it looks the same as any point translated by a integer value of the translation vectors, namely

$$\mathbf{r}' = \mathbf{r} + n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2 + n_3 \mathbf{a}_3 \quad (7)$$

where  $n_{1,2,3}$  are arbitrary integers. The lattice vectors are said to be primitive translation vector if it is the smallest unit cell,  $\mathbf{a}_1 \cdot \mathbf{a}_2 \times \mathbf{a}_3$ , that satisfy equation (7) for that crystal. The primitive cell is the cell defined by the primitive translation vectors. An alternative to the primitive cell defined by the translation vectors is the Wigner-Seitz primitive cell. This is made by connection the lattice points to all nearby lattice points by lines and then construction planes normal to these lines. The smallest volume enclosed by these planes is the Wigner-Seitz primitive cell. The simplest 3-dimensional lattice is the cubic lattice. The three translation vectors are perpendicular to each other and of the same length, given by the lattice constant.

### 2.2.2 Reciprocal Lattice and Brillouin Zone

The reciprocal lattice is the Fourier transform of the real lattice. The vectors of the reciprocal lattice  $\mathbf{b}_i$  are given by

$$\mathbf{b}_1 = 2\pi \frac{\mathbf{a}_2 \times \mathbf{a}_3}{V} \quad (8)$$

where the indices are cycled trough to give all vectors and  $V$  is the volume of the (real) unit cell. The Brillouin Zone is defined as the Wigner-Seitz primitive cell, but in reciprocal space.

### 2.2.3 Density of States

The density of states (DOS) is in essence the number of states, or in our case molecular orbitals, available at a given energy that electrons can occupy. From the DOS one can identify if a material whether a system is matalic, a semi-conductor or a insulator based on the band gap. The band gap is the difference in energy between the HOMO and LUMO. For metals there is no gap, whereas for semi-conductors the gap is small and insulators it is large, typically 3 eV or higher.

## 2.3 Quantum Mechanics

In this section elementary quantum mechanics will be introduced together with the Born-Oppenheimer and Hartree-Fock approximations. The introduction is based on the textbook of Parr R. and Weitao Y. [12]. Atomic units are used, unless something else is stated.

### 2.3.1 Schrödinger Equation

To properly describe atoms and molecules at the most fundamental level, one needs to use quantum mechanics. A many particle system of atoms is described by the time-dependent Schrödinger equation

$$\hat{H}\Psi = i\frac{\partial}{\partial t}\Psi, \quad (9)$$

where  $\hat{H}$  is the Hamiltonian and  $\Psi$  is the wave function. If the Hamiltonian is time independent, the system can be described by the time-independent Schrödinger equation

$$\hat{H}\Psi = E\Psi, \quad (10)$$

---

where  $E$  is the energy of the system. The Hamiltonian can be split into several terms

$$\hat{H} = \hat{H}_{nn} + \hat{H}_{ee} + \hat{H}_{en}. \quad (11)$$

The terms are the nucleus nucleus interaction  $\hat{H}_{nn}$ , the electron electron interaction  $\hat{H}_{ee}$  and the electron neutron interaction  $\hat{H}_{en}$ . The three parts are given by

$$\hat{H}_{nn} = \hat{V}_{nn} + \hat{T}_n = \sum_{\alpha<\beta} \frac{Z_\alpha Z_\beta}{|\mathbf{R}_\alpha - \mathbf{R}_\beta|} - \sum_\alpha \frac{1}{2M_\alpha} \nabla_\alpha^2, \quad (12)$$

$$\hat{H}_{ne} = \hat{V}_{ne} = - \sum_{\alpha,i} \frac{Z_\alpha}{|\mathbf{R}_\alpha - \mathbf{r}_i|}, \quad (13)$$

$$\hat{H}_{ee} = \hat{V}_{ee} + \hat{T}_e = \sum_{i<j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} - \sum_i \frac{1}{2} \nabla_i^2. \quad (14)$$

Here  $\mathbf{R}_\alpha$  is the positions of nucleus  $\alpha$ ,  $\mathbf{r}_i$  is the position of electron  $i$ ,  $\nabla_\alpha^2$  and  $\nabla_i^2$  is the Laplace-operator with respect to nucleus  $\alpha$  or electron  $i$  respectively and  $Z_\alpha$  and  $M_\alpha$  are the charge and mass of the nucleus.

### 2.3.2 Born-Oppenheimer Approximation

The mass of the nuclei are about a thousand times more massive the electrons. This means that the electrons moves orders of magnitude faster than the nuclei. Therefore when modeling the atomic system one can approximate the nuclei as frozen classical particles and only threat the electrons quantum mechanically. This leads to the following equation for the electronic wave function  $\Psi_e$

$$\hat{H}'\Psi_e = (\hat{H}_{ee} + \hat{H}_{ne})\Psi_e = E_e\Psi_e \quad (15)$$

The total energy of the atomic system can then be found by adding the nucleus nucleus energy the electron energy  $E_e$ . The nucleus electron interaction can be expressed as

$$\hat{H}_{ne} = \sum_i v(\mathbf{r}_i) = \sum_i \left( - \sum_\alpha \frac{Z_\alpha}{|\mathbf{R}_\alpha - \mathbf{r}_i|} \right). \quad (16)$$

The term  $v(\mathbf{r}_i)$  is the external potential acting on electron  $i$ . From this point on only the electron formulation will be considered and  $\Psi$  will refer to  $\Psi_e$ ,  $\hat{H}$  to  $\hat{H}'$  and  $E$  to  $E_e$  unless specified otherwise. Solving the Schrödinger equation (15) for a many electron system becomes rapidly unfeasible with an increased number of electrons and approximate methods are needed.

### 2.3.3 Properties of the Wave Function

The wave function of  $N$  electrons,  $\Psi = \Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)$ , is a function of the position  $\mathbf{r}_i$  and spin  $s_i$  compactly noted as  $\mathbf{x}_i$ . A given set of solutions of equation (15) is described by eigen-wave functions  $\Psi_k$  with corresponding energy eigenvalues  $E_k$ . The ground state of the system is the eigen-wave function with the lowest energy, denoted  $\Psi_0$  and  $E_0$ . The set of eigen-wave functions are complete and can always to be made orthonormal

$$\int \Psi_k^* \Psi_l d\mathbf{x}^N = \langle \Psi_k | \Psi_l \rangle = \delta_{kl}, \quad (17)$$

by re-normalizing. Here the Dirac notation for the inner product has been used and  $\delta_{kl}$  is the Kronecker-delta. The wave function for fermions have to be anti-symmetric with the exchange of two coordinates, i.e.

$$\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) = -\Psi(\mathbf{x}_2, \mathbf{x}_1, \dots, \mathbf{x}_N). \quad (18)$$

This the origin of the Pauli exclusion principle. This because if two electrons would be in the same state the wave function would be symmetric upon exchange of them.

### 2.3.4 Expectation Values

The expectation value of an observable  $\hat{O}$  is given by

$$\langle \hat{O} \rangle_{\Psi} = \frac{\int \Psi^* \hat{O} \Psi d\mathbf{x}^N}{\int \Psi_k^* \Psi_j d\mathbf{x}^N} = \frac{\langle \Psi | \hat{O} | \Psi \rangle}{\langle \Psi | \Psi \rangle}. \quad (19)$$

Then the kinetic and potential energies are given by

$$T[\Psi] = \langle \hat{T} \rangle_{\Psi} = \frac{\langle \Psi | \hat{T} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \quad (20)$$

and

$$V[\Psi] = \langle \hat{V} \rangle_{\Psi} = \frac{\langle \Psi | \hat{V} | \Psi \rangle}{\langle \Psi | \Psi \rangle}. \quad (21)$$

Here one say that  $T[\Psi]$  and  $V[\Psi]$  are functionals as they are uniquely determined by the wave function.

### 2.3.5 Variational Principle for the Ground State

For a system in state  $\Psi$  the energy is given by

$$E[\Psi] = \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle}. \quad (22)$$

Given a complete set of wave functions  $|\Psi_k\rangle$ , one can construct a general wavefunction  $|\Psi\rangle$ , as a linear combinations of these states,  $|\Psi\rangle = \sum_k a_k |\Psi_k\rangle$ . The energy is then given by

$$E[\Psi] = \frac{\sum_{k,l} \langle \Psi_k | a_k^* \hat{H} a_l | \Psi_l \rangle}{\sum_{k,l} \langle \Psi_k | a_k^* a_l | \Psi_l \rangle} = \frac{\sum_{k,l} a_k^* a_l E_l \delta_{kl}}{\sum_{k,l} a_k^* a_l \delta_{kl}} = \frac{\sum_k |a_k|^2 E_k}{\sum_k |a_k|^2} \geq E_0. \quad (23)$$

This means that the true ground state can be found by doing a minimization over all  $\Psi$ . Namely the fround state  $\Psi_0$  is the wave function solving

$$E_0 = \min_{\Psi} E[\Psi]. \quad (24)$$

Every eigen-wave function  $\Psi$  is an extremum of  $E[\Psi]$ . This means that one can rephrase equation (15) as

$$\delta E[\Psi] = 0. \quad (25)$$

If one also wants to imply normalization one can add this constrain using Lagrange multiplier, namely one should solve

$$\delta \left[ \langle \Psi | \hat{H} | \Psi \rangle - E \langle \Psi | \Psi \rangle \right] = 0. \quad (26)$$

### 2.3.6 Hartree-Fock Approximation

Hartree-Fock approximates the wave function by the product of  $N$  orthonormal single-particle wave functions  $\psi_i(\mathbf{x})$ , which are composed of a spatial part  $\phi_i(\mathbf{r})$  and a spin function  $\sigma(s)$ . This product is called the Slater determinant

$$\Psi_{HF} = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(\mathbf{x}_1) & \psi_2(\mathbf{x}_1) & \dots & \psi_N(\mathbf{x}_1) \\ \psi_1(\mathbf{x}_2) & \psi_2(\mathbf{x}_2) & \dots & \psi_N(\mathbf{x}_2) \\ \vdots & \vdots & & \vdots \\ \psi_1(\mathbf{x}_N) & \psi_2(\mathbf{x}_N) & \dots & \psi_N(\mathbf{x}_N) \end{vmatrix}. \quad (27)$$

---

The Hartree-Fock method finds the  $N$  single-particle wave functions that solve equation (24). By construction the Slater determinant is normalized and the expectation energy is given by

$$E_{\text{HF}} = \sum_{i=1}^N H_i + \frac{1}{2} \sum_{i,j=1}^N (J_{ij} - K_{ij}) \quad (28)$$

where

$$H_i = \int \psi_i^*(\mathbf{x}) \left( -\frac{1}{2} \nabla^2 + v(\mathbf{x}) \right) \psi_i(\mathbf{x}) d\mathbf{x} \quad (29)$$

$$J_{ij} = \iint \psi_i(\mathbf{x}_1) \psi_i^*(\mathbf{x}_1) \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \psi_j^*(\mathbf{x}_2) \psi_j(\mathbf{x}_2) d\mathbf{x}_1 d\mathbf{x}_2 \quad (30)$$

$$K_{ij} = \iint \psi_i^*(\mathbf{x}_1) \psi_j(\mathbf{x}_1) \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \psi_i(\mathbf{x}_2) \psi_j^*(\mathbf{x}_2) d\mathbf{x}_1 d\mathbf{x}_2. \quad (31)$$

The  $J_{ij}$  are the Coulomb integrals and corresponds to the Coulomb interaction between the charge distribution of  $\psi_i$  and  $\psi_j$ . The  $K_{ij}$  are the exchange integrals. They can be view as the term responsible for the Pauli exclusion principle. The Coulomb and exchange integrals are real and  $J_{ij} \geq K_{ij} \geq 0$ . In the sum of equation (28) the terms  $J_{ii}$  are present which would correspond to an electron interacting with itself. The reason they can be included is that they are exactly canceled by  $K_{ii}$ . The minimizing of equation (28) with respect to the single-particle wave functions gives the  $N$  Hartree-Fock differential equations

$$\hat{F}\psi_i(\mathbf{x}) = \left( -\frac{1}{2} \nabla^2 + v + \hat{g} \right) \psi_i = \sum_{j=1}^N \varepsilon_{ij} \psi_j(\mathbf{x}). \quad (32)$$

Here the Coulomb-exchange operator  $\hat{g}(\mathbf{x})$  is given by

$$\hat{g} = \hat{j} - \hat{k}. \quad (33)$$

These operators are the mean field approximation of the true operators. For an arbitrary function  $f(\mathbf{x})$  they are given by

$$\hat{j}(\mathbf{x})f(\mathbf{x}) = \sum_{k=1}^N \int \psi_k^*(\mathbf{x}') \psi_k(\mathbf{x}') \frac{1}{|\mathbf{x} - \mathbf{x}'|} f(\mathbf{x}) d\mathbf{x}' \quad (34)$$

and

$$\hat{k}(\mathbf{x})f(\mathbf{x}) = \sum_{k=1}^N \int \psi_k^*(\mathbf{x}') f(\mathbf{x}') \frac{1}{|\mathbf{x} - \mathbf{x}'|} \psi_k(\mathbf{x}') d\mathbf{x}' \quad (35)$$

The matrix  $\varepsilon$  is hermitian and consists of the Lagrange multipliers associated with the orthonormalization of the single-particle wave functions. The orbital energies are given by

$$\varepsilon_i = \varepsilon_{ii} = H_i + \sum_{j=1}^N (J_{ij} - K_{ij}). \quad (36)$$

Then the total energy is given by

$$E_{\text{HF}} = \sum_{i=1}^N \varepsilon_i - \frac{1}{2} \sum_{i,j=1}^N (J_{ij} - K_{ij}). \quad (37)$$

To solve equations (32) one has to proceed iteratively since the single-particle wave functions are also present in the operators in equation (34) and (35). The equations are solve until self-consistency, meaning that the single-particle wave functions have converged. Because of this the method is also called self-consistent field method.

## 2.4 Density Functional Theory

This section will introduce the Density Functional Theory (DFT). First the main idea of DFT and the Hohenberg-Kohn theorems are shown. Then then Kohn-Sham (KS) formulation of DFT is introduced. The presentation is based on the textbook of Parr R. and Weitao Y. [12].

### 2.4.1 Electron Density

DFT is an alternate description to the Schrödinger equation (15). The main idea of DFT is to replace the  $3N$  dimensional wave function by the 3 dimensional electron density

$$\rho(\mathbf{r}) = N \int |\Psi(\mathbf{r}, s_1, \mathbf{x}_2, \dots, \mathbf{x}_N)|^2 ds d\mathbf{x}_2 \dots d\mathbf{x}_N. \quad (38)$$

The probability to find an electron in a volume element  $dV$  is given by  $p(\mathbf{r} + dV) = \rho(\mathbf{r})dV$ . The total number of electrons  $N$ , in the system is given by

$$\int \rho(\mathbf{r}) d\mathbf{r} = N. \quad (39)$$

### 2.4.2 Hohenberg-Kohn Theorems

The DFT is assured to give the correct description of the ground state a many electron system by the Hohenberg-Kohn theorems [13]. The first theorem states that: *The external potential  $v(\mathbf{r})$  is determined, within a trivial additive constant, by the electron density  $\rho(\mathbf{r})$ .* The number of electrons also is determined from the density by equation (39). The proof of this theorem comes from the minimum-energy principle for the ground state. Consider a non-degenerate ground state electron density  $\rho(\mathbf{r})$  of some  $N$ -electron system. Assume that there exists two external potentials  $v(\mathbf{r})$  and  $v'(\mathbf{r})$  that differ by more than a constant and have  $\rho(\mathbf{r})$  as the ground state electron-density. Having differing external potentials is the same as having different Hamiltonians  $\hat{H}$  and  $\hat{H}'$ , with different ground state wavefunctions  $|\Psi\rangle$  and  $|\Psi'\rangle$ . By the variation principle, equation (23), one finds

$$E_0 < \langle \Psi' | \hat{H} | \Psi' \rangle = \langle \Psi' | \hat{H}' | \Psi' \rangle + \langle \Psi' | \hat{H} - \hat{H}' | \Psi' \rangle = E'_0 + \int \rho(\mathbf{r}) [v(\mathbf{r}) - v'(\mathbf{r})] d\mathbf{r}, \quad (40)$$

where the last equality holds since only the external potential differ between the systems. Similarly one gets

$$E'_0 < \langle \Psi | \hat{H}' | \Psi \rangle = \langle \Psi | \hat{H} | \Psi \rangle + \langle \Psi | \hat{H}' - \hat{H} | \Psi \rangle = E_0 - \int \rho(\mathbf{r}) [v(\mathbf{r}) - v'(\mathbf{r})] d\mathbf{r}. \quad (41)$$

Combining equation (40) and (41) one obtains the contradiction  $E_0 + E'_0 < E'_0 + E_0$  and the external potentials have to be equal. Because of this the energy can be expressed as a functional of the electron density

$$E[\rho] = V_{ne}[\rho] + T[\rho] + V_{ee}[\rho] = \int \rho(\mathbf{r}) v(\mathbf{r}) d\mathbf{r} + F[\rho], \quad (42)$$

Where  $F[\rho]$  is an unknown universal functional that is independent of the system and describes the kinetic energy and electron-electron potential energy. The second theorem says: *For a trial density  $\tilde{\rho}(\mathbf{r})$  such that  $\tilde{\rho}(\mathbf{r}) \geq 0$  and  $\int \tilde{\rho}(\mathbf{r}) d\mathbf{r} = N$ ,*

$$E_0 \leq E[\tilde{\rho}] \quad (43)$$

where  $E[\tilde{\rho}]$  is the energy functional of equation (42). Thereby the variational principle also can be applied to this formulation. Assuming that  $E[\rho]$  is differentiable, the variational principle requires that the ground state satisfy

$$\delta \left( E[\rho] - \mu \left[ \int \rho(\mathbf{r}) d\mathbf{r} - N \right] \right) = 0 \quad (44)$$

where  $\mu$  is the chemical potential. This further gives the Euler-Lagrange equation

$$\mu = v(\mathbf{r}) + \frac{\delta F[\rho]}{\delta \rho(\mathbf{r})}. \quad (45)$$

The problem now arises to determine the functional  $F$ . If it were known exactly, the solution of equation (45) would yield the exact ground-state density. Unfortunately no such form of the functional is known and one needs to approximate it.

### 2.4.3 Kohn Sham

The direct numerical implementation of equation (45) has proven to be a difficult task. Another method was invented by Kohn and Sham in 1965 that is simpler to implement [14]. This Kohn-Sham (KS) method introduces orbitals analogous to what is done by HF. The kinetic energy can then be computed to a high accuracy, only leaving a small term that can be approximated separately. The introduction of orbitals also makes the evaluation of the electrostatic electron-electron interaction simpler. It is given by equation (30). The kinetic energy is given by

$$T = \sum_i^M n_i \langle \psi_i | -\frac{1}{2} \nabla^2 | \psi_i \rangle, \quad (46)$$

where  $n_i$  is the occupation number,  $\psi_i$  is the spin orbitals and  $M$  is the number of orbitals. The density is given by

$$\rho(\mathbf{r}) = \sum_i^M \sum_s n_i |\psi_i(\mathbf{r})|^2. \quad (47)$$

To simplify Kohn and Sham showed that one can build a theory by restricting the occupations numbers to be either  $n_i = 0$  or  $n_i = 1$ . With this the sums in equations (46) and (47) only need to go over the  $N$  occupied orbitals. Considering a non-interacting reference system with Hamiltonian

$$\hat{H}_s = \sum_i^N -\frac{1}{2} \nabla^2 + v_{\text{eff}}(\mathbf{r}_i), \quad (48)$$

where  $v_{\text{eff}}$  is the KS effective potential, given by

$$v_{\text{eff}}(\mathbf{r}) = v(\mathbf{r}) + \frac{\delta J[\rho]}{\delta \rho(\mathbf{r})} + \frac{\delta E_{\text{XC}}[\rho]}{\delta \rho(\mathbf{r})} = v(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + v_{\text{XC}}(\mathbf{r}). \quad (49)$$

where  $E_{\text{XC}}, v_{\text{XC}}$  is the exchange-correlation energy and potential. This system has the ground-state density given by equation (47) with  $n_i$  restricted as described. This can be seen by that the exact solution is given by the Slater determinant in equation (27), where the spin orbitals are the  $N$  lowest eigenstates of

$$\hat{h}_s \psi_i = \left( -\frac{1}{2} \nabla^2 + v_{\text{eff}}(\mathbf{r}) \right) \psi_i = \varepsilon_i \psi_i, \quad (50)$$

where  $\hat{h}_s$  is the one-electron Hamiltonian. The idea of Kohn and Sham was then to consider the approximate form of the kinetic energy  $T_s[\rho]$  as the exact kinetic energy of the system. This is done through rewriting equation (42) as

$$F[\rho] = T_s[\rho] + J[\rho] + E_{\text{XC}}[\rho]. \quad (51)$$

The exchange-correlation energy is defined as

$$E_{\text{XC}}[\rho] = T[\rho] - T_s[\rho] + V_{ee}[\rho] - J[\rho]. \quad (52)$$

Thus it contains the non-classical electron-electron potential and the difference between  $T$  and  $T_s$ . This now gives the energy the form

$$E[\rho] = T_s[\rho] + J[\rho] + E_{\text{XC}}[\rho] + \int \rho(\mathbf{r}) v(\mathbf{r}) d\mathbf{r}. \quad (53)$$

The KS-method can be considered a modification to the HF method. By introducing operator to include the correlation energy. The Kohn-Sham equations have to be solved iteratively in the same manner as the HF equations.

---

#### 2.4.4 Unrestricted Kohn Sham

So far only one electron density has been used, the total one. Then each single-particle wave function is implicitly assumed to be occupied by a spin up and spin down electron. The total number of electrons  $N$  is therefore split into  $N/2$  spin up and  $N/2$  down electrons. In unrestricted KS the number of spin up  $N_\uparrow$  and down orbitals  $N_\downarrow$  are allowed to fluctuate, but keeping the total number of electrons fixed, i.e.  $N = N_\uparrow + N_\downarrow$ . For unrestricted KS also the functionals must be modified to be dependent on the electron density of both spins. The spin up electron density is given by

$$\rho_\uparrow(\mathbf{r}) = \sum_{i=1}^{N_\uparrow} |\psi_{i,\uparrow}(\mathbf{r})|^2, \quad (54)$$

and spin down electron density

$$\rho_\downarrow(\mathbf{r}) = \sum_{i=1}^{N_\downarrow} |\psi_{i,\downarrow}(\mathbf{r})|^2, \quad (55)$$

where  $\psi_{i,\uparrow}, \psi_{i,\downarrow}$  is the  $i$ 'th wave function with spin up and down respectively. In this unrestricted approach the single-particle wave functions can have different spatial dependences and energies. One typically use the spin polarization

$$\zeta(\mathbf{r}) = \frac{\rho_\uparrow(\mathbf{r}) - \rho_\downarrow(\mathbf{r})}{\rho(\mathbf{r})} = \frac{\rho_\uparrow(\mathbf{r}) - \rho_\downarrow(\mathbf{r})}{\rho_\uparrow(\mathbf{r}) + \rho_\downarrow(\mathbf{r})}, \quad (56)$$

and the total electron density in spin dependent functionals, instead of the projected electron densities. The energy will in the unrestrictcted approach be given by

$$E[\rho, \zeta] = T_s[\rho] + J[\rho] + E_{XC}[\rho, \zeta] + \int \rho(\mathbf{r}) v(\mathbf{r}) d\mathbf{r}. \quad (57)$$

## 2.5 Exchange Correlation Functional

The exchange correlation functional of equation (52) has no known analytical form and therefore one needs to find suitable approximations. There exists a vast amount of different approximations, both empirical and non-empirical models. Here the non-empirical approximations will be introduced.

### 2.5.1 Local Density Approximation

The local density approximation (LDA) is the simplest of approximations. It assumes that the exchange-correlation only depends locally on the density. This implies that the exchange-correlation functional is given by

$$E_{XC}[\rho] = \int \rho(\mathbf{r}) \varepsilon_{XC}^{\text{LDA}}(\rho) d\mathbf{r}. \quad (58)$$

where  $\varepsilon_{XC}^{\text{LDA}}(\rho)$  is the exchange-correlation energy per particle of a uniform electron gas of density  $\rho$ . This term can be split into the exchange and the correlation part  $\varepsilon_{XC}^{\text{LDA}} = \varepsilon_X^{\text{LDA}} + \varepsilon_C^{\text{LDA}}$ . The exchange part was found analytically by P. Dirac [15], and is given by

$$\varepsilon_X^{\text{LDA}}(\rho) = -\frac{3}{4} \left( \frac{3}{\pi} \rho \right)^{1/3}. \quad (59)$$

### 2.5.2 Local Spin Density Approximation

When the system is spin-polarized, the functional is a function of both the density of spin up electron,  $\rho_\uparrow$ , and spin down electrons,  $\rho_\downarrow$ . When one extends the LDA to include spin, this is

called the local spin density approximation (LSDA). The common way of expressing the energy functionals for the LSDA is through the total density and The LSDA expression for the exchange energy per particle is

$$\varepsilon_X^{\text{LSDA}}(\rho, \zeta) = \varepsilon_X^0(\rho) + [\varepsilon_X^1(\rho) - \varepsilon_X^0(\rho)] f(\zeta), \quad (60)$$

where  $\varepsilon_X^0(\rho) = \varepsilon_X^{\text{LDA}}(\rho)$  is the un-spin-polarized energy and corresponds to  $\zeta = 0$ .  $\varepsilon_X^1(\rho)$  is the completely spin-polarized case, corresponding to  $\zeta = 1$  and it is given by  $\varepsilon_X^1(\rho) = 2^{1/3} \varepsilon_X^{\text{LDA}}(\rho)$ . The function  $f(\zeta)$  is given by

$$f(\zeta) = \frac{(1 + \zeta)^{4/3} + (1 - \zeta)^{4/3} - 2}{2(2^{1/3} - 1)}. \quad (61)$$

The correlation part is available thanks to quantum Monte Carlo simulations of the electron gas by Ceperley and Alder [16] and has its analytical form due to interpolation [17]. It is given by

$$\varepsilon_C^{\text{LDA}}(\rho, \zeta) = \varepsilon_c^0(\rho) + \alpha(\rho) \frac{f(\zeta)}{f''(\zeta)} [1 + \beta(\rho) \zeta^2] \quad (62)$$

where  $\alpha(\rho)$  is the spin stiffness and  $\beta(\rho)$  is chosen to satisfy

$$1 + \beta(\rho) = f''(0) \frac{\varepsilon_C^1(\rho) - \varepsilon_C^0(\rho)}{\alpha(\rho)}. \quad (63)$$

The  $\varepsilon_C^{0,1}$  are the un-spin-polarized and completely spin-polarized correlation energy per particle. The functions  $\alpha(\rho)$  and  $\beta(\rho)$  have tabulated values in the original paper.  $\varepsilon_C^{0,1}$  both have the approximate form

$$\varepsilon_C(\rho) = \frac{A}{2} \left\{ \ln \frac{x}{X(x)} + \frac{2b}{Q} \arctan \frac{Q}{2x + b} - \frac{bx_0}{X(x_0)} \left[ \ln \frac{(x - x_0)^2}{X(x)} + \frac{2(b + 2x_0)}{Q} \right] \arctan \frac{Q}{2x + b} \right\} \quad (64)$$

with  $x = (4\pi\rho/3)^{1/6}$ ,  $X(x) = x^2 + bx + c$  and  $Q = (4c - b^2)^{1/2}$ .  $A, x_0, b, c$  are constant that differ between the two correlation energies.

As the main approximation of the LDA and LSDA is that locally the electron-density behaves like a uniform gas, this means that is only a valid approximation for systems where the density is slowly varying. The electron density of atoms and molecules if far from this, but still the LDA and LSDA has proven successful in numerical applications.

### 2.5.3 General Gradient Approximation

The natural extension of the LDA is General Gradient Approximation (GGA). In GGA the local exchange-correlation energy per particle is assumed to also depend on the derivative of the electron-density. This gives a functional form of

$$E_{\text{XC}}[\rho] = \int \rho(\mathbf{r}) \varepsilon_{\text{XC}}^{\text{GGA}}(\rho, \nabla\rho) d\mathbf{r}. \quad (65)$$

As with in the LDA approximation the energy can be separated into a exchange and a correlation part. The common way of expressing the exchange functional is

$$E_X^{\text{GGA}}[\rho] = \int \rho(\mathbf{r}) \varepsilon_X^{\text{LDA}}(\rho) F_X^{\text{GGA}}(s) d\mathbf{r} \quad (66)$$

where  $\varepsilon_X^{\text{LDA}}$  is the exchange energy per particle given in equation (59) and  $F_X^{\text{GGA}}(s)$  is the exchange enhancement factor in reference to the LDA case. The dimensionless variable  $s$  is called the reduced density gradient, and is given by

$$s = \frac{|\nabla\rho|}{2(3\pi^2)^{1/3} \rho^{4/3}}. \quad (67)$$

---

The Perdew-Burke-Ernzerhof(PBE) functional [8] is one of the most common functionals and its exchange enhancement factor is given by

$$F_X^{\text{PBE}}(s) = 1 + \kappa - \kappa \left( 1 + \frac{\mu s^2}{\kappa} \right) \quad (68)$$

where  $\mu \approx 0.21951$  and  $\kappa = 0.804$  and are numerical constants derived for physical constrains. The correlation functional is dependent on the partial spin densities  $\rho_\uparrow, \rho_\downarrow$  an not the total density as for the exchange functional. The functional form is

$$E_C^{\text{GGA}} = \int \rho [\varepsilon_C^{\text{LDA}}(r_s, \zeta) + H(r_s, \zeta, t)] \quad (69)$$

where  $r_s$  is the local Seitz radius and  $t$  is a dimensionless gradient, which also is dependent on a spin scaling factor  $\phi$  and the Thomas-Fermi screening vector  $k_s$ . They are all given by

$$r_s = \left( \frac{4\pi\rho}{3} \right)^{1/3}, \quad t = \frac{|\Delta\rho|}{2\phi k_s \rho}, \quad \phi(\zeta) = \frac{(1+\zeta)^{2/3} + (1-\zeta)^{2/3}}{2}, \quad k_s \sqrt{\frac{4k_F}{\pi a_0}}, \quad (70)$$

where  $a_0$  is the Bohr radius. By considering several limiting cases a the form of  $H$  can be found to be

$$H = \frac{e^2}{a_0} \gamma \phi^3 \ln \left[ 1 + \frac{\beta}{\gamma} t^2 \left( \frac{1+At^2}{1+At^2+A^2t^4} \right) \right], \quad (71)$$

where

$$A = \frac{\beta}{\gamma} \left( e^{-\varepsilon_C^{\text{LDA}} / (\gamma \phi^3 e^2 / a_0)} - 1 \right)^{-1} \quad (72)$$

and  $\beta \approx 0.066725$ ,  $\gamma \approx 0.031091$  are numerical constants.

## 2.6 Basis Sets and Pseudopotentials

In order to use either the Hartree-Fock or the Kohn-Sham method, one needs to approximate the orbitals in some manner. The unknown orbitals are expanded in a known set of basis functions. This procedure is exact as long as the basis functions one expands in are complete. But in order to guarantee this, one needs to include an infinite number of basis functions. The number of basis functions to include is determined by the required accuracy. As a larger accuracy implies a larger computational cost, one wants to use as small of a basis set as possible given the accuracy needed. The number of basis functions necessary for a given accuracy and system, depends heavily on the type of basis functions used. For larger atoms there are several electrons close to the core that are inert and do not contribute to the interaction with other atoms. Therefore one would like to not do calculations for the core electrons while still keeping the description of the other electrons correct. This is the goal of pseudopotentials. In this section common basis sets and pseudopotentials are introduced based on the textbooks by Jensen [18] and Leach [19].

### 2.6.1 Slater

Slater functions are centered on the atoms. They have the functional form

$$\chi(r, \theta, \varphi) = CY_{l,m}(\theta, \varphi)r^{n-1}e^{-\xi r}. \quad (73)$$

Here  $C$  is a normalization constant,  $Y_{l,m}$  are the spherical harmonics,  $n$  is the order and  $\xi > 0$  is a constant dependent on the charge of the atom. An advantage of the Slater function is that they give the correct electron density close to the nuclei. The electron density should have a cusp at the position of the nuclei, as asserted by Kato's theorem [20]. The theorem states that for generalized Coulomb potentials

$$Z_k = \lim_{\mathbf{r} \rightarrow \mathbf{R}_k} -\frac{a_0}{2\rho(\mathbf{r})} \frac{d\rho(\mathbf{r})}{dr}, \quad (74)$$

where  $Z_k, \mathbf{R}_k$  is the atomic number and position of the central atom. A disadvantage of the Slater functions is that they are very difficult to evaluate accurately. This makes them impossible to use for many electron systems.

### 2.6.2 Gaussian

A Gaussian basis function is a function on the form

$$f(\mathbf{r}) = Cx^a y^b z^c e^{-\alpha(x^2+y^2+z^2)} \quad (75)$$

where  $\alpha > 0$  determines the width,  $a, b, c \in \{\mathbb{N}, 0\}$  determine the spatial symmetries of the function and  $C$  is a normalization constant. The functions are grouped by orders with the  $n$ 'th order given by functions which satisfy  $a + b + c = n$ . The major advantage of Gaussian basis functions over Slater functions is the fact that there exist way more efficient methods for evaluating integrals of Gaussian basis functions. Thus they are better suited for many electron problems. The Gaussian basis function does not give the physically correct description of the electron density cusp, as given by equation (74). To give a more accurate description a linear combination of Gaussian basis function with different  $\alpha$  values can be used to approximate a Slater function. However this will for a finite set of Gaussian basis function, not change the limiting behavior of the electron density.

### 2.6.3 Plane Wave

Both Slater and Gaussian basis function describe localized electrons. For metallic systems and crystal structures the valance electrons are usually delocalized from the nuclei. To better describe these electrons plane waves can be used. For a system with periodic boundary conditions one can utilise the Bloch's Theorem. It states that the wave function of a periodic system can be described by Bloch waves [21]

$$\psi_{\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} u_{\mathbf{k}}(\mathbf{r}) \quad (76)$$

where  $u_{\mathbf{k}}(\mathbf{r})$  has the same periodicity as the system. It can be Fourier expanded with regards to the reciprocal lattice vector  $\mathbf{G}$  which gives [11]

$$\psi_{\mathbf{k}} = \sum_{\mathbf{G}} C(\mathbf{k} - \mathbf{G}) e^{i(\mathbf{k}-\mathbf{G})\cdot\mathbf{x}} \quad (77)$$

where in principle the sum of  $\mathbf{G}$  is over all lattice vectors and  $C(\mathbf{k} - \mathbf{G})$  is the weighting of the plane waves. Conventionally only plane waves with energy less than a cutoff energy  $E_{\text{cut}}$ , meaning that they satisfy

$$\frac{1}{2} |\mathbf{k} - \mathbf{G}|^2 \leq E_{\text{cut}}, \quad (78)$$

are used in a basis set. For plane waves basis set to effectively describe the core electrons it requires a large cutoff energy to describe the fast oscillations. This would also require a large computational cost, thus making it a unfavorable basis set for core electrons.

### 2.6.4 Augmented-plane-wave

Since the Gaussian basis functions provide a satisfactory description of the localized electrons and the plane wave basis set a good description for the delocalized electrons the idea of the augmented-plane-wave approach is to combine these two descriptions. Around each nuclei a augmentation sphere is introduced and inside these augmentation spheres the Gaussian basis functions are used to describe the wave function. In the space in-between these augmentation spheres the wave function is described by plane waves. The two wave functions have to be joined continuously at the surfaces of the augmentation spheres [22]. A natural extension of the augmented-plane-wave method is the linearized augmented-plane-wave method. It requires that both the wave function and its derivative is continuous over the surfaces of the augmentation spheres [23].

### 2.6.5 Pseudopotential

For heavier atoms there is a increased amount of electrons. Most of the electrons are core electrons that does not contribute to the interaction with other atoms. Therefore one would like to disregard

---

the orbitals description of the core electrons. Simply removing these orbitals will alter the valence orbitals drastically. Thus one need to replace the removed orbitals by some approximate method to obtain the correct description of valence electrons. Another effect of removing the explicit description of the innermost electrons is that the basis set can be reduced. Since the core electrons are the most localized and requires additional basis function to be properly described. For heavier atoms relativistic effects also become important. This is due to the high velocity of the most tightly bond electrons. To give a correct description this effect has to be taken into account. Both of these problems can be addressed using pseudopotentials. The idea is to use an effective core potential that approximately capture the two above mentioned properties. Allowing for only the valence electron orbitals to be included. This significantly reduces the computational cost of systems with heavier atoms [24].

### 2.6.6 Projector Augmented-Wave

The projector-augmented wave method was introduced by Blöch [25]. The method combines ideas from both the linearized augmented-plane-wave method and pseudopotentials making it an efficient, close to all-electron orbital, method [24]. This is achieved by using a linear transformation from the true orbitals onto so called pseudo-orbitals inside the augmentation spheres. Outside the augmentation spheres the true orbitals and pseudo-orbitals are identical. This transformation is obtained using projector functions which are orthonormal to the pseudo-orbitals and localized to the augmentation spheres.

## 2.7 Force Fields

The advancements in DFT calculations have made it possible to study systems with hundreds of atoms with high accuracy and reasonable computation time. But sometimes one does not need this accuracy or want to simulate larger systems, or wants do do molecular simulation over long time spans. Then computationally cheaper methods are needed. Force field methods does not solve the system on the quantum mechanical level. Instead force field methods approximate the true interactions between atoms by effective interactions. This makes the computational cost for one evaluation drastically smaller. Force field does not describe the electrons and only depend on the positions of the nuclei. The forces field are parametric functions. The parameters are either fitted by experimental data or some higher level computations, e.g. DFT.[18]

### 2.7.1 Effective Medium Theory

The Effective Medium Theory(EMT) is based on the idea of calculation the energy of an atom in an arbitrarily environment by first calculating it in a reference system, the effective medium. Then the the energy difference between the reference system and the actual system is estimated [26]. One writes the total energy of the system as

$$E = \sum_{i=1}^N E_{c,i} + \left( E - \sum_{i=1}^N E_{c,i} \right), \quad (79)$$

where the  $E_{c,i}$  is the energy in the reference system of atom  $i$ . The essential part is now that the reference system is close in nature to the real system such that the last term of equation (79) can be considered with perturbation theory or some other approximation. For the bonding in the transition metals having face-center cubic crystals, which includes Ni, the EMT has been parameterized by Jacobsen et al. [26]. The parametrization was based on DFT calculations. The energy corrections could be separated into two terms

$$E = \sum_{i=1}^N E_{c,i} + \Delta E_{AS} + \Delta_{1el} \quad (80)$$

where  $\Delta E_{AS}$  is called the atomic-sphere correction and  $\Delta_{1el}$  the one-electron correction. The atomic-sphere interaction correct for the difference in electrostatic and exchange-correlation energy

---

for the system compared to the reference, they modeled this using a effective pair interaction. The one-electron correction is the difference between the one electron energies in the system and the reference. This correction is also modeled as a pair potential, and thus the whole correction is approximated as a pair interaction.

### 3 Numerical Methods

In this section the numerical methods used to analyse the Ni clusters are introduced. To begin with, a local optimizer will be introduced. Then the main aspects of using DFT in VASP are discussed. Finally the global energy minimization method is explained together with its constituent algorithms.

#### 3.1 Local Optimizers

The local optimizers find the geometry of the nearest energy minima. When using teh Born-Oppenheimer approximation, a system of  $M$  atoms has an energy given be the geometry, i.e.  $E(\mathbf{R}_1, \dots, \mathbf{R}_M)$ , where  $\mathbf{R}_i$  are the nuclei-positions. The energy is a multi-dimensional complex function and has several minima. The minimum with lowest energy is the global minimum, while the other minima are local. The local optimization methods are iterative methods that are repeated until some convergence criteria is reached. All local optimization methods presented in this section require the gradient

$$\mathbf{g} = \nabla E = \left( \frac{\partial E}{\partial x_1}, \frac{\partial E}{\partial y_1}, \frac{\partial E}{\partial z_1}, \dots, \frac{\partial E}{\partial y_M}, \frac{\partial E}{\partial z_M} \right), \quad (81)$$

where  $x_i, y_i, z_i$  refer to the Cartesian components of the position of atom  $i$ . Some methods also require the Hessian  $\mathbf{H}$ , given by the matrix

$$\mathbf{H} = \begin{bmatrix} \frac{\partial^2 E}{\partial x_1^2} & \frac{\partial^2 E}{\partial x_1 \partial y_1} & \dots & \frac{\partial^2 E}{\partial x_1 \partial z_M} \\ \frac{\partial^2 E}{\partial y_1 \partial x_1} & \frac{\partial^2 E}{\partial y_1^2} & \dots & \frac{\partial^2 E}{\partial y_1 \partial z_M} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 E}{\partial z_M \partial x_1} & \frac{\partial^2 E}{\partial z_N \partial y_1} & \dots & \frac{\partial^2 E}{\partial z_M^2} \end{bmatrix} \quad (82)$$

The description of the methods are based on the textbook of Jensen [18].

##### 3.1.1 Steepest Descent

The Steepest Descent is the simplest method given a gradient  $\mathbf{g}$ . Since the gradient points in the direction the function grows fastest, one can lower the value by stepping in the opposite direction. How large the step in this direction should be, could either be determined by a line search along that direction to find the optimal size, or have some predetermined value. There exists several methods for choosing this predetermined step size. Some methods also have a adaptive step size which based on the previous step's energies.

##### 3.1.2 Conjugate Gradient Algorithm

The conjugate gradient algorithm addresses one of the main problems with the steepest descent algorithm, the fact that the subsequent steps partially cancel each other. This is done by making the next steps perpendicular to the previous steps. In particular, this can be summarized by the step  $i$ 'th step direction  $\mathbf{d}_i$ , which is given by

$$\mathbf{d}_i = -\mathbf{g}_i + \beta_i \mathbf{g}_{i-1}. \quad (83)$$

---

Here  $\beta_i$  is a weighting of the previous steps. Several weighting methods exist. When the steps should be truly perpendicular all  $\beta_i = 1$ . In general one also might need to restart the values as all the perpendicular directions has been used.

### 3.1.3 Newton Method

In the above methods, only the gradient has been used to find the optimal direction. The Newton method improves on this by utilizing the Hessian  $\mathbf{H}$ . The method relies on the Taylor expansion of the function to be minimized around the current position  $x_i$ , namely

$$E(\mathbf{x}) \approx E(\mathbf{x}_i) + \mathbf{g}_i^T(\mathbf{x} - \mathbf{x}_i) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_i)^T \mathbf{H}_i(\mathbf{x} - \mathbf{x}_i), \quad (84)$$

where  $\mathbf{g}_i = \mathbf{g}(\mathbf{x}_i)$  and  $\mathbf{H}_i = \mathbf{H}(\mathbf{x}_i)$ . If one assumes that the function is purely quadratic, i.e.  $\mathbf{H}_i = \mathbf{H}(\mathbf{x})$ . The minimum is found by optimizing the step. The step is then given by

$$\mathbf{s}_i = -\mathbf{H}_i^{-1}\mathbf{g}_i. \quad (85)$$

The Newton method has quadratic convergence and is therefore better at locating the minima. The drawback is that the computation of the Hessian and the inversion of it for large systems is computationally expensive.

### 3.1.4 Broyden–Fletcher–Goldfarb–Shanno algorithm

The Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm is a quasi-Newton method. This means that it uses the same update method as the Newton method. However, instead of calculating the Hessian at each iteration and inverting it, the inverse of the Hessian is iteratively approximated using the gradients from previous steps [27]. This is significantly cheaper than calculating and inverting the Hessian. A good approximation for the inverse of the Hessian is most beneficial close to the minimum energy. The algorithm typically has had enough iterations to get a good estimation of the inverse of the Hessian before it is close to the energy minimum. The benefit of a Hessian is therefore present without the need to compute it at each iteration.

## 3.2 Density Functional Theory in the Vienna Ab-initio Simulation Package

Vienna Ab-initio Simulation Package (VASP) is a software package for doing atomic scale modeling using first principles [28–30]. This is done using DFT within the Kohn-Sham approach solved in the self-consistent manner presented earlier. The exchange-correlation energy functional used is the PBE functional[8]. One can select from several different pseudopotentials and basis sets. In this work the Projector augmented wave method [25] is used. For geometry optimizations native to VASP, the Conjugate Gradient Algorithm is used.

### 3.2.1 Periodic Boundary Conditions

VASP utilises periodic boundary conditions (PBC) in all directions. This means that it is effective at calculating bulk properties of crystals. When wanting to investigate non-periodic structures in one or several directions within this PBC framework, one can add vacuum in the non-periodic directions. This addition of vacuum effectively removes the structures self-interaction in these directions since the distance becomes large. The disadvantage of this is that one now uses computational resources on modeling vacuum. If the cutoff energy is kept fixed while increasing the size unit cell, this increases the computational cost while the energy resolution remains the same. When modeling a free structure, one needs vacuum padding on all sides. However, this padding should be kept to a minimum, so one does not increase the computational cost. For finding this optimal value, a cubic unit cell for the Ni clusters is used, with a varying lattice constant.

---

### 3.2.2 Smearing

Using the Kohn-Sham method for a periodic system requires evaluation of many integrals of periodic functions over the Brillouin Zone. These integrals have slow convergence rates for systems having discontinuous state occupancy close to the Fermi surface. As is the case for metals, like Ni. Another problem can also arise in the self-consistency iterations. The energy level of a state can shift in energy during an iteration causing the state that was occupied, to be unoccupied and vice versa. This can cause convergence problems as the charge distribution change strongly between two consecutive self-consistent steps. The smearing method was introduced by Methfessel and Paxton to address these issues [31]. The method consists of smoothing out the discontinuous occupancy function by a smooth function. In this work, Gaussian functions are used for smearing, and the method is therefore called Gaussian smearing. The smearing width of the Gaussian functions is  $\sigma$ . When doing calculations, this value needs to be set. A larger value leads to faster convergence, but less accurate results. Thus one needs to optimize smearing width to the system in question and the desired accuracy.

### 3.2.3 Parameter Tuning

As mentioned, when using VASP for a particular system one needs to tune the parameters for the calculation. The most important parameters when considering a free system is the lattice constant, the energy cutoff for the plane waves used as in (78) and the smearing width  $\sigma$  used in the Gaussian smearing method. These parameters are found by varying one at a time until the calculation has the desired properties. For the lattice constant and energy cutoff, this is a converged energy. However, for the smearing width the difference between the energy without entropy and the energy approximation for  $\sigma \rightarrow 0$  yielded by the calculations, should be small.

## 3.3 Global Energy Minimization Method

Finding the ground state geometry of a Ni cluster is a global energy minimization problem. The local optimizer presented earlier find the nearest energy minimum. The lowest energy minimum found from a set could be the global minimum, but there is no guarantee that it is. In order to be confident that the lowest energy minimum is the global minimum, one needs to sample many times. This should be done while trying to minimize the computational cost. A DFT evaluation is the significantly most expensive operation and therefore it should be kept to a minimum. To achieve this a four step model has been used. First a random geometry search is done using the EMT force field. This search yields many potential candidates. Secondly these potential candidates are filtered using the Oganov Fingerprint and their relative energy. The lowest energy candidates are fully relaxed using DFT. Thirdly the lowest energy candidates and their relaxation trajectory are used as training data for the Global Optimization with First-principles Energy Expressions (GOFEE) algorithm [9]. The algorithm also gives several potential candidates. The forth step is to filter the candidates form GOFEE in the same manner as for the ones form the random search. The remaining candidates are fully relaxed using DFT. The geometry that has the lowest energy of all the DFT evaluated geometries, is taken to be the global minimum. In the rest of this section the different constituent part of this method are presented.

### 3.3.1 Random Geometry Search

A random search is the simplest and least biased search method. A implementation of the method using the *Atomistic Global Optimization X* (AGOX) [32] python package is shown in appendix A. Searching for a optimal geometry consists of two main steps. Firstly a starting geometry if generated and then it is locally optimized. The starting geometries are generated by first placing an atoms at random position in the unit cell. Then iteratively a new atom is placed randomly inside any of the spherical shells centered at the already present atoms with a inner radius of 75% and outer radius 125% of the covalent bonding radius between the center atom and the atom to be

---

placed. The final geometry is made by placing all atoms in this way. As the structures generated are randomly generated they have no bias towards certain structures. This is both an advantage and a disadvantage. It assured that many qualitative unique geometries are locally optimized, but this also implies that several geometries are locally optimized even though they have high in energies.

### 3.3.2 Oganov Fingerprint

In order to filter out similar geometries from a set of candidates the Oganov fingerprint is used[33]. The fingerprint is based on the pair of atom types and their radial distribution. The whole fingerprint is a matrix of function where each entry if given by

$$F_{AB}(R) \sum_{A_i} \sum_{B_j} \frac{\delta(R - R_{ij})}{4\pi R_{ij}^2 (N_A N_B / V) \Delta} - 1, \quad (86)$$

where  $i, j$  runs over all the atoms of type  $A, N_A$  and type  $B, N_B$  respectively,  $R_{ij}$  is the distance between the pair of atoms,  $V$  is the volume of the unit cell and  $\delta$  is the Dirac delta function. Each peak is smeared by a Gaussian of width  $\sigma = 0.02$  Å and accumulated into a histogram of bin size  $\Delta = 0.05$  Å. The good properties of the Oganov fingerprint as a descriptor is that it is independent of translation and rotation of the structure, ordering of atoms and inversion or mirroring of the structure.

### 3.3.3 Global Optimization with First-principles Energy Expressions

The algorithm is developed by Malthe K. Bisbo and Bjørk Hammer [9]. A flowchart of the algorithm is presented in figure 1 and an implementation using AGOX is shown in appendix B. The working principle of GOFEE is to minimize the number of first-principle (FP) calculations needed to be done by utilising machine-learning (ML) techniques. Due to the fact that ML models require reliable data to interpolate well this makes using a predefined database hard. Therefore a more suitable approach is to start from a small incomplete training database and then progressively add more samples as to augment and improve the ML model, which is called active learning. When choosing which next candidate to propose for doing a FP calculation this is where the GOFEE algorithm comes into play.

The model consists of the following steps each iteration. 1; Several candidates are generated. These are split between two generation methods random generation and rattle generation. 2; All candidates are relaxed in the surrogate potential. 3; The acquisition function determines which of the relaxed candidates that should be evaluated with FP calculations. 4; FP calculations are done with a number of relaxations steps, using the BFSG algorithm. 5; The surrogate potential is updated with the new data. 6; The results are written to the database.

#### *Generators*

The generators purpose to propose candidates that are subject to further evaluation. The two used here are the random generator and the rattle generator. The random generator is the same generation process as described in section 3.3.1. The rattle generator works by perturbing an already existing structure. This is done by picking structures from a K-means sampler. This K-means sampler performs K-means clustering on the previously evaluated candidates and then chooses candidates from distinct clusters to assure diverse sampled candidates [34]. When the rattle generator then perturbs the structure it displaces a certain number of atoms and checks that the still satisfy the confinements of the problem. This is perturbed structure does not stay within the confinements, another perturbation of the original structure is done. In this work there are no confinements so all perturbations are accepted.

#### *Gaussian Process Regression*

Gaussian process (GP) regressions used as the model for the surrogate potential. The GP is used to find a probability distribution over energy surrogate functions  $p(E_{\text{sur}} | \mathbf{X}, \mathbf{E})$ , where  $\mathbf{X} =$

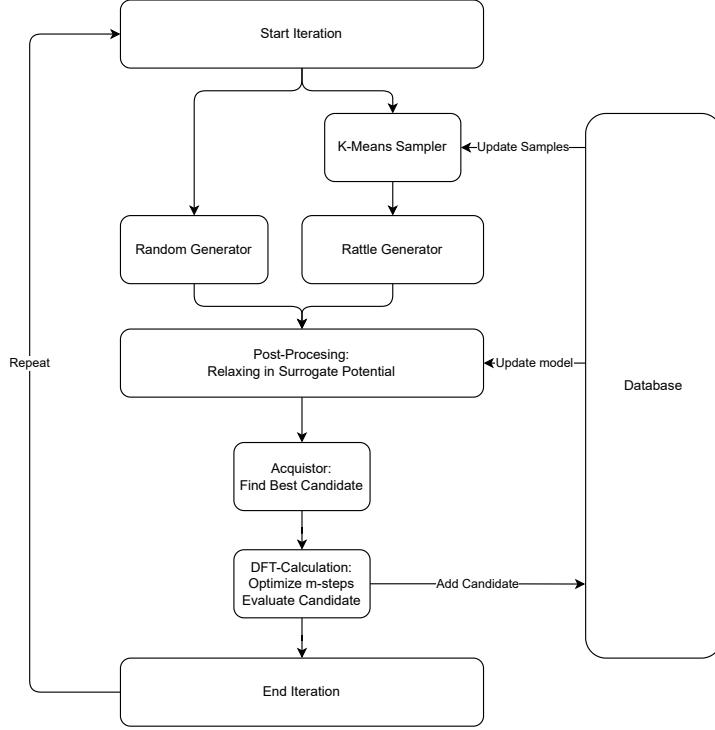


Figure 1: Flowchart of the GOFEE-algorithm.

$(\mathbf{x}_1, \dots, \mathbf{x}_N)^T$  and  $\mathbf{E} = (\mathbf{E}_1, \dots, \mathbf{E}_N)^T$  are the previously observed atomic configurations and energies respectively. The atomic configuration  $x_i$  is a feature vector that can contain any feature that seems relevant. This can amount to using the Cartesian coordinates of the atoms, but in the GOFEE algorithm the Oganov fingerprint is used. The advantage of the GP model is that in addition to the prediction  $E_{\text{sur}}(\mathbf{x})$  there is also the predicted uncertainty  $\sigma_{\text{sur}}(\mathbf{x})$  of the prediction. The GP is specified by the prior mean  $\mu(\mathbf{x})$  and covariance  $k(\mathbf{x}, \mathbf{x}')$ . Given these priors a new structure is predicted using

$$E_{\text{sur}}(\mathbf{x}') = \mathbf{k}'^T (K + \sigma_n^2 I)^{-1} + \mu(\mathbf{x}), \quad (87)$$

$$\sigma_{\text{sur}}^2(\mathbf{x}') = k(\mathbf{x}', \mathbf{x}') - \mathbf{k}'^T (K + \sigma_n^2 I)^{-1} \mathbf{k}', \quad (88)$$

where  $\mathbf{k}' = k(\mathbf{X}, \mathbf{x}')$  and  $K = k(\mathbf{X}, \mathbf{X})$  where the  $\sigma_n^2 = 1 \times 10^{-5} \text{ eV}^2$  is for regularization (avoiding overfitting to training-data). The prior mean is given by

$$\mu(\mathbf{x}) \propto \sum_{i < j} \left( \frac{0.7 r_{\text{CD}ij}}{r_{ij}} \right)^2, \quad (89)$$

where  $r_{ij}$  and  $r_{\text{CD}ij}$  is the distance and the covalent distance respectively between atom  $i$  and  $j$ . The covariance function is given by

$$k(\mathbf{x}, \mathbf{x}') = \theta_0 \left[ (1 - \beta) e^{-(\mathbf{x}-\mathbf{x}')^2/(2\lambda_1^2)} + \beta e^{-(\mathbf{x}-\mathbf{x}')^2/(2\lambda_2^2)} \right], \quad (90)$$

where  $\theta_0$  is the maximum covariance,  $\beta = 0.01$  determines the weighting and  $\lambda_1$  and  $\lambda_2$  are the characteristic lengths. These parameters are automatically determined by maximizing the marginal likelihood [35] of the model given the data. They are also bounded by  $\lambda_1, \lambda_2 \in [1, 10^2]$  and  $\theta_0 \in [0, 10^6]$

#### Acquisition function

When the search is conducted most of the local relaxations are done in this surrogate potential. This means that the model itself guides the exploration of the configuration space and this means that unexplored predicted high energy areas tends to be unexplored. To cope with this an acquisition function  $f(\mathbf{x})$  is used, namely

$$f(\mathbf{x}) = E_{\text{sur}}(\mathbf{x}) - \kappa \sigma_{\text{sur}}, \quad (91)$$

| Cutoff energy | Smearing width | $E_{SC}$ | $F_{GO}$  | LC Ni <sub>8</sub> | LC Ni <sub>12</sub> | LC Ni <sub>30</sub> |
|---------------|----------------|----------|-----------|--------------------|---------------------|---------------------|
| 400 eV        | 0.02 eV        | 0.1 meV  | 0.01 eV/Å | 12 Å               | 14 Å                | 16 Å                |

Table 1: The table shows the parameters used for the calculations in VASP. LC is the lattice constant of the cubic unit cell.  $E_{SC}$  is the convergence criteria for the energy in the self-consistent loop.  $F_{GO}$  convergence criteria for the largest force when doing geometry optimizations.

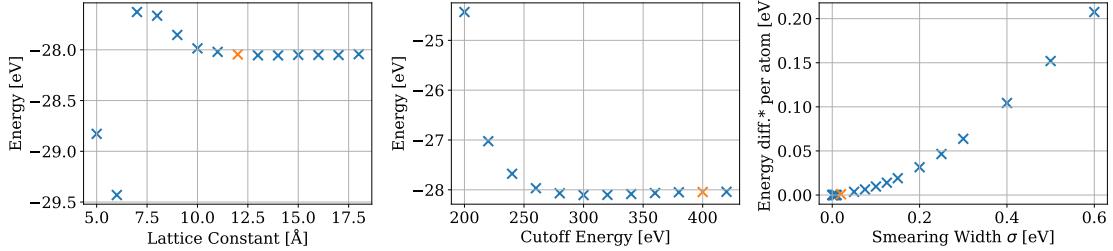


Figure 2: The figures show the parameter search for the Ni<sub>8</sub> cluster. The chosen value highlighted in orange. \*This is the difference in the energy without entropy and the energy approximated for when the smearing width goes to zero.

where  $\kappa$  is a tunable parameter determining the importance of exploration. The candidates are ranked according to their acquisition function value and the one with the best ranking (lowest value) is chosen for FP calculations.

## 4 Results and Discussion

### 4.1 Parameter Search

The parameter search using the Ni<sub>8</sub> cluster is shown in figure 2. The figures are in accordance with what one would expect from theory. For the search over lattice constants, a small lattice constant makes the cluster effectively interact with itself, which lowers the energy. When the lattice constant is increased, the self interaction disappears and the energy converges. The energy is seen to be converging for a lattice constant of 12 Å. The effect of an increased cutoff energy is to increase the basis set size, since more plane waves are included. Because of the variational principle, the increase should make the energy go down. This is the main trend in the figure. However, the energy is slightly lower for a cutoff energy around 300 eV as compared to the higher cutoff energy values. This is because VASP has implemented a correction term that helps acquire faster convergence. The energy has converged for a cutoff energy of 400 eV, therefore this value is used. The difference between the energy without entropy and the energy approximated for when the smearing width goes to zero, decreases monotonically with decreasing smearing width as one would expect. The break condition for the self-consistency loop is 0.1 meV. The smearing width chosen is 0.02 eV and it gives a difference in energy slightly below 1 meV. For the geometry optimizations, the break condition is set to be that all forces are below 0.01 eV/Å.

Since the Ni<sub>12</sub> and Ni<sub>30</sub> clusters have the same interactions as the Ni<sub>8</sub> cluster, the same cutoff energy and smearing width is used. The increase in number of atoms also implies a larger structure. In order to prevent the increased size to make the cluster interact with itself, the lattice constant of the cubic unit cell is increased by 2 Å and 4 Å for the Ni<sub>12</sub> and Ni<sub>30</sub> cluster respectively. The parameters used are summarized in table 1.

| System | E [eV] | CE [eV] | IP [eV] | EA [eV] | BL [ $\text{\AA}$ ] | M [ $\mu_B$ ] | Geometry  |
|--------|--------|---------|---------|---------|---------------------|---------------|---|
| B      | 0.0    | -3.044  | 4.075   | -2.697  | 2.36(0.1)           | 8             |  |
| T-BCO  | 0.067  | -3.036  | 4.073   | -2.704  | 2.33(0.04)          | 8             |  |
| CPBP   | 0.205  | -3.018  | 4.071   | -2.65   | 2.36(0.08)          | 8             |  |
| TCTBP  | 0.299  | -3.007  | 4.127   | -2.605  | 2.34(0.08)          | 8             |  |
| CBCO   | 0.358  | -2.999  | 4.151   | -2.805  | 2.33(0.01)          | 8             |  |

Table 2: Summary of the properties of the 5 lowest energy structures of  $\text{Ni}_8$ . The properties are E energy relative to the lowest one, CE cohesive energy, IP ionization potential, EA electron affinity, BL Bond length with standard deviation in parenthesis and M total magnetic moment. The system names are shortcuts for their geometric description. They are from the top Bidisphenoid, Trans-bibappedoctahedron, Cappedpentagonalbipyramid, Tricappedtriagonaltipyramid and Cis-bicappedoctahedron.

## 4.2 $\text{Ni}_8$ cluster

The properties of the five lowest energy geometries for the  $\text{Ni}_8$  cluster found from the full search are shown in table 2. For the energy, all structures are within a range of 0.36 eV. Therefore also the cohesive energies are comparable. For the IP and EA, the energy range is 0.1 eV and 0.2 eV respectively. Also the bond lengths are very similar, whereas all magnetisations are equal.

There are three core geometric structures that are present in the low energy  $\text{Ni}_8$  cluster geometries. They are the octahedron, pentagonalbipyramid and tetrahedron. The octahedron is the core structure of the bidisphenoid (B), the cis-bicappedoctahedron (C-BCO) and the trans-bicappedoctahedron (T-BCO). In all three structures, a main octahedron has been capped at different sites. For C-BCO and T-BCO, this does not change the shape of the octahedron, but for B it is altered. The pentagonalbipyramid is the core of the cappedpentagonalbipyramid (CPBP). The tetrahedron is the only sub-structure of the tricappedtriagonaltipyramid (TCTBP). The capped atoms of T-BCO, C-BCO and CPBP also have a tetrahedron structure.

The lowest energy structure was the bidisphenoid, also called snub disphenoid or triangular dodecahedron. The structure is in agreement with a previous study by Song et al. using DFT with the PBE functional [36]. Song et al. also reported cohesive energy, IP, EA and magnetization for the bidishpenoid. Their reported magnetization was equal to the one found in this work. Their cohesive energy, IP and EA where all around 0.5 eV higher in absolute values. Their IP and EA calculations were done adiabatic, wheres in this work they are done vertically. This could explain the difference, but they reported that the ionic system had identical geometries. The calculations of Song et al. are done with a smaller cutoff energy of 269.5 eV, which might be responsible for some of the discrepancies. They also did not report what smearing method was used so this also might contribute to the difference.

The DOS for the bidisphenoid is shown in figure 3 and the DOS for the other  $\text{Ni}_8$  structures are shown in appendix C figure 8. Here the DOS for all the geometries are similar. For the bidisphenoid, the HOMO-LUMO gap for the up states is 1.45 eV and in the down sates the gap is 0.10 eV.

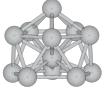
| System | E [eV] | CE [eV] | IP [eV] | EA [eV] | BL [ $\text{\AA}$ ] | M [ $\mu_B$ ] | Geometry  |
|--------|--------|---------|---------|---------|---------------------|---------------|---|
| TI     | 0.0    | -3.284  | 3.994   | -2.622  | 2.39(0.06)          | 8             |  |
| TM     | 0.19   | -3.268  | 4.061   | -2.891  | 2.38(0.07)          | 10            |  |
| DOL    | 0.198  | -3.268  | 4.083   | -2.85   | 2.35(0.04)          | 10            |  |
| MPBP   | 0.455  | -3.246  | 4.068   | -2.876  | 2.38(0.1)           | 10            |  |
| BCDO   | 0.469  | -3.245  | 4.145   | -2.86   | 2.35(0.03)          | 10            |  |

Table 3: Summary of the properties of the 5 lowest energy structures of  $\text{Ni}_{12}$ . The properties are E energy relative to the lowest one, CE cohesive energy, IP ionization potential, EA electron affinity, BL Bond length with standard deviation in parenthesis and M total magnetic moment. The system names are shortcuts for their geometric description. They are from the top truncated icosahedron, tetrahedron mesh, double octahedronlike, merged pentagonal bipyramidal and bicapped double octahedron.

For all geometries, the effective band gap is always the gap between the down states and is in the range 0.1-0.2 eV. For the up states it is significantly larger from about 0.5 to 1.5 eV.

A display of the spatial distribution of the HOMO and LUMO for both spin are shown in appendix D figure 11. Here one can clearly see that the up-spin HOMO and LUMO are delocalized, while the down HOMO and LUMO are strongly localized on the Ni atoms.

### 4.3 $\text{Ni}_{12}$ cluster

The properties of the five lowest energy geometries for the  $\text{Ni}_{12}$  cluster found from the full search are shown in table 3. All geometries are within a range of 0.5 eV. The small difference in energy,

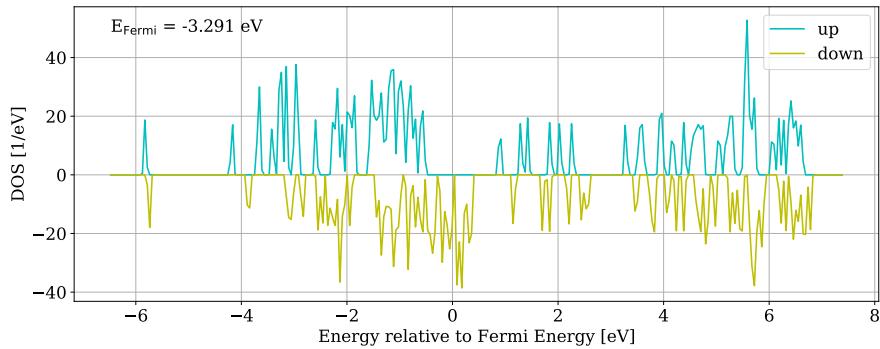


Figure 3: Density of state for Bidispheoid, the lowest energy geometry of the  $\text{Ni}_8$  cluster.

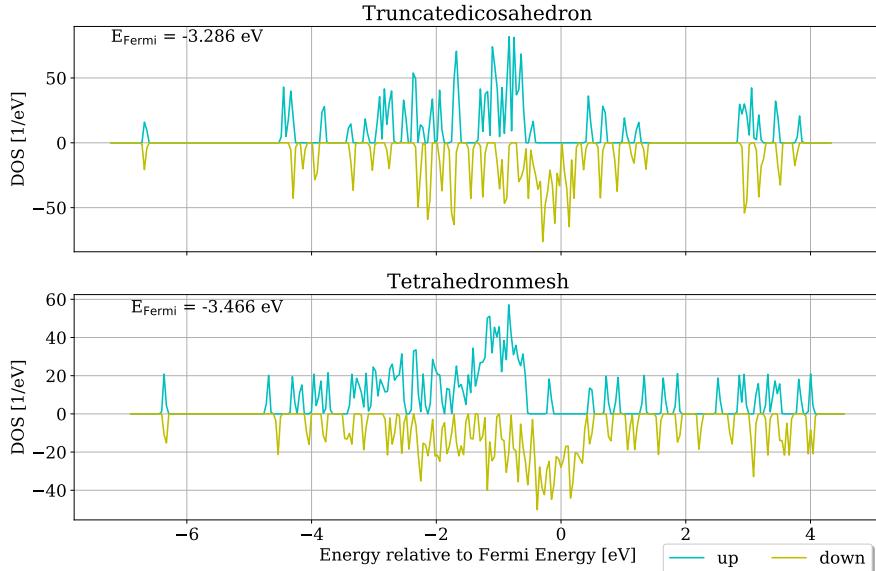


Figure 4: The density of state for the truncated icosahedron and the tetrahedron mesh geometries of the  $\text{Ni}_{12}$  cluster.

leads to similar cohesive energy for all. The structures also have comparable values of IP and EA, except for the EA of the truncated icosahedron (TI), which is 0.25 eV lower than the other EAs. The bond lengths have slightly increased compared to the  $\text{Ni}_8$  clusters. The TI also has a  $2\mu_B$  lower magnetization than the other structures.

The same core geometries are present in the  $\text{Ni}_{12}$  clusters, as in the  $\text{Ni}_8$ , the octahedron, the pentagonal bipyramid and the tetrahedron. There is also the icosahedron which can be made up of two pentagonal bipyramids sharing a top atom. This is the structure of the TI, but with one atom removed. The tetrahedron mesh (TM) is a combination of many tetrahedron sharing sides. Both the double octahedronlike (DOL) and the bicapped double octahedron (BCDO) consist of a main structure of two joint octahedrons. Their difference lies in how the additional atoms are placed. The merged pentagonalbipyramid (MPBP) resembles two pentagonal bipyramids that are joint at an angle relative to each other.

The lowest energy geometry is the TI. This is in agreement with the study by Song et al.[36]. Their reported magnetic moment agrees with the moment found here. As for the  $\text{Ni}_8$  cluster, the values for cohesive energy, IP and EA found by Song et al. are larger, 0.6 eV, 0.4 eV and 0.2 eV respectively. Thus, there are some discrepancies, even though the method for DFT calculation is equal, apart from the energy cutoff.

The DOS for the TI and the TM are shown in figure 4. There is a small gap for the TI, whereas for the TM the gap has closed and therefore it is metal like. There is a reduced gap in the up-spin states as compared to the  $\text{Ni}_8$  geometries. The DOS for all the structures are shown in appendix C figure 9. All other structures have a gap, except for the DOL.

The HOMO and LUMO for both spins are shown in appendix D figure 12. All the down orbitals are localized to the atoms, whereas all up orbitals are delocalized, except for the HOMO orbital on the TI. This orbital is localized in the same manner as the down orbitals.

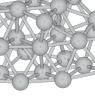
| System | E [eV] | CE [eV] | IP [eV] | EA [eV] | BL [ $\text{\AA}$ ] | M [ $\mu_B$ ] | Geometry  |
|--------|--------|---------|---------|---------|---------------------|---------------|---|
| IL1    | 0.0    | -3.758  | 3.551   | -2.789  | 2.49(0.24)          | 22            |  |
| IL2    | 0.213  | -3.751  | 3.613   | -2.85   | 2.48(0.21)          | 24            |  |
| 3LHCPL | 0.864  | -3.729  | 3.553   | -2.813  | 2.49(0.26)          | 24            |  |
| Song   | -0.792 | -3.784  | 3.548   | -2.776  | 2.48(0.22)          | 20            |  |
| Luo    | 0.628  | -3.737  | 3.691   | -2.903  | 2.45(0.13)          | 24            |  |

Table 4: Summary of the properties of the 3 lowest energy structures of  $\text{Ni}_{30}$  found and two reference structures from the literature Song [36] and Luo [37]. The properties for the reference structure are calculated using VASP and not the ones reported by the references. Only the reported geometric structure is used when calculating the properties. The properties are E energy relative to the lowest one(not including literature), CE cohesive energy, IP ionization potential, EA electron affinity, BL Bond length with standard deviation in parenthesis, M total magnetic moment and the geometry. From the top the shorthands are icosahedronlike variant 1, icosahedronlike variant 2 and 3 Layer HCPLike.

#### 4.4 $\text{Ni}_{30}$ cluster

The properties for three lowest energy  $\text{Ni}_{30}$  geometries found from the search are listed in 4. The lowest energy structure found by Song et al.[36] and from the research of Luo [37] are also included. The geometries are taken from their work, but the properties listed are based on calculations done in the same manner as for the other geometries. Luo used tight-binding molecular dynamic calculations with the simulated annealing method to obtain the result. The three structures found from the search have a larger energy spread than the  $\text{Ni}_8$  and  $\text{Ni}_{12}$  clusters. Since the energy difference per atom is anyhow quite small, the relative differences in the cohesive energies are also quite small, about 0.03 eV. The difference in IP and EA are also modest with a difference of around 0.06 eV for both. The bond lengths increased form the  $\text{Ni}_{12}$  structure up to  $2.48 - 2.49 \text{ \AA}$  and is similar to the bond length of bulk metallic nickel,  $2.491 \text{ \AA}$  [38]. The magnetization difference is  $2\mu_B$ .

The low energy geometries for the  $\text{Ni}_{30}$  cluster is considerably less ordered than the others. Both the icosahedronlike variant 1 (IL1) and 2 (IL2) have an icosahedron as constituent part, but the additional atoms are not added around the icosahedron in an orderly way. The Song et al. geometry also has an icosahedron as a building block. In contrast to the two icosahedron like variants, the additional atoms are structured in a more ordered manner. The Luo geometry is composed of four hexagonal bipyramids that are joint by the tops and rotated with respect to each other. The 3 layer HCPLike geometry does not resemble a polyhedron structure, but there are three distinct layers that are hexagonal closed packed (HCP). The HCP structure is close to the FCC structure of metallic Ni, but with the difference being in how the layers are shifted with respect to each other[39].

The lowest found energy candidate was the IL1. It had a 0.8 eV higher energy than the structure from Song et al.. Thus the full search did not find the lowest energy geometry. The properties reported for the Song et al. geometry in table 4 are calculated using the same parameters as for the other calculations. Comparing these results to the reported properties from Song et al. one finds that the magnetization and the ionization potential are the same. However, the cohesive energy is

0.6 eV higher and the EA is 0.3 eV higher than the results found here. This indicates that there is some method discrepancies, as it is unlikely that the difference is caused by the difference in cutoff energy. A possible explanation for the difference might be the smearing method or smearing width used. Given that the smearing width gives an error of the order of meV, this is also not likely to be the cause.

The energy of the geometry found by Luo is higher than the lowest energy found through the search, but lower in energy than the 3LHCPL. Therefore the Luo geometry would be a geometry expected to be found by the search. This again might indicate that the addition of some generator that favours symmetry might be useful. The lack of the Luo geometry from the search also shows that the configurations space is not well explored.

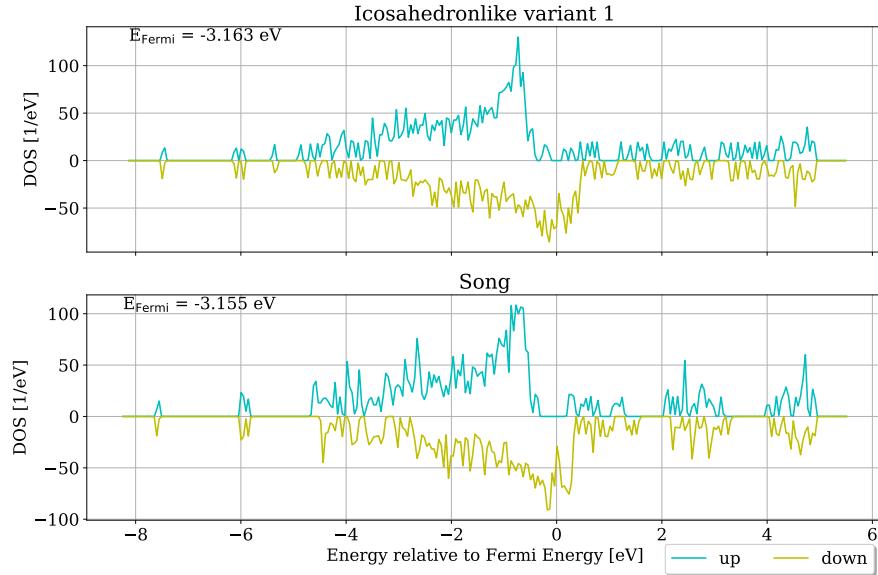


Figure 5: The density of state for the best  $\text{Ni}_{30}$  geometry, icosahedronlike variant 1 and the reference geometry of Song et al. [36].

The DOS for the IL1 and the Song reference structure is shown in figure 5. The DOS is similar for the two. There is no longer a gap between the Fermi energy and the systems have become nearly-metallic. There is a small gap in the up states of 0.32 eV and 0.61 eV for the IL1 and the Song reference respectively. The DOS for the other structures are similar to the ones presented and also have a closed gap. The other DOS are shown in appendix C figure 10

The HOMO and LUMO for both spins are shown in appendix D figure 13. The main takeaway is that the up HOMO and LUMO are strongly delocalized. This is in agreement with the fact that the system have become nearly metallic, since metals have nearly free electrons. The down orbitals on the other hand, are localized to the atoms.

## 4.5 Global Energy Minimization Method

Some preliminary use of the GOFEE algorithm showed that it did struggle with learning the surrogate landscape of DFT when no initial training data was used. This might be due to the need for several steps for VASP to relax the candidate significantly. This is why an initial screening using the EMT was adopted. A random geometry search was chosen since the force field is computationally cheap and the added computational cost of a machine learning method would likely not give much added performance. The fact that the search is unbiased is also good, because this gives a diverse pool of candidates.

Figure 6 shows the distribution of candidate energies produced by the GOFEE algorithm for the Ni<sub>30</sub> cluster. The algorithm samples energies nearly uniformly in a 5 eV range. From the maximum force on a individual atom in the proposed candidates, it is clear that the structures need to be post relaxed. This is in contrast to what is reported by the developers of AGOX [32]. They stated that GOFEE were able to find the lowest energy structures without post-relaxation. This difference might be due to the fact that they showed the performance using EMT as a replacement for first principle calculations. So the added complexity of DFT calculations is most likely harder to learn for the surrogate potential. The distribution of candidate energies and forces are similar for the other cluster sizes and are shown in appendix E.

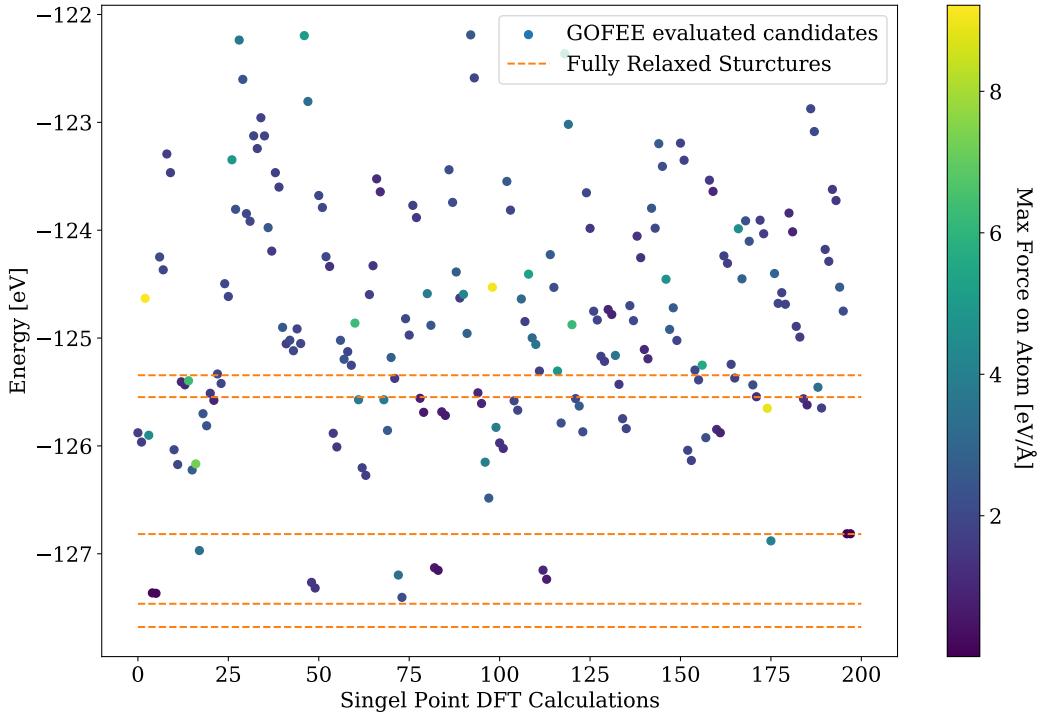
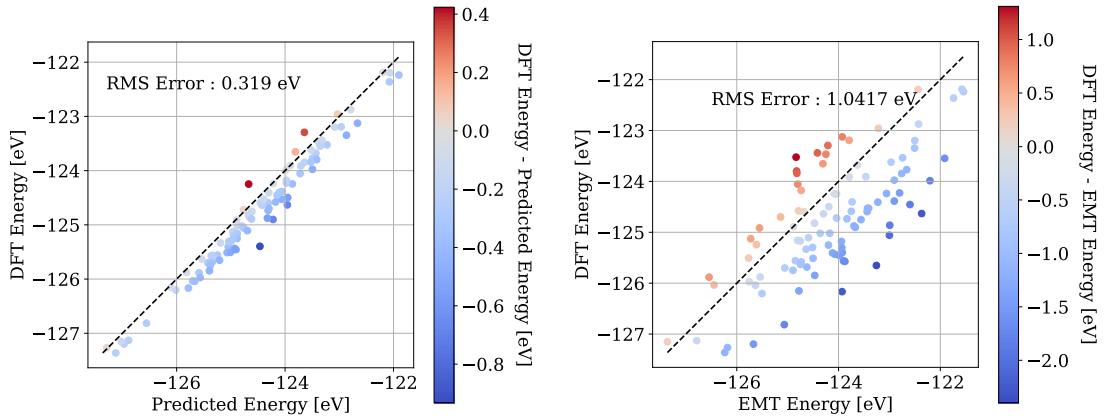


Figure 6: the figure shows the distribution of candidate (DFT) energies for the use of the GOFEE algorithm for the Ni<sub>30</sub> cluster with 100 iterations, 1 update step (2 DFT evaluations each iteration), 20 candidates each iteration and a kappa value of 2. The points are colored based on the largest force on an individual atom for that candidate. The dashed lines are the energies of the fully relaxed structures, that together with the relaxation trajectory, is used as training data for the surrogate potential.

In figure 7(a) a comparison between the surrogate model predictions and the DFT energies are shown. In figure 7(b) a comparison between EMT energies and DFT energies is shown for the same set of structure. From the two distributions one can clearly see that the surrogate potential model gives an improved description of the DFT energies, compared to the EMT. In general, the surrogate potential tends to model the energy as too high compared to DFT. The EMT on the other hand, over estimates some and underestimate other compared to DFT. The root mean square error of the surrogate potential is around a third of the EMT one. It is clear that the surrogate model gives an improved description of the DFT energies compared to the EMT. Thus making the computational expense of training and using the model worthwhile.

For the GOFEE algorithm the same setup of 100 iterations, 1 DFT update step, 20 candidates each iteration (14 rattle and 6 random) is used for all the cluster sizes. By having 20 candidates in each iteration the algorithm makes good use of the surrogate potential. This is because a large amount of candidates gives a wide range of geometries for the acquisition function to choose from. This increases the chance that the best one is relevant for the DFT calculations. Using one DFT



(a) The distribution of the predicted energies by the surrogate model and the DFT energies. The dashed line shows where the two energies are identical.

(b) The distribution of the EMT energies and the DFT energies. The dashed line shows where the two energies are identical. As the energy offset of the two theories are different they have been translated so the lowest EMT energy is the same as the lowest DFT energy.

Figure 7: The figure show the comparison between the surrogate potential against DFT and EMT againt DFT. The geometries that are used are the candidates proposed by the GOFEE algorithm for the Ni<sub>30</sub> cluster with 100 iterations, 1 update step (2 DFT evaluations each iteration), 20 candidates each iteration and a kappa value of 2.

update step lets the DFT calculations better guide the exploration as this moves the structure closer to a local minimum. To increase the effectiveness of the GOFEE algorithm these parameters could be tuned. Due to the large computational cost of actually determining optimal value, this was not seen as beneficial.

For the Ni<sub>8</sub> and Ni<sub>12</sub> cluster, the global search yielded the same optimal structure as Song et al.. Thus for these simpler systems the method works fine. For the Ni<sub>30</sub> cluster the lowest energy structure is 0.8 eV above the optimal structure found by Song er al.. This is an considerable amount. This shows that the current parameter is not adequate for the more complex system. As a first possible solution the number of iterations for the Ni<sub>30</sub> should be increased. It has a significantly larger configuration space than the smaller clusters and that needs to be reflected in the number of iterations. A possible reason for not finding the optimal structure might be the usage of an inaccurate force field for the initial screening. As EMT only is an approximation it could have some energetically favored DFT geometries be energetically unfavorable. If there is a significantly large discrepancy, no geometries close to the minimum would be locally optimized. This would cause the surrogate model to have a worse initial description and make it harder for GOFEE to find the global minimum.

For the geometries found for the Ni<sub>30</sub> cluster there is an apparent lack of symmetries. This might be due to a poor exploration of possible geometries and a longer run would eventually find more symmetric geometries. Nevertheless there might be an idea to also include some candidate generator that also enforce some symmetries on the generated candidates.

## 5 Conclusion

Global energy minimization of three different sized Ni clusters has been done at the density functional theory level using the Pedrew-Bruke-Ernzerhof functional. The Ni cluster sizes were of 8, 12 and 30 atoms. For the Ni<sub>8</sub> and Ni<sub>12</sub> clusters, the lowest energy geometries coincided with the

---

literature. The cohesive energy, ionization potential, electron affinity, bond length and magnetization for the lowest geometries was reported. The magnetic moment and bond lengths agree with the literature references, whereas there were some discrepancies in the cohesive energy, ionization potential and electron affinity. These discrepancies might partly be due to differing calculation parameters. In addition the difference in the ionization potential and electron affinity is possibly caused by the calculation being done vertical and not adiabatic as in the literature. The density of states was computed for all low energy geometries found. There was a tendency for the system to be more metallic as the cluster size increased. This also agrees with the computed highest occupied and lowest unoccupied molecular orbitals that were increasingly delocalized.

A global energy minimization method has been used to reduce the computational cost. The method used effective medium theory and a random geometry search to do an initial screening for candidates. The most promising and distinct candidates were fully relaxed using density functional theory and served as training data for a surrogate potential. The surrogate potential was a part of the Global Optimization with First-principles Energy Expressions algorithm that also generated candidates. The algorithm generated high force geometries, and therefore the best were chosen to fully relax. This was in contrast to the reported effectiveness of the algorithm by the creators.

The presented method for finding the global energy minimum geometry of nickel clusters performed sufficiently for smaller cluster sizes of 8 and 12 atoms. For the Ni<sub>30</sub> cluster, the method did not find the global minimum. This might be due to a poorly explored configuration space and could be improved by increasing the number of iterations. Another reason for not finding the optimal geometry, might be the use of effective medium theory in the screening, as it does not necessarily have the same lowest energy structures as density functional theory. The idea of introducing a generator that proposes candidates with certain symmetries, might also be an improvement.

The outlook for the master thesis will be to improve the energy minimization method such that the correct minimum energy geometry for the Ni<sub>30</sub> cluster is found. The method will further be used to map out adsorption sites for the syngas and for the reactants. The trajectories between these states will be mapped out so the whole Fisher-Tropsch synthesis is simulated.

## References

- [1] Hannah Ritchie, Max Roser and Pablo Rosado. ‘Energy’. In: *Our World in Data* (2022). URL: <https://ourworldindata.org/energy>.
- [2] Mitchell P. Jones, Theresa Krexner and Alexander Bismarck. ‘Repurposing Fischer-Tropsch and natural gas as bridging technologies for the energy revolution’. In: *Energy Conversion and Management* 267 (Sept. 2022). ISSN: 01968904. DOI: 10.1016/j.enconman.2022.115882.
- [3] Yongjun Jiang et al. ‘Recent advances in thermocatalytic hydrogenation of carbon dioxide to light olefins and liquid fuels via modified Fischer-Tropsch pathway’. In: *Journal of CO<sub>2</sub> Utilization* 67 (Jan. 2023), p. 102321. ISSN: 22129820. DOI: 10.1016/j.jcou.2022.102321.
- [4] Hans Schulz. ‘Short history and present trends of Fischer-Tropsch synthesis’. In: *Applied Catalysis A: General* 186 (1999), pp. 3–12.
- [5] Michela Martinelli et al. ‘An overview of Fischer-Tropsch Synthesis: XtL processes, catalysts and reactors’. In: *Applied Catalysis A: General* 608 (Nov. 2020). ISSN: 0926860X. DOI: 10.1016/j.apcata.2020.117740.
- [6] Zahra Teimouri, Nicolas Abatzoglou and Ajay K. Dalai. ‘Kinetics and selectivity study of fischer-tropsch synthesis to c5+ hydrocarbons: A review’. In: *Catalysts* 11.3 (Mar. 2021), pp. 1–32. ISSN: 20734344. DOI: 10.3390/catal11030330.
- [7] Jyoti Kapil, Pramila Shukla and Ashish Pathak. ‘Review Article on Density Functional Theory’. In: *Recent Trends in Materials and Devices*. Ed. by V.K. Jain, S. Rattan and A. Verma. Vol. 256. Singapore: Springer, 2020, pp. 211–220. DOI: 10.1007/978-981-15-8625-5\\_\\_22.
- [8] John P Perdew, Kieron Burke and Matthias Ernzerhof. ‘Generalized Gradient Approximation Made Simple’. In: *Phys. Rev. Letters* 77.18 (1996), pp. 3865–3868.

- 
- [9] Malthe K. Bisbo and Bjørk Hammer. ‘Efficient Global Structure Optimization with a Machine-Learned Surrogate Model’. In: *Phys. Rev. Letters* 124.086102 (Feb. 2020). ISSN: 10797114. DOI: 10.1103/PhysRevLett.124.086102.
  - [10] Steven S. Zumdahl. *Chemical Principles*. 6th. CENAGE Learning, 2009.
  - [11] Charles Kittel. *Introduction to Solid State Physics*. 3rd ed. John Wiley & Sons, Inc, 2005. ISBN: 0-471-41526-X.
  - [12] Robert G Parr and Yang Weitao. *Density-Functional Theory of Atoms and Molecules*. English. International Series of Monographs on Chemistry Vol. 16. Oxford University Press, 1989. ISBN: 9780195092769. URL: <https://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=169751&site=ehost-live&scope=site>.
  - [13] P Hohenberg and W Kohn. ‘Inhomogeneous Electron Gas’. In: *Phys. Rev.* 136.3B (Nov. 1964), B 864–B 871. DOI: 10.1103/PhysRev.136.B864. URL: <https://link.aps.org/doi/10.1103/PhysRev.136.B864>.
  - [14] W Kohn and L J Sham. ‘Self-Consistent Equations Including Exchange and Correlation Effects’. In: *Phys. Rev. A* 140.4 (1965), pp. 1133–1128.
  - [15] P. A.M. Dirac. ‘Note on Exchange Phenomena in the Thomas Atom’. In: *Mathematical Proceedings of the Cambridge Philosophical Society* 26.3 (1930), pp. 376–385. ISSN: 14698064. DOI: 10.1017/S0305004100016108.
  - [16] D M Ceperley and B J Alder. ‘Ground State of the Electron Gas by a Stochastic Method’. In: *Phys. Rev. Letters* 45.7 (Aug. 1980), pp. 566–569. DOI: 10.1103/PhysRevLett.45.566. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.45.566>.
  - [17] L Wilk A N and D M Nusair. ‘Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis’. In: *J. Phys.* 58 (1980), pp. 1200–1211.
  - [18] Frank Jensen. *Introduction to Computational Chemistry*. Second. Wiley, 2007.
  - [19] Andrew R. Leach. *Molecular Modeling: Principles and Applications*. 2nd ed. Prentice Hall, 2001.
  - [20] N. H. March. ‘Spatially dependent generalization of Kato’s theorem for atomic closed shells in a bare Coulomb field’. In: *Phys. Rev. A* 33.1 (1986), pp. 88–89.
  - [21] Felix Bloch. ‘Über die Quantenmechanik der Elektronen in Kristallgittern’. In: *Zeitschrift für physik* 52.7 (1929), pp. 555–600.
  - [22] John C Slater. ‘Wave Functions in a Periodic Potential’. In: *Phys. Rev.* 51.10 (1937), pp. 846–851. DOI: 10.1103/PhysRev.51.846.
  - [23] Krogh O. Andersen. ‘Linear methods in band theory’. In: *Phys. Rev. B* 12.8 (1975), pp. 3060–3083. ISSN: 01631829. DOI: 10.1103/PhysRevB.12.3060.
  - [24] Peter Schwerdtfeger. ‘The pseudopotential approximation in electronic structure theory’. In: *ChemPhysChem* 12 (Dec. 2011), pp. 3143–3155. ISSN: 14397641. DOI: 10.1002/cphc.201100387.
  - [25] P E Blöchl. ‘Projector augmented-wave method’. In: *Phys. Rev. B* 50.24 (1994), pp. 17953–17979. DOI: 10.1103/PhysRevB.50.17953.
  - [26] K W Jacobsen, P Stoltze and J K Norskov. *A semi-empirical effective medium theory for metals and alloys*. Tech. rep. 1996, pp. 394–402.
  - [27] J. F. (Joseph Frédéric) Bonnans. *Numerical optimization : theoretical and practical aspects*. Springer, 2006, p. 490. ISBN: 354035445X.
  - [28] G Kresse and J Hafner. ‘Ab. initio molecular dynamics for liquid metals’. In: *Phys. Rev. B* 47.1 (1993), pp. 558–561. DOI: 10.1103/PhysRevB.47.558. URL: <https://link.aps.org/doi/10.1103/PhysRevB.47.558>.
  - [29] G Kresse and J Furthmüller. ‘Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set’. In: *Computational Materials Science* 6 (1996), pp. 15–50. DOI: 10.1016/0927-0256(96)00008-0. URL: <http://www.sciencedirect.com/science/article/pii/0927025696000080>.

- 
- [30] G Kresse and J Furthmüller. ‘Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set’. In: *Phys. Rev. B* 54.16 (1996), pp. 11169–11186. DOI: 10.1103/PhysRevB.54.11169. URL: <https://link.aps.org/doi/10.1103/PhysRevB.54.11169>.
  - [31] M Methfessel and A T Paxton. ‘High-precision sampling for Brillouin-zone integration in metals’. In: *Phys. Rev. B* 40.6 (1989), pp. 3616–3621. DOI: 10.1103/PhysRevB.40.3616.
  - [32] Mads-Peter Verner Christiansen, Nikolaj Rønne and Bjørk Hammer. ‘Atomistic Global Optimization X: A Python package for optimization of atomistic structures’. In: *J. Chem. Phys* 157 (2022). DOI: 10.1063/5.0094165. URL: <http://dx.doi.org/10.1063/5.0094165>.
  - [33] Mario Valle and Artem R. Oganov. ‘Crystal fingerprint space - A novel paradigm for studying crystal-structure sets’. In: *Acta Cryst. A* 66 (Sept. 2010), pp. 507–517. ISSN: 01087673. DOI: 10.1107/S0108767310026395.
  - [34] Lindsay R Merte et al. ‘Structure of an Ultrathin Oxide on Pt 3 Sn(111) Solved by Machine Learning Enhanced Global Optimization’. In: *Angew. Chem. Int.* 61.e2022042 (2022). DOI: 10.26434/chemrxiv-2021-mtbq2.
  - [35] Carl Edward. Rasmussen and Christopher K. I. Williams. *Gaussian processes for machine learning*. MIT Press, 2006, p. 248. ISBN: 026218253X.
  - [36] Wei Song et al. ‘Magnetic and electronic properties of the nickel clusters Ni<sub>n</sub> (n<=30)’. In: *Comput. Theor. Chem* 978.1-3 (Dec. 2011), pp. 41–46. ISSN: 2210271X. DOI: 10.1016/j.comptc.2011.09.028.
  - [37] Chenglin Luo. ‘A possible packing sequence of nickel clusters: Ni 13-Ni 32’. In: *New J. Phys.* 4.10 (2002), pp. 1–8. URL: <http://www.njp.org/>.
  - [38] John W Annthony et al. ‘Nickel’. In: *Handbook of Mineralogy. Vol. I. Chantilly, VA, US: Mineralogical Society of America* (1990).
  - [39] William D. Callister Jr. and David G. Rethwisch. *Materials Science and Engineering: An Introduction*. 10th. Wiley, 2018. ISBN: 9781119321590.

---

## Appendix

### A Code - Random Geometry Search

This appendix shows the implementation of the random geometry search method using the Atomistic Global Optimization X package.

```
import matplotlib

matplotlib.use("Agg")

import numpy as np
from agox import AGOX
from agox.databases import Database
from agox.environments import Environment
from agox.evaluators import LocalOptimizationEvaluator
from agox.generators import RandomGenerator
from agox.postprocessors import CenteringPostProcess

from ase import Atoms

from argparse import ArgumentParser

parser = ArgumentParser()
parser.add_argument("-i", "--run_idx", type=int, default=0)
args = parser.parse_args()

# Calculator
from ase.calculators.emt import EMT

calc = EMT()

# System & general settings:
template = Atoms("", cell=np.eye(3) * 16, pbc=True)
confinement_cell = np.eye(3) * 16
confinement_corner = np.array([0, 0, 0])
environment = Environment(
    template=template,
    symbols="Ni30",
    confinement_cell=confinement_cell,
    confinement_corner=confinement_corner,
)

# Database
db_path = "db{}.db".format(args.run_idx) # From input argument!
database = Database(filename=db_path, order=7, write_frequency=1)

# Search Settings:
random_generator = RandomGenerator(
    environment=environment,
    **environment.get_confinement(),
    sets={"set_key": "candidates"},
    order=1
)
# Centers the cluster in the cell
centering = CenteringPostProcess(order=2)
```

---

```
evaluator = LocalOptimizationEvaluator(
    calc,
    gets={"get_key": "candidates"},
    optimizer_kwargs={"logfile": None},
    store_trajectory=False,
    optimizer_run_kwargs={"fmax": 0.01, "steps": 400},
    order=5,
)

# Run the program
agox = AGOX(random_generator, database, evaluator, centering, seed=args.run_idx)
agox.run(N_iterations=500)
```

---

## B Code - Global Optimization with First-principles Energy Expressions

This appendix shows the implementation of the Global Optimization with First-principles Energy Expressions algorithm using the Atomistic Global Optimization X package.

```
import matplotlib
matplotlib.use('Agg')

import numpy as np
from agox import AGOX
from agox.databases import Database
from agox.environments import Environment
from agox.evaluators import LocalOptimizationEvaluator
from agox.generators import RandomGenerator, RattleGenerator
from agox.samplers import KMeansSampler
from agox.models import ModelGPR
from agox.acquisitors import LowerConfidenceBoundAcquisitor
from agox.postprocessors import RelaxPostprocess, CenteringPostProcess
from agox.postprocessors import ParallelRelaxPostprocess
from agox.evaluators import SinglePointEvaluator
from agox.collectors import StandardCollector
from agox.collectors import ParallelCollector
from agox.candidates import StandardCandidate
import os

from ase.io.vasp import read_vasp_xml
from ase import Atoms

# personally made functions to interact with vaspfiles
from vasputils import*

# setup VASP calculator
from ase.calculators.vasp import Vasp
calc = Vasp(
    prec = 'Accurate',
    istart = 0, # start new calculation
    xc = "PBE", # set PBE exchange correlation functional
    ismear = 0, # Gaussian smearing
    sigma = 0.02, # smearing constant
    encut = 400, # energy cutoff
    nelm = 80, # max number of SC steps default = 60
    isym = -1, # ignore symmetries
    ispin = 2, # for spin calculations
)

# creat empty template
template = Atoms(' ', positions= [], cell=np.eye(3)*14, pbc = True)
# set whole cell as the available space
confinement_cell = np.eye(3) * 14
confinement_corner = np.array([0, 0, 0])
# describe the enviornment for AGOX
environment = Environment(template=template, symbols='Ni12',
                           confinement_cell=confinement_cell, confinement_corner=confinement_corner)

# Database
db_path = 'db.db'
```

---

```

database = Database(filename=db_path, order=7, write_frequency=1, initialize =
    ↵  True)

# fill database with relaxed structures
file_path = [os.path.join("xml", f"{i}.xml") for i in range(7)]
for file in file_path:
    print("Done with :", file)
    for i,_ in enumerate(get_all_energies_xml(file)[0]):
        struc = next(read_vasp_xml(filename=file, index = i))
        candidate = StandardCandidate(template=template, **struc.todict())
        candidate.calc = struc.calc
        database.store_candidate(candidate)

# initialize default Gaussian Process Regression model
model = ModelGPR.default(environment, database)
model.iteration_start_training = 0

# setup generator through a sampler
sample_size = 20
sampler = KMeansSampler(feature_calculator=model.get_feature_calculator(),
    database=database, sample_size=sample_size, order=1)

rattle_generator = RattleGenerator(use_mic = False,
    ↵  **environment.get_confinement())
random_generator = RandomGenerator(use_mic = False,
    ↵  **environment.get_confinement())
generators = [random_generator, rattle_generator]
num_candidates = {0:[6, 14]} # distribution between candidates

# controls the sampling from both generators
collector = StandardCollector(generators=generators, sampler=sampler,
    environment=environment, num_candidates=num_candidates, order = 2)

# Center the candidate in the unit cell
centering = CenteringPostProcess(order = 3)

# Acuistor function initialization
acquisitor = LowerConfidenceBoundAcquisitor(model_calculator=model,
    kappa=2, order=4)

# Relaxing in surrogate potential
relaxer = RelaxPostprocess(model=acquisitor.get_acquisition_calculator(),
    constraints=environment.get_constraints(), order = 5, start_relax=0)

# VASP evaluation and one step
evaluator = LocalOptimizationEvaluator(calc,
    gets={'get_key':'prioritized_candidates'},
    optimizer_kwargs={'logfile':None}, store_trajectory=True,
    optimizer_run_kwargs={'fmax':0.01, 'steps':1}, order=6)

# collect agox modules and run algorithm
agox = AGOX(collector, acquisitor,centering, relaxer, database, evaluator, seed =
    ↵  1)
agox.run(N_iterations=100)

```

---

## C Density of States

In this appendix the computed density of state for all relevant geometries are shown.

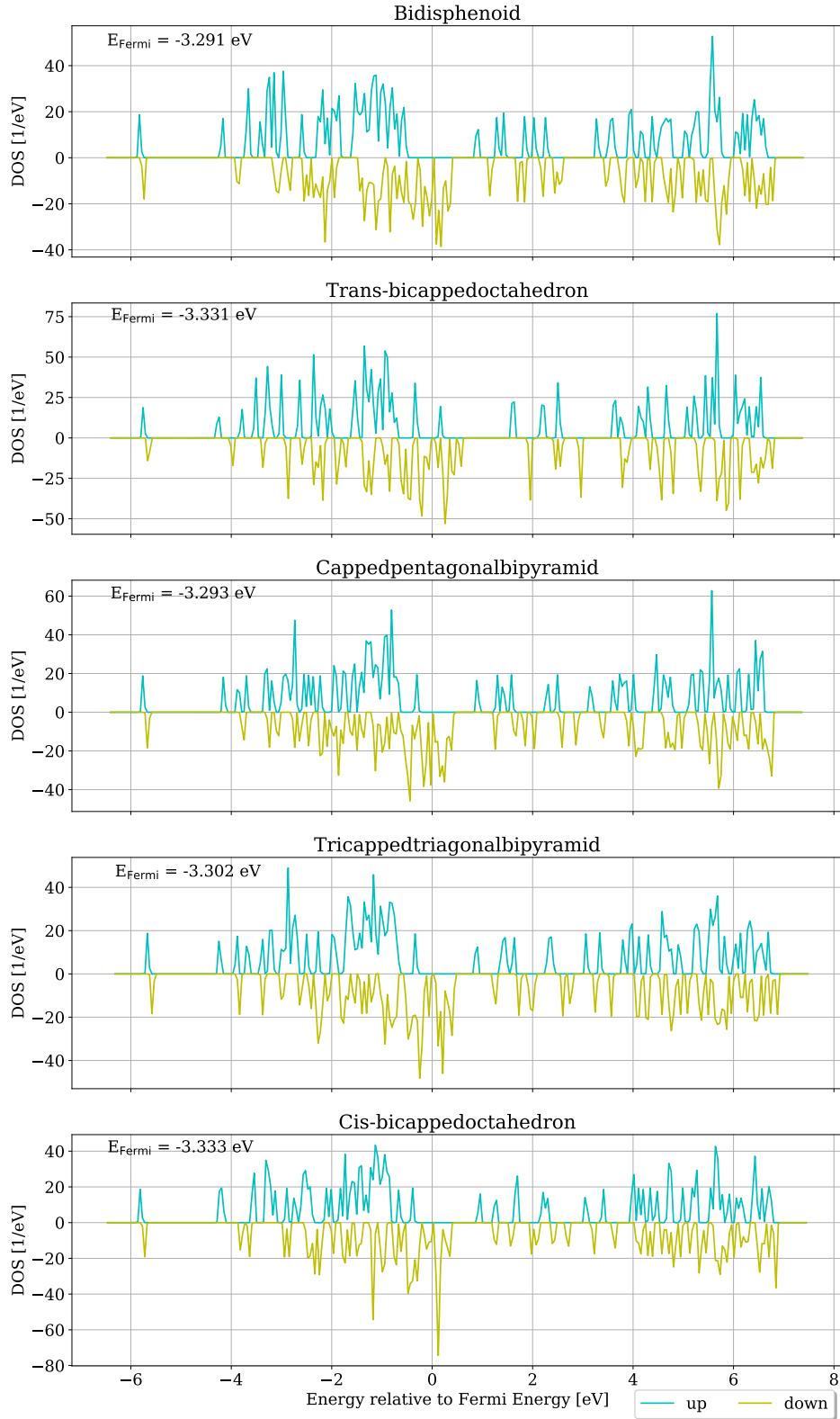


Figure 8: The density of state for the  $\text{Ni}_8$  geometries. The energy is relative to the Fermi energy.

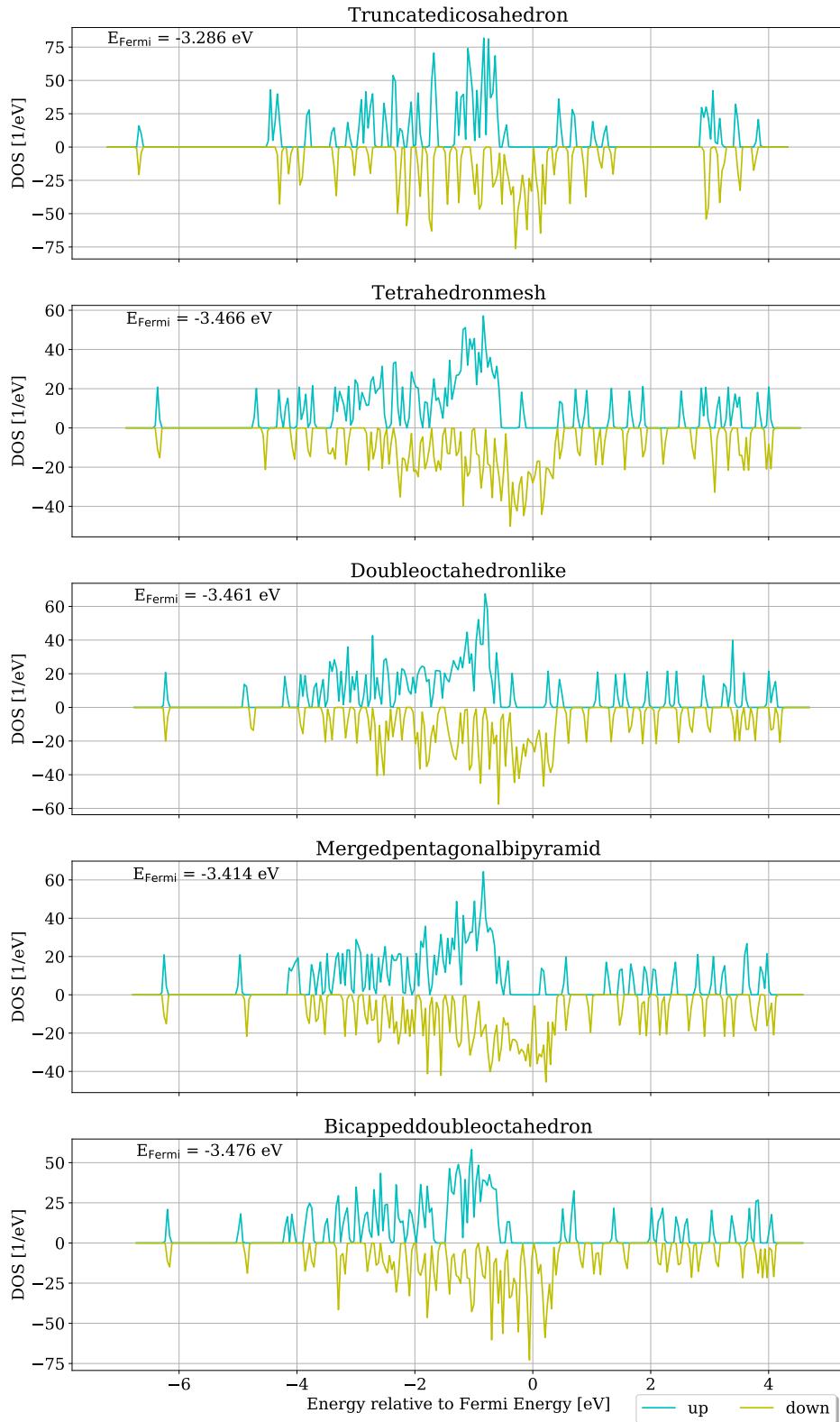


Figure 9: The density of state for the  $\text{Ni}_{12}$  geometries. The energy is relative to the Fermi energy.

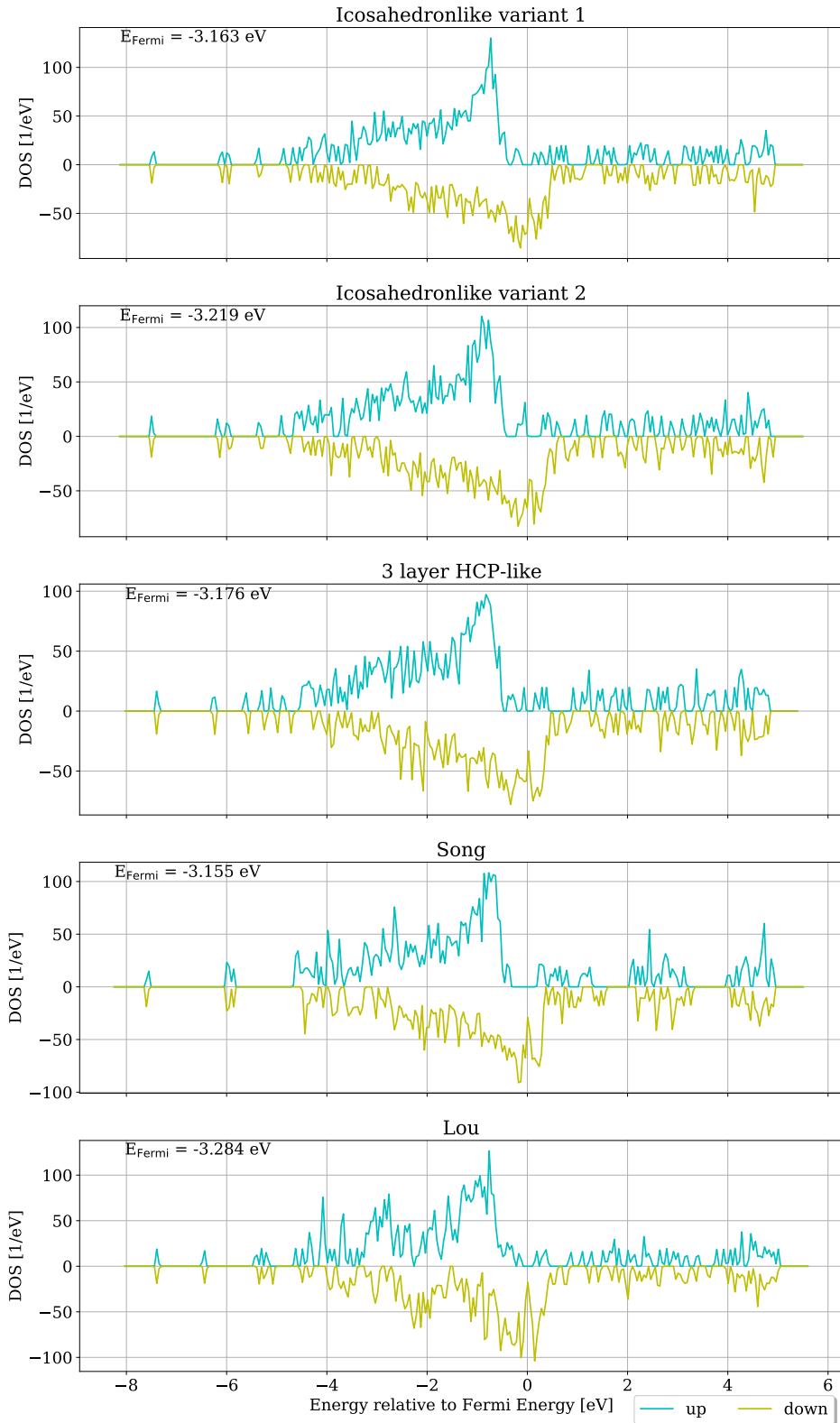


Figure 10: The density of state for the  $\text{Ni}_{30}$  geometries. The energy is relative to the Fermi energy.

---

## D Orbitals

In this appendix the charge distribution HOMO and LUMO orbitals for the structures investigated are shown.

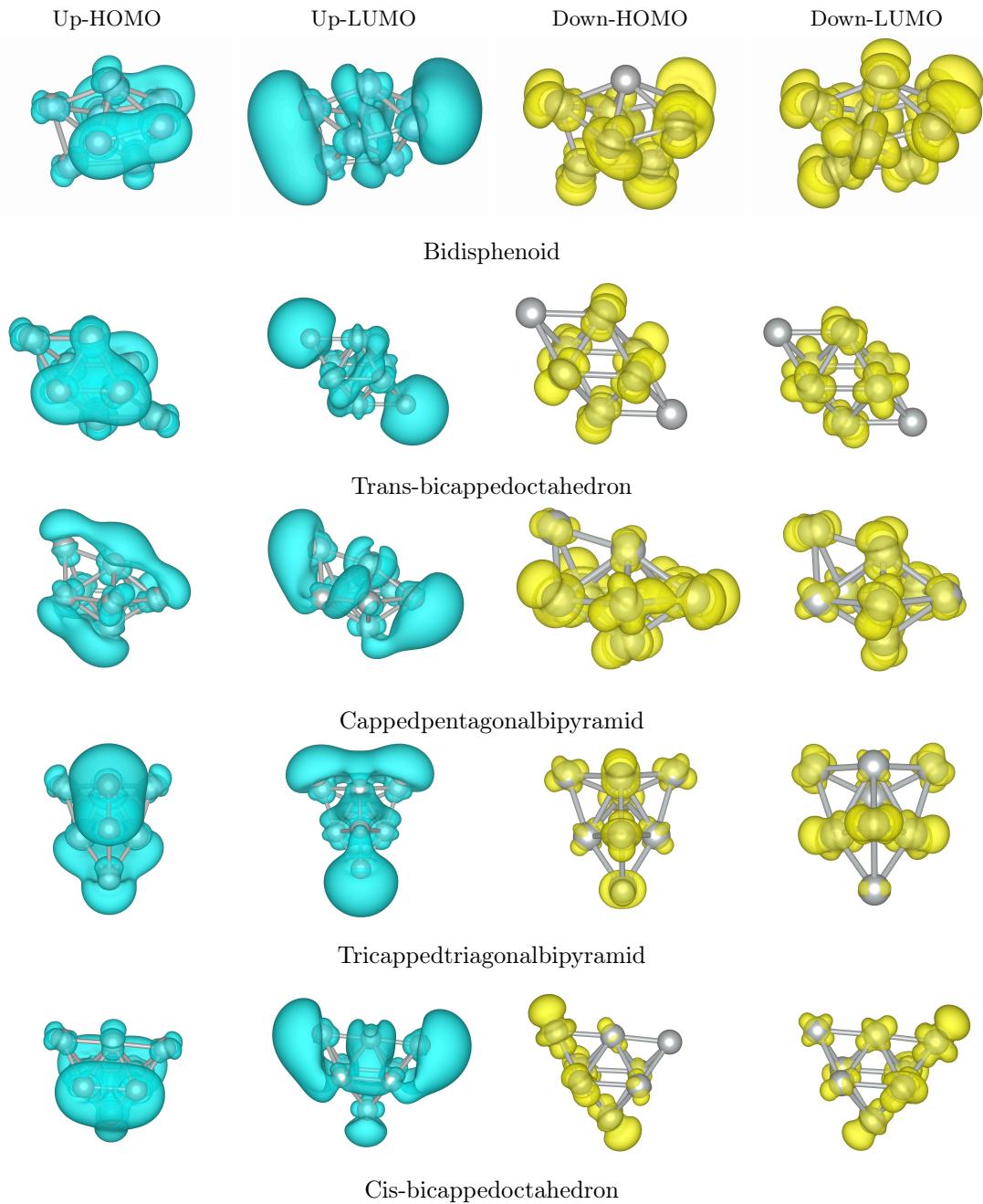


Figure 11: The different charge distributions for the HOMO and LUMO of the Ni<sub>8</sub> geometries. Each spin is included and the down spin HOMO and LUMO gap is the effective gap.

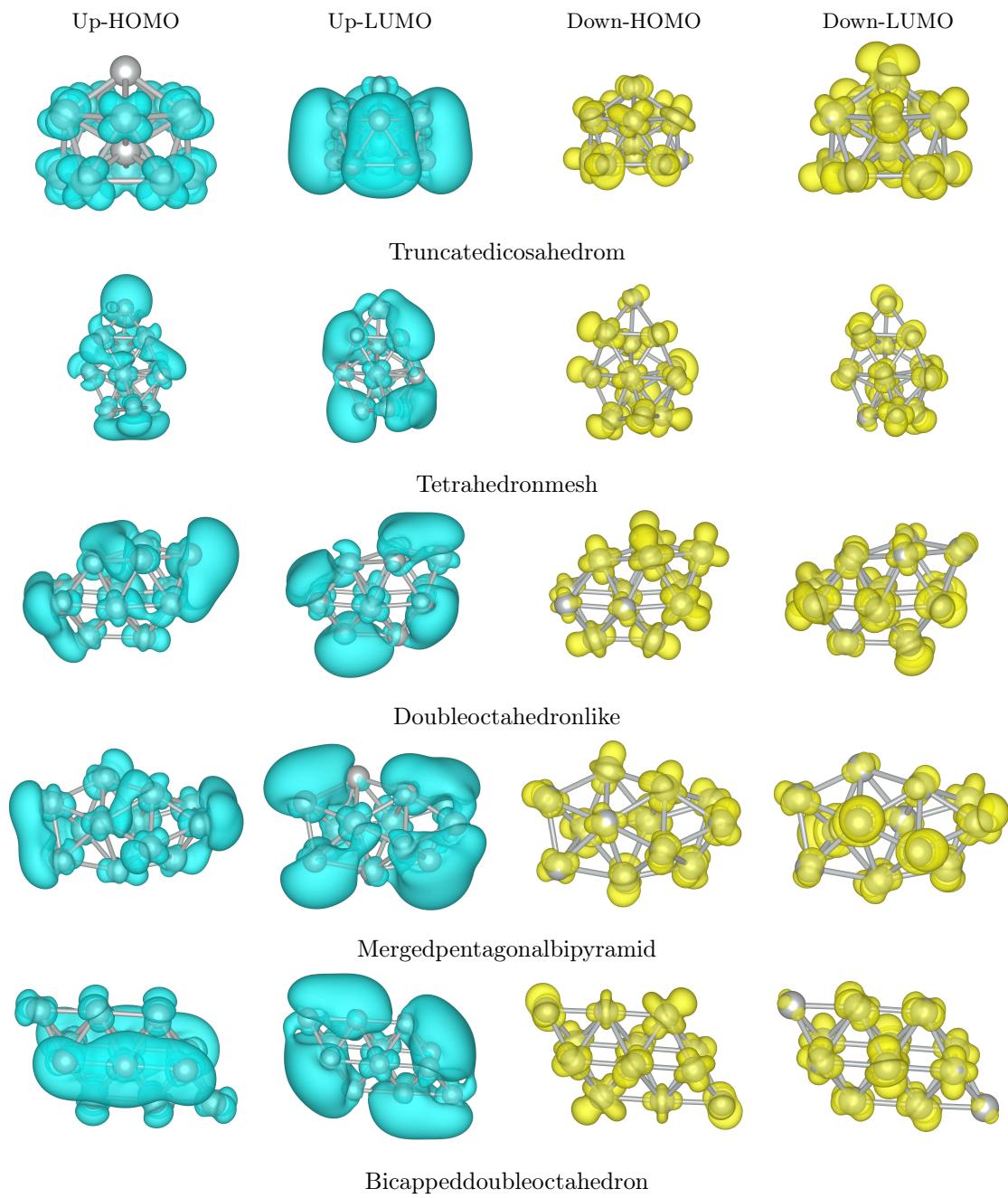


Figure 12: The different charge distributions for the HOMO and LUMO of the  $\text{Ni}_{12}$  geometries. Each spin is included and the down spin HOMO and LUMO gap is the effective gap.

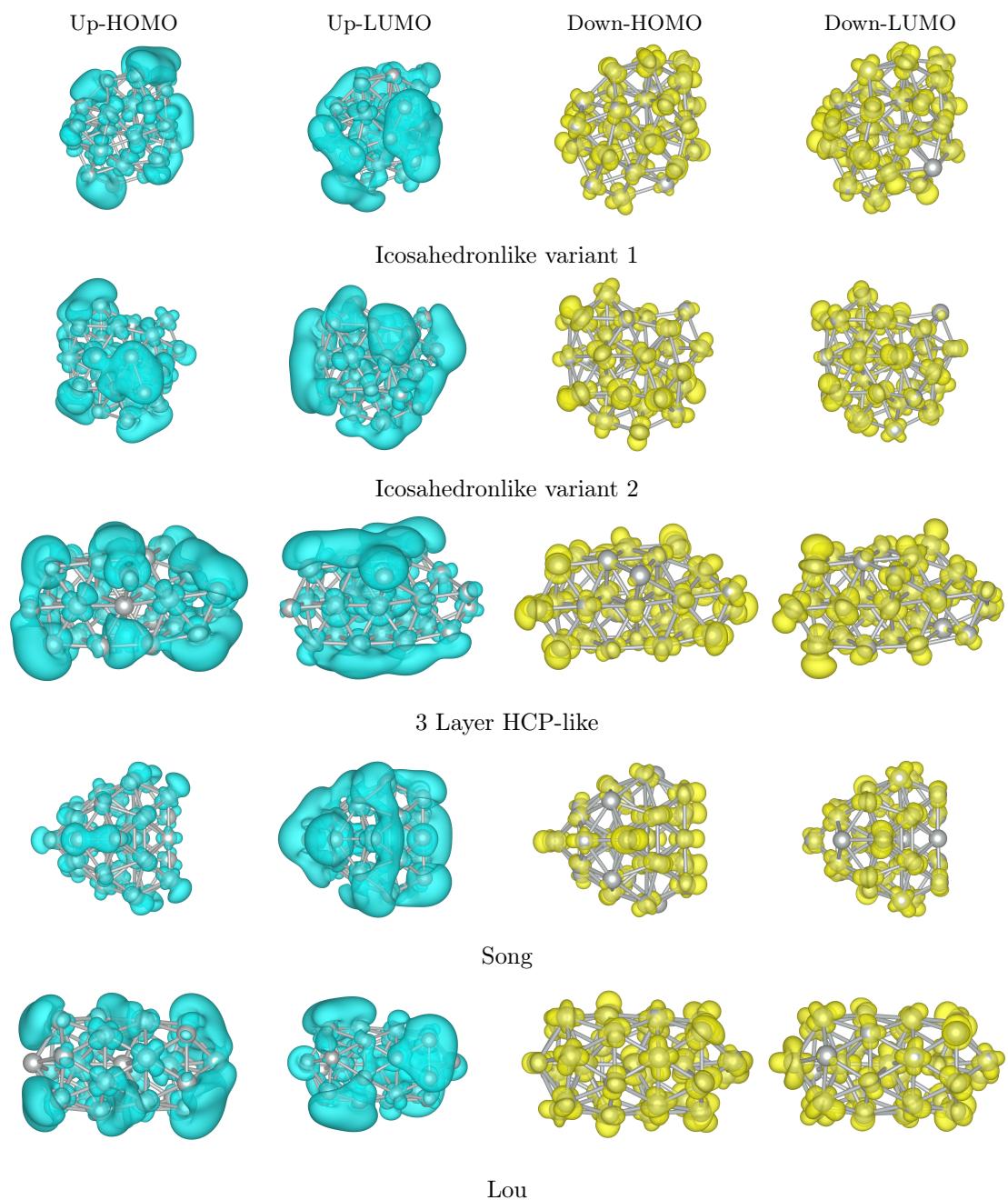
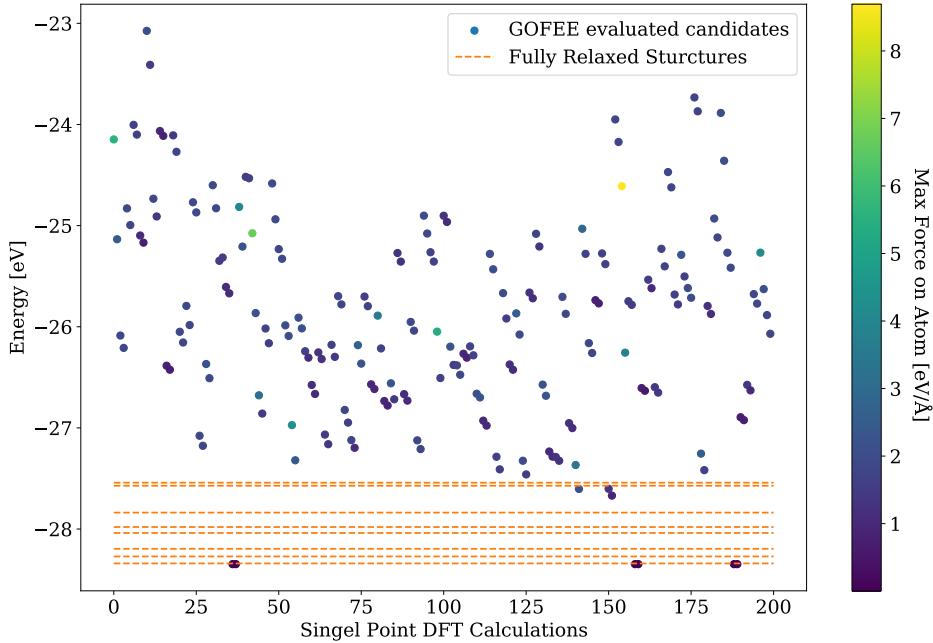


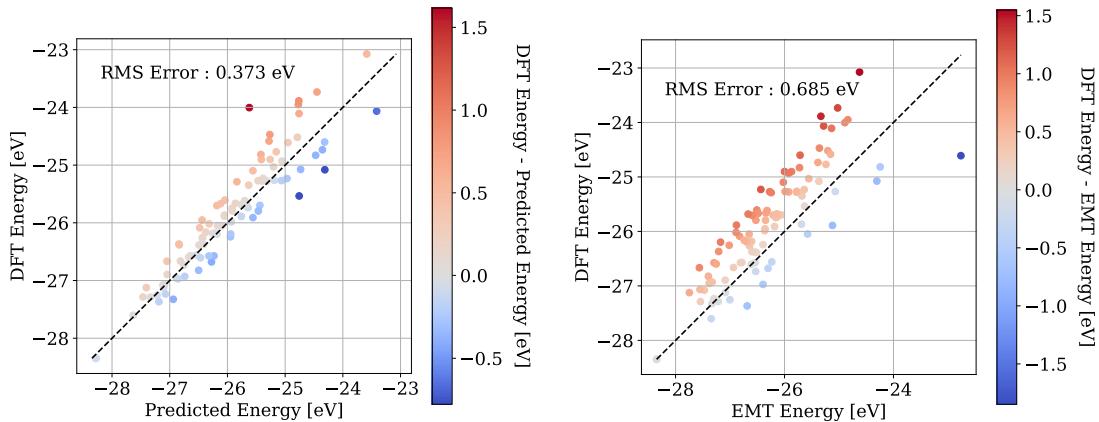
Figure 13: The different charge distributions for the HOMO and LUMO of the  $\text{Ni}_{30}$  geometries. Each spin is included and the down spin HOMO and LUMO gap is the effective gap.

## E Global Optimization with First-principles Energy Expressions Performance Showcase

In this section the performance of GOFEE is shown together with the comparisons of the surrogate potential, EMT and DFT energies.



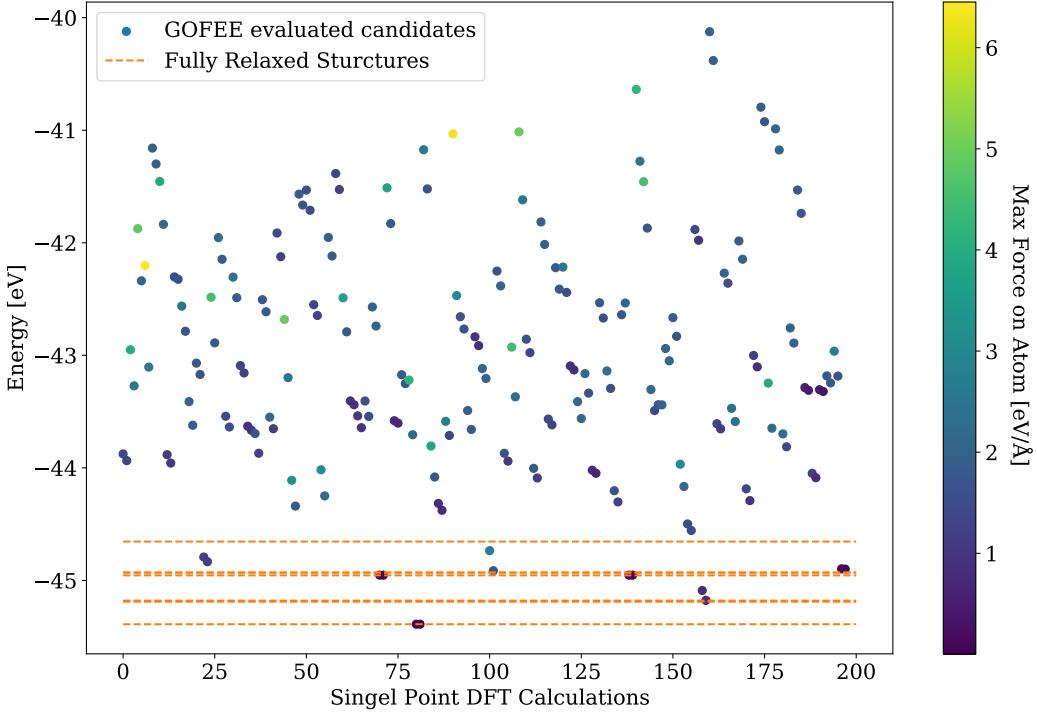
(a) The distribution of candidate (DFT) energies. The points are colored based on the largest force on an individual atom for that candidate. The dashed lines are the energies of the fully relaxed structures, that together with the relaxation trajectory, is used as training data for the surrogate potential.



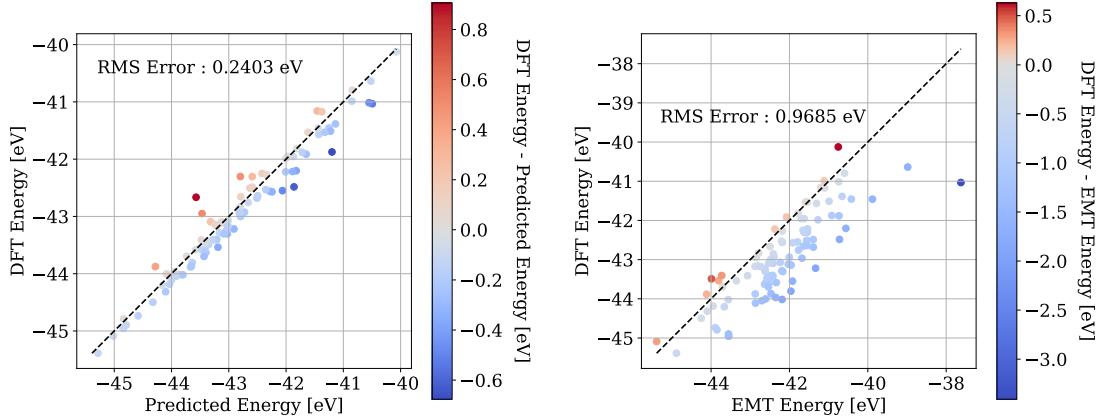
(b) The distribution of the predicted energies by the surrogate model and the DFT energies. The dashed line shows where the two energies are identical.

(c) The distribution of the EMT energies and the DFT energies. The dashed line shows where the two energies are identical. As the energy offset of the two theories are different they have been translated so the lowest EMT energy is the same as the lowest DFT energy.

Figure 14: The figures show the performance of the GOFEE-algorithm for the Ni<sub>8</sub> cluster with 100 iterations, 1 update step (2 DFT evaluations each iteration), 20 candidates each iteration and a kappa value of 2.



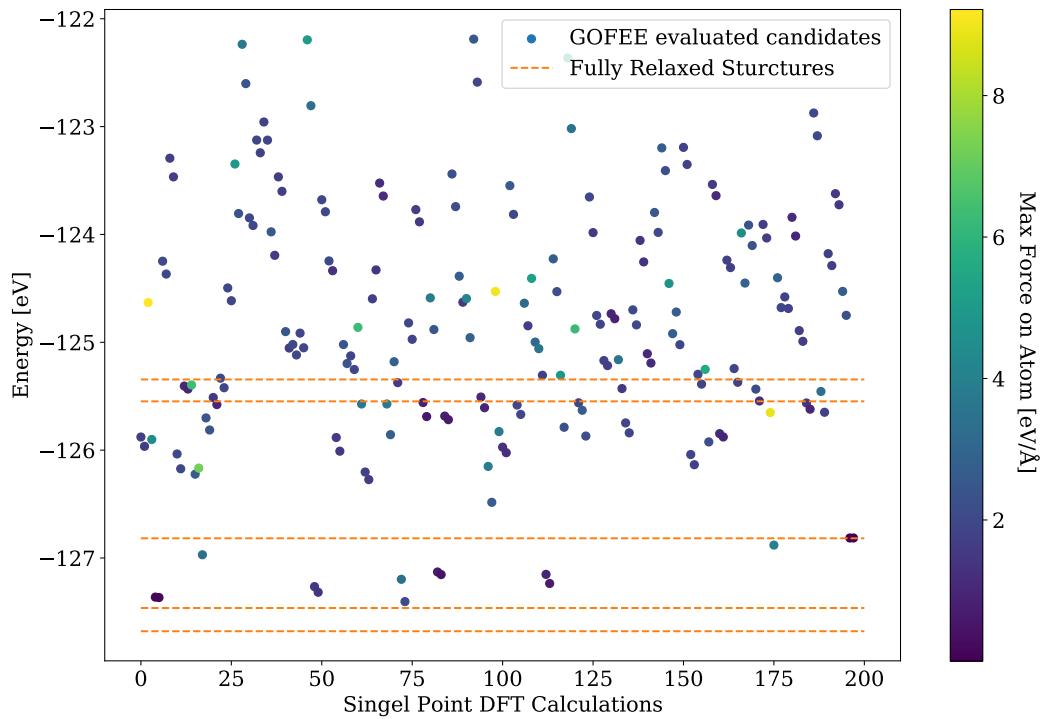
(a) The distribution of candidate (DFT) energies. The points are colored based on the largest force on an individual atom for that candidate. The dashed lines are the energies of the fully relaxed structures, that together with the relaxation trajectory, is used as training data for the surrogate potential.



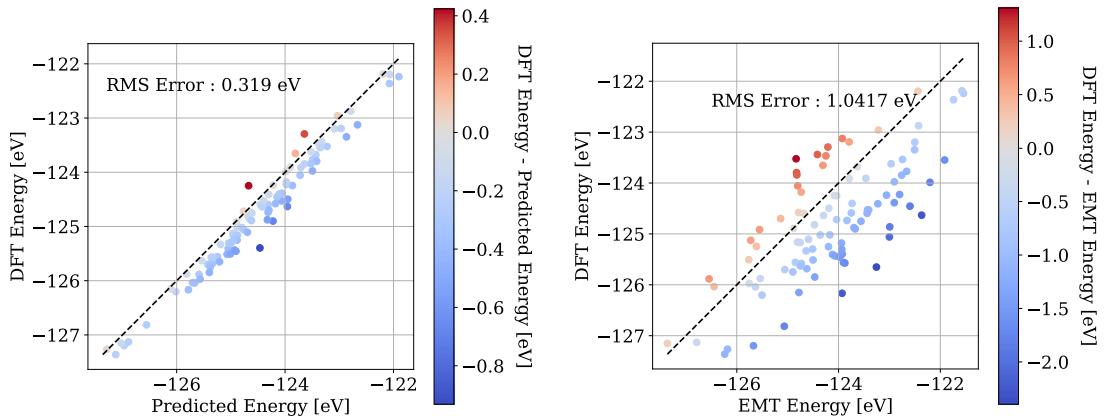
(b) The distribution of the predicted energies by the surroage model and the DFT energies. The dashed line shows where the two energies are identical.

(c) The distribution of the EMT energies and the DFT energies. The dashed line shows where the two energies are identical. As the energy offset of the two theories are different they have been translated so the lowest EMT energy is the same as the lowest DFT energy.

Figure 15: The figures show the performance of the GOFEE-algorithm for the  $\text{Ni}_{12}$  cluster with 100 iterations, 1 update step (2 DFT evaluations each iteration), 20 candidates each iteration and a kappa value of 2.



(a) The distribution of candidate (DFT) energies. The points are colored based on the largest force on an individual atom for that candidate. The dashed lines are the energies of the fully relaxed structures, that together with the relaxation trajectory, is used as training data for the surrogate potential.



(b) The distribution of the predicted energies by the surroage model and the DFT energies. The dashed line shows where the two energies are identical.

(c) The distribution of the EMT energies and the DFT energies. The dashed line shows where the two energies are identical. As the energy offset of the two theories are different they have been translated so the lowest EMT energy is the same as the lowest DFT energy.

Figure 16: The figures show the performance of the GOFEE-algorithm for the Ni<sub>30</sub> cluster with 100 iterations, 1 update step (2 DFT evaluations each iteration), 20 candidates each iteration and a kappa value of 2.