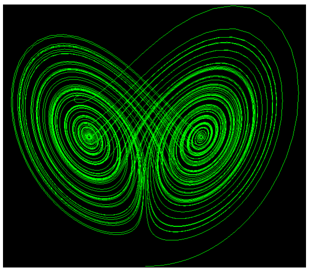
Departments of Mathematics, Physics and DACSO

MASTER THESIS

Modelling for Science and Engineering

Electronic Structure Calculations for Large Metallic Systems

Marc Eixarch Fernández



Setembre 2017

Universitat Autònoma de Barcelona

Master Thesis

Electronic structure simulations for large metallic systems

September 2017

Author: Marc Eixarch Fernández
NIU: 1212502
Supervisors: Stephan Mohr and Mervi Mantsinen

Fusion Group, BSC
Master on Modeling for Science and Engineering
Universitat Autònoma de Barcelona

Abstract

Motivated by its possible application in fusion materials research, we validate the linear scaling approach of density functional theory (DFT) for large metallic systems using the BigDFT code. After proving its validity and applicability in terms of accuracy and computational cost, a real application to fusion materials research is presented. The single vacancy formation energy of a Tungsten bcc lattice is analyzed and compared to the state-of-the-art results. In reference calculations reported in the literature, those values are obtained using the traditional cubic approach of DFT with a relatively large unit cell containing the vacancy and a small k-points grid. The use of the Monkhorst-Pack k-points grid allows to correctly take into account the periodicity of the cell and thus to perform simulations that are equivalent to many repeated copies of the unit cell. In this way one can simulate system sizes that lead to well converged results. However this is based upon one assumption: the effect introduced by the vacancy does not reach distances larger than half the unit cell length. The linear scaling approach has allowed us to perform Γ-only calculations with system sizes that are equivalent to those using a k-point grid. This has enabled us to analyze the size dependence of the defects in more detail and provided us with new insights regarding the validity of the aforementioned assumption.

Acknowledgements

This work has been carried out within the Fusion Group in the Barcelona Supercomputing Center. First of all I would like to thank the whole group for the marvelous opportunity they offered me and the wonderful treatment they gave me. In particular, I would like to thank Mervi Mantsinen for her constant attention and interest in my work.

Secondly, I would like to thank Stephan Mohr, not only for its excellent supervision of my work and constantly helping me when anything did not make sense, but also for the enjoyable conversations we have had over the past seven months. Up to now, there has not been a single time he has refused assisting me *immediately* after I had a doubt.

Finally, we acknowledge EUROfusion for providing us with the CPU-time that we have used to perform most of our simulations, in the Marconi-Fusion supercomputer in Italy. This work has been carried out within the framework of the EUROfusion Consortium and has received funding from the Euratom research and training programme 2014-2018 under grant agreement No 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

També a vosaltres, Guillem, Mercè, Jordi i, sobretot a tu, Elena.

Publications and presentations

This MSc thesis project resulted in the research results presented in the following contributions:

- Invited talk entitled ”Linear scaling DFT simulations for large systems – application to vacancy defects in W” at the Integrated Radiation Effects Modelling and Experimental Validation (IREMEV) Meeting of the EUROfusion Work Package Materials (WPMAT), in June 2017 in the Institutu Ruđer Bošković, Zagreb (Croatia).
- Contributed talk entitled ”Accurate and universally applicable DFT calculations for large metallic systems – Application to defects” at EUROMAT 2017, European Congress and Exhibition on Advanced Materials and Processes. To be presented in September 2017 in the Thessaloniki Concert Hall, Thessaloniki (Greece).
- Contributed talk entitled ”Accurate and universally applicable DFT calculations for large metallic systems – Application to defects” at ICFRM-18, International Conference on Fusion Reactor Material. To be presented in November 2017 in the Link Station Hall, Aomori (Japan).
- Original journal paper in preparation.

The MSc candidate carried out most of the simulations presented in this MSc thesis, analyzed the simulation results and had the full responsibility of writing it.

Contents

1 Introduction 3

2 Electronic Structure Calculations 7

2.1 Basics on Density Functional Theory 8

2.1.1 The many-body problem and the Born-Oppenheimer approximation 8

2.1.2 The Hohenberg-Kohn Theorems 9

2.1.3 The Kohn-Sham formalism and the exchange-correlation functional 13

2.1.4 Numerical and computational techniques 15

2.2 Linear scaling approach 16

2.3 Real *vs* reciprocal space simulations 19

2.3.1 Interests and challenges regarding metal calculations 22

3 Implementation of DFT in the BigDFT code 25

3.1 Wavelets 25

3.2 Specific linear scaling implementation 31

4 Validation of the linear version of BigDFT for large metallic systems 37

4.1 Minimal cell study and parameter optimization 39

4.1.1 Grid optimization 40

4.1.2 Lattice parameter optimization 42

4.1.3 k-points convergence 43

4.2 Towards large systems with the linear scaling approach 47

4.2.1 Tungsten calculations up to 2000 atoms 48

4.2.2 Comparison with the cubic approach 51

4.2.3 Computing performance and scalability 53

5 An application: Tungsten Vacancy Formation Energy 57
5.1 Point defects 58
5.2 Simulations and analysis of results 60
5.2.1 Multidefects analysis 65
5.2.2 Charge analysis 67

6 Conclusions 73

7 Bibliography 77

1 | Introduction

The rapid increase of the energy demand over the world has become an urgent and critical issue for humanity. For different causes, the current energy extraction methodologies are not real candidates to fulfill the future needs. This is the main reason why over the last decades the scientific community has been developing big projects on fusion power reactors. Fusion energy has lots of advantages and benefits over the other approaches, providing a *sustainable, secure* and *safe* solution. However, because of the huge complexity and extreme conditions required to achieve a controlled fusion reaction, no design has yet achieved positive net energy gain. The study of suitable materials to be used on the reactor walls is one of the challenges that still needs to be overcome.

In fusion reactors, high-energy neutrons are produced in nuclear reactions between fusion fuel particles. When colliding with the material wall of the reactor, these neutrons can cause damage, leading to, for example, transmutation of atoms or creation of radiation *defects.* Understanding of such defects both qualitatively and quantitatively is a key point of fusion materials research, since even very small defects could evolve over time into big clusters of defects and eventually change the microstructure of the material. At this point, the lifetime of the material could be dramatically reduced or it could behave unexpectedly.

Unfortunately, the lack of a suitable high-flux source of high-energy neutrons ($\sim$ 14.1 MeV) required to reproduce the conditions found inside a reactor vessel makes it impossible to experimentally test with different materials. Furthermore, long-term effects would only be appreciable after many years of experimental testing. Therefore, it is needed to use various mathematical models and implement them in computer codes to simulate those situations at various temporal and spatial scales. The analysis of the obtained results may give to the community the required insight in order to properly predict which materials would be more appropriate to form the reactor wall and to foretell its degradation rate.

The study of such fusion (and fission) materials started in the early 1970's and many different approaches have been studied until now. A particular theoretical model, known as Density Functional Theory (DFT) [1], dominates nowadays at small scales (from tens to some hundred atoms) thanks to its good balance between accuracy, generality and scalability with respect to the system size. In particular, DFT-based simulation codes show a very good performance not only for small finite systems but also for large and purely periodic ones. The good performance in the latter case is thanks to a special and widely used technique known as k-point sampling, which takes advantage of the periodicity and only considers a *minimal cell*.

Unfortunately, three main issues make fusion material simulations hard and computationally expensive. Firstly, as the system size or the minimal cell size for periodic systems grow, DFT simulations become unfeasible even running on supercomputers, since it scales to the power of 3 with the system size. Secondly, metallic systems, which are the most promising candidates for the fusion reactor walls, represent further difficulty in DFT simulations due to their zero HUMO-LUMO gap [2]. Finally, the fact that a radiation defect breaks the periodicity of a perfect lattice hinders the use of the k-points method and demands the use of big cells. These reasons impose a maximum size for DFT simulations on fusion material systems of around some hundred atoms and force scientists to make extra assumptions or approximations in order to overcome this restriction.

In the last few years DFT-based simulation codes that exhibit linear scaling with respect to the system size have been developed. They involve an additional assumption in the theoretical background. Known as *nearsightedness* [3], this approximation basically imposes a cutoff radius for each particle, meaning that it only "sees" information from particles closer than this radius. Far from being questionable, this assumption has been theoretically proven to be well justified. While the convergence and applicability of this linear scaling approach have already been shown with non-metallic system, no code has yet proven to be reliable and accurate for metallic ones.

The first objective of this thesis is to validate the linear scaling DFT approach for metallic systems up to a few thousand atoms using the BigDFT code [4], which presents an implementation of an universally applicable linear scaling algorithm. To do so, a particular metallic crystal structure, namely a perfect Tungsten bcc lattice, will be tested for different sizes with both cubic and linear approaches. This metal has been chosen mainly due to its relevance in the fusion material research field. As an example, its good behavior under radiation has lead to its future usage in the material wall of the ITER nuclear fusion reactor [5]. Although other materials have been also tested, this thesis focuses on the Tungsten simulations. From the aforementioned simulations, a set of material properties will be derived and compared both qualitatively and quantitatively.

Once the viability of the linear scaling approach of DFT for the Tungsten system is proven, the second objective is addressed: perform electronic structure calculations with interest to the fusion community with the linear scaling version of BigDFT for Tungsten. For this purpose we have chosen the monovacancy point defect in a Tungsten lattice as a target structure. This phenomena may occur when a high-energy neutron collides with the material wall of the fusion device and results in the possible migration of an atom from its usual place, thus leaving a vacancy. The key quantity associated to the vacancy defect is the so-called vacancy formation energy (VFE) and it is calculated using the total energy of the perfect and defect structures. In order to complete the study, we have also investigated another point defect, namely the self interstitial atom (SIA). Unfortunately and mainly due to CPU-time issues, we have not been able to gather enough SIA simulation results to include its analysis in this thesis. We expect to obtain these results in the upcoming months and to add them to the publication in preparation.

Even though accurate values of VFEs for a wide range of materials were already published during the last decade using DFT codes (see for example [6]), we want to re-investigate this topic with our linear scaling approach to get additional and valuable insights regarding the points defects. Particularly, the standard method used to compute defect formation energies makes an assumption: the interaction range of a single vacancy does not reach distances larger than about 7 - 8 Å. Using the linear version of BigDFT we are able to reach system sizes that are larger than those used in the literature and do thus not need to make the latter assumption. This introduces a new perspective on the calculations of point defects using DFT. We find that in the Γ-only approach the Tungsten VFE does not converge for 2000 atoms, which clearly indicates that larger cells are needed. In the k-points approach, we find that a unit cell consisting only on 128 atoms with a single vacancy and a k-points grid may introduce a small interaction between neighboring vacancies which would lead to a not so accurate value of the VFE. The latter needs to be further studied with more simulation data.

The thesis is structured as follows: in chapter 2 a review of Electronic Structure Calculations will be given, emphasizing the theoretical background of DFT and the extra assumption and development of the linear scaling approach. In chapter 3 the specific implementation of DFT into the BigDFT code will be explained, along with some technicalities. In chapter 4 we present the simulations performed for the validation of the linear scaling version of BigDFT for large metallic systems, along with the analysis of the results in terms of accuracy and computational cost. In chapter 5 we study the vacancy point defect in a Tungsten lattice. In particular, we focus on its key quantity, the formation energy. Finally, in chapter 6, we present the conclusions of this work.

2 Electronic Structure Calculations

Since the first qualitative calculations of electronic structures in the 1930's (for example, by Fock [7]), Electronic Structure Calculations (ESC) have been constantly improving in terms of complexity, accuracy and speed. The main goal of ESC is to obtain the exact electronic configuration (the electron wave functions) of any system of atoms and/or molecules by means of quantum mechanics. Once this is achieved, any other magnitude related to the system can be derived "easily", such as its energy, band structure or thermal conductivity among many others. Although during the first years ESCs were mainly theoretical, nowadays its applications are plentiful, ranging from structural biology to fusion research.

There are plenty of different methods and approximations in order to perform ESC. In this thesis we will work within the Density Functional Theory (DFT) framework. DFT is becoming nowadays more and more popular mainly because of its balance between accuracy, generality and scalability with respect to the system size. Its use is focused on systems ranging from typically tens to hundreds of atoms. It is an *ab initio* method, meaning that it uses fundamental principles of quantum theory in order to perform the calculations. While providing very accurate results, the complexity and cost of the calculations are generally much higher than those for non-*ab initio* methods.

In this chapter we start, in section 2.1, by giving a short review of the basics and more relevant approximations of DFT. Secondly, in section 2.2, we present the theoretical background of the aforementioned linear scaling approach for DFT, developed in the recent years and needed to reach system sizes that a few years ago were unfeasible. Next, in section 2.3, we discuss about the differences and characteristics between real and reciprocal space simulations. Finally, in section

2.3.1, the extra challenges that metallic systems present will be covered and will serve as starting points for chapters 3 and 4. For the discussion in section 2.1, reference [1] has been very useful and extensively consulted, while for the discussion of the linear scaling in section 2.2 reference [8] has been also widely used.

2.1 Basics on Density Functional Theory

2.1.1 The many-body problem and the Born-Oppenheimer approximation

The electronic structure problem is a many-body system consisting of electrons and nuclei which needs to be solved by quantum mechanics. Considering that no external potential is present and that relativistic, magnetic and electrodynamics effects are not considered, the Hamiltonian of the system is written as:

$$\begin{aligned}\hat{H} = &-\frac{\hbar^2}{2m_e}\sum_j \nabla_j^2 - \sum_{j,J}\frac{Z_J e^2}{|\mathbf{r}_j - \mathbf{R}_J|} + \frac{1}{2}\sum_{j\neq k}\frac{e^2}{|\mathbf{r}_j - \mathbf{r}_k|} \\ &- \sum_J \frac{\hbar^2}{2M_J}\nabla_J^2 + \frac{1}{2}\sum_{J\neq K}\frac{Z_J Z_K e^2}{|\mathbf{R}_J - \mathbf{R}_K|},\end{aligned} \tag{2.1}$$

where lower case subscripts refer to electrons and upper case subscripts to nuclei, Z_I is the charge of nuclei I and m_e and e are the electron's mass and charge, respectively. The first and fourth terms account for the kinetic energy of the electrons and nuclei, respectively, while the rest of terms account for electric repulsion or attraction among electrons and nuclei themselves or between them.

By solving the eigenstates and eigenvalues problem

$$\hat{H}\left|\Psi\right\rangle = E\left|\Psi\right\rangle, \tag{2.2}$$

one may obtain the exact solutions of the problem, consisting in a combined electron-nuclei wave function $\Psi(\{\mathbf{r}_i\}, \{\mathbf{R}_I\})$. Equation (2.2) is the so-called *time-independent Schrödinger equation.*

Unfortunately, this problem can be solved analytically only for the simplest case (hydrogen-like systems) and numerically for slightly more complex ones. Therefore, the search and derivation of useful and valid approximation is the first step in ESC.

The first and standard approximation is the well known *Born-Oppenheimer* or *adiabatic* approximation. Shortly, it allows to treat the position of the nuclei as fixed parameters and therefore solve the time-independent Schrödinger equation only for the electrons, reducing significantly the number of variables and operators. It can be easily justified by expanding the general electron-nuclei wave function in an orthonormal basis (see [1] or [8]) and using the fact that the contribution of the term with M_J (with much higher value than any other constant) in the Hamiltonian (2.1) is negligible. Taking into account this approximation and using the Hartree atomic units $\hbar = m_e = e = 4\pi/\epsilon_0 = 1$, the Hamiltonian (2.1) can be rewritten as:

$$\hat{H} = \hat{T} + \hat{V}_{\text{ext}} + \hat{V}_{\text{int}} + E_{JJ}, \tag{2.3}$$

with

$$\begin{aligned} \hat{T} &= -\frac{1}{2}\sum_j \nabla_j^2, \\ \hat{V}_{\text{ext}} &= \sum_{j,J} \frac{Z_J}{|\mathbf{r}_j - \mathbf{r}_J|}, \\ \hat{V}_{\text{int}} &= \frac{1}{2}\sum_{j\neq k} \frac{1}{|\mathbf{r}_j - \mathbf{r}_k|} \end{aligned} \tag{2.4}$$

and E_{JJ} contains any term contributing to the total energy of the system other than electronic interaction, such as the interaction between nuclei.

2.1.2 The Hohenberg-Kohn Theorems

The latter expressions are used not only in DFT, but also in other ESC methods. From now on, particular DFT approximations are presented, starting with the *Hohenberg-Kohn* theorems [9], which form the basis of DFT. These widely-known theorems were proved in 1964 by Hohenberg and Kohn and their usefulness lies on their simplicity and ingeniousness. Basically, it changes completely the perspective of the many-body problem by giving an analogous way to solve it using the *particle* *density.* Thus, the H-K theorems form the cornerstone of all the DFT theory. Before the statement and proof of both theorems we introduce formally the particle density $n(\mathbf{r})$, which plays a key role in DFT. It is defined as the expectation value of the density operator $\hat{n}(\mathbf{r}) = \sum_{i=1,\ldots,N} \delta(\mathbf{r} - \mathbf{r}_i)$, where N is the total number of particles in the system. $n(\mathbf{r})$ is expressed (both in Dirac's and integral notation) as

$$n(\mathbf{r}) = \frac{\langle \Psi | \hat{n}(\mathbf{r}) | \Psi \rangle}{\langle \Psi | \Psi \rangle} = N \frac{\int \mathrm{d}^3 r_2 \ldots \mathrm{d}^3 r_N |\Psi(\mathbf{r}, \mathbf{r}_2, \ldots, \mathbf{r}_N)|^2}{\int \mathrm{d}^3 r_1 \ldots \mathrm{d}^3 r_N |\Psi(\mathbf{r}_1, \mathbf{r}_2 \ldots, \mathbf{r}_N)|^2}. \tag{2.5}$$

We notice that spin has been omitted and the integral notation is obtained using the fact that wave functions are symmetric in all the electron coordinates. Also, the particle density of a particular state $n_j(\mathbf{r})$ is obtained by substituting Ψ by its particular wave function Ψ_j. In a similar way, the total energy of the system is defined as the expectation value of the Hamiltonian,

$$E = \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \equiv \langle \hat{H} \rangle = \langle \hat{T} \rangle + \langle \hat{V}_{\text{int}} \rangle + \int d^3 r V_{\text{ext}}(\mathbf{r}) n(\mathbf{r}) \; + E_{JJ}, \tag{2.6}$$

where the expectation value of the external potential has been explicitly written. With the latter definitions in mind, the *Hohenberg-Kohn* theorems are presented, which are quoted from reference [1]:

Theorem 2.1. *For any system of interacting particles in an external potential* $V_{ext}(\mathbf{r})$*, the potential* $V_{ext}(\mathbf{r})$ *is determined uniquely, except for a constant, by the ground state particle density* $n_0(\mathbf{r})$*.*

Proof. This proof works by *reductio ad absurdum.* Let us assume that there exist two different external potentials, namely $V_{\text{ext}}^{(1)}(\mathbf{r})$ and $V_{\text{ext}}^{(2)}(\mathbf{r})$, which differ by more than a constant and which lead to the same ground state density $n(\mathbf{r})$. First, using the assumption and equation (2.3), one notices that the two external potentials lead to two different Hamiltonians, $\hat{H}^{(1)}$ and $\hat{H}^{(2)}$, which at the same time have different ground state wave functions, $\Psi^{(1)}$ and $\Psi^{(2)}$ (we omit the subscript 0 for convenience). Since $\Psi^{(2)}$ is not the ground state of $\hat{H}^{(1)}$ (also supposing the ground states are non-degenerate and states are normalized), it follows that

$$E^{(1)} = \left\langle \Psi^{(1)} \middle| \hat{H}^{(1)} \middle| \Psi^{(1)} \right\rangle < \left\langle \Psi^{(2)} \middle| \hat{H}^{(1)} \middle| \Psi^{(2)} \right\rangle. \tag{2.7}$$

Rewriting the last term in the later inequality and using (2.6):

$$\begin{aligned}\langle \Psi^{(2)}|\hat{H}^{(1)}|\Psi^{(2)}\rangle &= \langle \Psi^{(2)}|\hat{H}^{(2)}|\Psi^{(2)}\rangle + \langle \Psi^{(2)}|\hat{H}^{(1)} - \hat{H}^{(2)}|\Psi^{(2)}\rangle \\ &= E^{(2)} + \int \mathrm{d}^3r \left[V_{\text{ext}}^{(1)}(\mathbf{r}) - V_{\text{ext}}^{(2)}(\mathbf{r})\right] n_0(\mathbf{r}).\end{aligned} \tag{2.8}$$

This leads to

$$E^{(1)} < E^{(2)} + \int \mathrm{d}^3r \left[V_{\text{ext}}^{(1)}(\mathbf{r}) - V_{\text{ext}}^{(2)}(\mathbf{r})\right] n_0(\mathbf{r}). \tag{2.9}$$

We notice that exactly the same procedure exchanging superscript 1 by 2 leads to the same expression taking into account the substitution,

$$E^{(2)} < E^{(1)} + \int \mathrm{d}^3r \left[V_{\text{ext}}^{(2)}(\mathbf{r}) - V_{\text{ext}}^{(1)}(\mathbf{r})\right] n_0(\mathbf{r}). \tag{2.10}$$

Combining (2.9) and (2.10) leads to the contradiction $E^{(1)} + E^{(2)} < E^{(1)} + E^{(2)}$, proving that indeed the density uniquely determines the external potential. □

Corollary 2.1. *All properties of the system described in theorem 2.1 are completely determined given only the ground state density* $n_0(\mathbf{r})$.

Proof. By theorem 2.1, the external potential is already determined. From equation (2.3) it follows that therefore the full Hamiltonian is also completely determined. Finally, the system is completely characterized by the many-body wave function, which is obtained by means of the Hamiltonian. □

These results are theoretically remarkable, since it gives a completely different point of view of the problem. However, it is of little use in terms of applicability, since the many-body wave function would still need to be found. The second theorem tackles this problem.

Theorem 2.2. *A universal functional for the energy* $E[n]$ *in terms of the density* $n(\mathbf{r})$ *can be defined, valid for any external potential. Given a particular external potential, the exact ground state energy of the system is the global minimum value of this functional and the density* $n(\mathbf{r})$ *that minimizes the functional is the exact ground state density* $n_0(\mathbf{r})$.

Proof. This proof is restricted to densities that are *V-representable*, although the proof can be extended to a larger set of densities. For a detailed discussion see [1]. Since all properties of the system are uniquely determined if $n(\mathbf{r})$ is specified, then each of them can be expressed as a functional of $n(\mathbf{r})$, in particular the total energy functional:

$$\begin{aligned} E_{HK}[n] &= T[n] + E_{\text{int}}[n] + \int \mathrm{d}^3r \, V_{\text{ext}}(\mathbf{r})n(\mathbf{r}) \; + E_{JJ} \\ &\equiv F_{HK}[n] + \int \mathrm{d}^3r \, V_{\text{ext}}(\mathbf{r})n(\mathbf{r}) \; + E_{JJ}. \end{aligned} \tag{2.11}$$

Let us consider a system with a ground state density $n^{(1)}(\mathbf{r})$, corresponding to an external potential $V^{(1)}_{\text{ext}}(\mathbf{r})$. Then, the functional (2.11) evaluated at $n^{(1)}(\mathbf{r})$ is equal to the expectation value of the Hamiltonian in the ground state wave function, $\Psi^{(1)}$,

$$E^{(1)} = E_{HK}[n^{(1)}] = \left\langle \Psi^{(1)} \middle| \hat{H}^{(1)} \middle| \Psi^{(1)} \right\rangle . \tag{2.12}$$

Let us now consider a different density $n^{(2)}(\mathbf{r})$, which corresponds to a different wave function $\Psi^{(2)}$. By evaluating its energy as in (2.6) we obtain

$$E^{(1)} = \left\langle \Psi^{(1)} \middle| \hat{H}^{(1)} \middle| \Psi^{(1)} \right\rangle < \left\langle \Psi^{(2)} \middle| \hat{H}^{(1)} \middle| \Psi^{(2)} \right\rangle = E^{(2)} \tag{2.13}$$

Since $n^{(2)}$ has been arbitrarily chosen (only imposing that is different than $n^{(1)}$), then it follows that the functional takes its minimum when evaluated at n_0. ☐

Corollary 2.2. *The functional $E[n]$ alone is sufficient to determine the exact ground state energy and density.*

Proof. Assume that $E[n]$ is known. Then by minimizing the total energy of the system (2.11) with respect to variations in the density function $n(\mathbf{r})$ one would find the exact ground state density of the system. ☐

Remark 1. One may notice that this constructions give you information about the ground state. Further steps should be performed to find excited states' properties.

Summarizing, the two H-K theorems present a new approach for solving the electronic structure problem by means of the *particle density* $n(\mathbf{r})$ and offer a method to solve the problem without the need of solving the many-body equation. However, one may see that the problem lies now on finding the *universal* functional $F_{HK}[n]$ and, unfortunately, its exact form is not known.

2.1.3 The Kohn-Sham formalism and the exchange-correlation functional

This issue was partially solved by Kohn and Sham one year after the theorems were stated, by introducing the so-called Kohn-Sham formalism [10]. Nowadays this approach is the more extended method within the DFT framework, because of its simplicity and great progress. The main idea of the formalism is to convert the problem of interacting particles to a system of non-interacting particles with the same ground state density. This *ansatz* also includes the effect of combining all the difficult many-body terms in an *exchange-correlation* functional of the density $E_{\text{xc}}[n]$. As a consequence the many-body electron wave function would be split into N single electron independent wave functions. The system of non-interacting particles is known as the Kohn-Sham auxiliary system and it rests upon two assumptions [1]:

1. The exact ground state density can be represented by the ground state density of an auxiliary system of non-interacting particles. This is called *non-interacting-V-representability.*
2. The auxiliary Hamiltonian is chosen to have the usual kinetic operator and an effective local potential $V_{\text{eff}}(\mathbf{r})$ acting on an electron at point $\mathbf{r}$.

These assumptions lead to the following Hamiltonian for the auxiliary system:

$$\hat{H}_{\text{aux}} = -\frac{1}{2}\nabla^2 + V(\mathbf{r}). \tag{2.14}$$

Furthermore, the fact that now the wave functions are independent, allows us to rewrite some quantities in terms of the orbital eigenfunctions $\psi_i(\mathbf{r})$,

$$\begin{gathered} n(\mathbf{r}) = \sum_{i=1}^{N} |\psi_i(\mathbf{r})|^2, \\ T = \frac{1}{2}\sum_{i=1}^{N} \int \mathrm{d}^3r |\nabla\psi_i(\mathbf{r})|^2. \end{gathered} \tag{2.15}$$

We also define the classical Coulomb interaction energy of the electron density interacting with itself (the so-called *Hartree energy*) as

$$E_{\text{Hartree}}[n] = \frac{1}{2}\int \mathrm{d}^3r\mathrm{d}^3r' \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}. \tag{2.16}$$

Finally, the Kohn-Sham expression for the energy is

$$E_{\text{KS}} = T[n] + \int \mathrm{d}\mathbf{r} V_{\text{ext}}(\mathbf{r})n(\mathbf{r}) + E_{\text{Hartree}}[n] + E_{JJ} + E_{\text{xc}}[n]. \tag{2.17}$$

Now the kinetic operator is particle-independent and the internal interactions among the particles are rewritten in terms of the particle density or included in the exchange-correlation functional, which allows to solve the system numerically.

However this comes with a price: the quality of the simulation depends uniquely on how precise the exchange-correlation functional $E_{\text{xc}}[n]$ is. This functional is meant to include all many-body effects of exchange and correlation, and it actually can be explicitly written in terms of the Hohenberg-Kohn approach as

$$E_{\text{xc}}[n] = F_{\text{HK}}[n] - (T[n] + E_{\text{Hartree}}[n]), \tag{2.18}$$

or in a more illustrative manner as

$$E_{\text{xc}}[n] = \langle \hat{T} \rangle - T[n] + \langle \hat{V}_{\text{int}} \rangle - E_{\text{Hartree}}[n]. \tag{2.19}$$

From the last equality one notice that the exchange and correlation functional is the difference of the kinetic and internal interaction energies between the true interacting many-body system and the fictitious independent-particle one. This is the reason why there is no exact form of the total energy functional and therefore, an approximation of $E_{\text{xc}}[n]$ need to be done to solve the problem. Before addressing this issue, we present the final Kohn-Sham equations that need to be solved to obtain the ground state density and energy, which are a direct result of minimizing the functional via variations under orthonormalization constraint and using Lagrange multiplier. The K-S Schrödinger-like equation is

$$(H_{\text{KS}} - \epsilon_j)\psi_j(\mathbf{r}) = 0, \tag{2.20}$$

where ϵ_i are the eigenvalues and H_{KS} is the effective Hamiltonian

$$H_{\text{KS}}(\mathbf{r}) = -\frac{1}{2}\nabla^2 + V_{\text{KS}}(\mathbf{r}), \tag{2.21}$$

with

$$\begin{aligned} V_{\mathrm{KS}}(\mathbf{r}) &= V_{\mathrm{ext}}(\mathbf{r}) + \frac{\delta E_{\mathrm{Hartree}}}{\delta n(\mathbf{r})} + \frac{\delta E_{\mathrm{xc}}}{\delta n(\mathbf{r})} \\ &= V_{\mathrm{ext}}(\mathbf{r}) + V_{\mathrm{Hartree}}(\mathbf{r}) + V_{\mathrm{xc}}(\mathbf{r}). \end{aligned} \tag{2.22}$$

We briefly comment here how the exchange-correlation functional is obtained. Different methods have been proposed to obtain an approximation of $E_{\mathrm{xc}}[n]$. Nowadays there are two main methods, the *Local Density approximation* (LDA) and the Generalized-Gradient approximation (GGA). Both of them, at the same time, have different particular implementations and details. The basic idea is to use the exchange and correlation function of an homogeneous electron gas with the same charge density as the real target system. Furthermore, the LDA only takes into account *local* values (that is, only evaluating $n(\mathbf{r})$), and gives surprisingly good results. The GGA takes into account, apart from the local term, derivatives terms (that is, gradients of $n(\mathbf{r})$) and for more complex systems give better values than the LDA at the expense of computational time.

2.1.4 Numerical and computational techniques

To finish section 2.1 we point out two interesting and important concepts in the framework of DFT but that arise from a numerical (or computational) point of view:

- The *self-consistent cycle* is a numerical approach for solving the K-S equations iteratively. Basically, one starts with a guess for the particle density $n(\mathbf{r})$ and solve the KS equations, from which one can obtain again a new value of $n(\mathbf{r})$. If the initial guess for the density was the correct one, then it is called *self-consistent* and the calculation is finished. If the initial guess and the output density are different, then it is not *self-consistent* and the output density is used as initial guess for the next cycle until their difference is closer than a certain threshold value.
- Fortunately (in terms of computational complexity) it turns out that the electrons which are close to the core region are chemically inert, meaning that they are not involved in chemical reactions. This allows to describe them using what is known as a *pseudopotential*, which replaces the atomic nucleus

and core electrons by a constructed potential whose charge is reduced by the number of core electrons. This is extremely useful, since for calculations of atoms with high atomic number (which would need lots of orbital wave functions) the number of *active* electrons is reduced to a value close to its valence number.

2.2 Linear scaling approach

The DFT formulation described in the section 2.1 has been one of the main developments of ESC in the past decades and has allowed researchers to obtain a huge range of properties of different systems that very likely would have been impossible to test experimentally. Moreover, its accuracy and reliability is very good compared with other methods. However, if one wants to go even further and calculate electronic structures of systems with a relatively high number of atoms (greater than a few hundred atoms) then DFT becomes unfeasible even running on supercomputers. The reason is that, for large systems, the Kohn-Sham approach scales to the power of three with the number of atoms in the system ($\sim N^3$). This effect may not be noticed for small system sizes, since there are other tasks with lower scaling but larger prefactor. In fact, the task that scales with N^3 consists in calculating scalar products among all orbitals of the system, needed to impose orthogonality on the system, and its contribution to the total computational time does not become important until a few hundred of atoms are reached. However, when this point is reached, only by increasing the system size a bit, the calculation become extremely expensive or even unfeasible.

In the path to overcome this barrier, in the last years some research groups have managed to achieve a *linear* scaling with respect to the system size by developing a new approach and computational method which is justified by the so-called *nearsightedness* (term first introduced by W. Kohn in [3]). The main idea of the approach consists in only considering interactions between electrons that are closer to each other than a threshold radius. In this section we present formally this extra approximation and its justification, along with the new formalism.

The starting point of linear scaling DFT is the Kohn-Sham approach. However, it is very useful and illustrative to change the formalism and work with *density matrices* rather than with the usual orbital wave functions. To this aim, we define the density matrix as

$$\rho(\mathbf{r},\mathbf{r}') = \sum_i f(\varepsilon_i)\psi_i(\mathbf{r})\psi_i(\mathbf{r}'), \tag{2.23}$$

where $f(\varepsilon_i)$ is the Fermi function given by

$$f(\varepsilon_i) = \frac{1}{1+e^{(\varepsilon_i-\mu)/(k_B T)}} \tag{2.24}$$

and determines the occupation of the i-th orbital. In (2.24), μ is the chemical potential, k_B the Boltzmann constant and T the electronic temperature, often assumed to be zero. By modifying equations (2.15) and (2.16), one can rewrite the central quantities of DFT in terms of the density matrix as

$$\begin{aligned}
T &= -\frac{1}{2}\int \nabla^2 \rho(\mathbf{r},\mathbf{r}')\,|_{\mathbf{r}=\mathbf{r}'}\,\mathrm{d}\mathbf{r}',\\
E_{\text{pot}} &= \int V_{\text{KS}}(\mathbf{r}')\rho(\mathbf{r}',\mathbf{r}')\mathrm{d}\mathbf{r}',\\
E_{\text{tot}} &= \int H(\mathbf{r}')\rho(\mathbf{r},\mathbf{r}')\,|_{\mathbf{r}=\mathbf{r}'}\,\mathrm{d}\mathbf{r}',\\
n(\mathbf{r}) &= \rho(\mathbf{r},\mathbf{r}).
\end{aligned} \tag{2.25}$$

The latter operators can also be discretized using a finite orthonormal basis set $\phi_\alpha(\mathbf{r})$, leading to

$$\begin{aligned}
H_{\alpha\beta} &= \int \phi_\alpha(\mathbf{r})H(\mathbf{r})\phi_\beta(\mathbf{r})\mathrm{d}\mathbf{r},\\
K_{\alpha\beta} &= \iint \phi_\alpha(\mathbf{r})\rho(\mathbf{r},\mathbf{r}')\phi_\beta(\mathbf{r}')\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}'.
\end{aligned} \tag{2.26}$$

With this analogous formalism, one can obtain both the energy and the total number of particles in the system by computing matrix traces,

$$\begin{aligned}
E &= \mathrm{tr}(\mathbf{HK}),\\
N &= \mathrm{tr}(\mathbf{K}).
\end{aligned} \tag{2.27}$$

Up to this point, a new formalism has been introduced, but no extra approximation has been applied. Thus, the computational cost of the calculations should be the same. In order to reduce the computational cost of DFT it is first needed to discuss what *locality* exactly means. In classical mechanics, properties are *local*, meaning that two bodies are localized and can only interact either in contact or if they interact with a self-generated potential. However, quantum mechanics introduces the *non-locality*, that is, particles are an extended wave function in space and could interact even far away from each other. A clear example to illustrate this idea is the antisymmetry of a many-electron wave function, applied to any pair of electrons in the system regardless the distance between them. However, the question that may arise is: to what extent does this *non-local* property really affect a quantity of a single electron? In other words, is a quantity of a given particle p_i modified by the effect of other particles p_j far away from p_i (so that the modification is noticeable)? If the answer were *no*, *i.e.* the effect of distant particles were negligible, it would be justified not to compute every binary interaction among the electrons and reach the aforementioned linear scaling.

It turns out that already by 1964 it had been proven [11] that the density matrix elements decay exponentially with respect to the distance $|\mathbf{r} - \mathbf{r}'|$ between particles for insulators and metals at finite temperature, and algebraically for metals at zero temperature. The reason for the decay properties of the density matrix lies in the interference among the various eigenfunctions, thereby canceling contributions where $\mathbf{r}$ and $\mathbf{r}'$ are far away [8]. In figure 2.1 is depicted this effect in a particular system that we have studied, a perfect body centered cubic lattice of Tungsten with 729 cells and 1458 atoms. It is clear that regions with low values for the density matrix (in light blue) are larger than regions with higher values (in dark blue). Thus, this matrix representation demonstrates its sparsity and the locality of the system. In chapter 4 this will be extensively commented for different system sizes, proving that indeed there is a huge localization property to exploit.

Due to the rapid decay of the matrix elements $\rho(\mathbf{r}, \mathbf{r}')$ it is justified to *cut* the density matrix at a given radius, *i.e.* imposing $\rho(\mathbf{r}, \mathbf{r}') = 0$ for $|\mathbf{r} - \mathbf{r}'| > \lambda$, where λ is some system-dependent constant that characterizes the decay behavior. This procedure transforms the density matrix into a *sparse* matrix with mostly zero elements, and allows to use new numerical algorithms that scale only linearly with the system size.

There are different methods to exploit the sparsity of the density matrix and achieve the linear scaling. In this thesis we will only discuss one particular method and implementation in the BigDFT code. This will be covered in chapter 3.

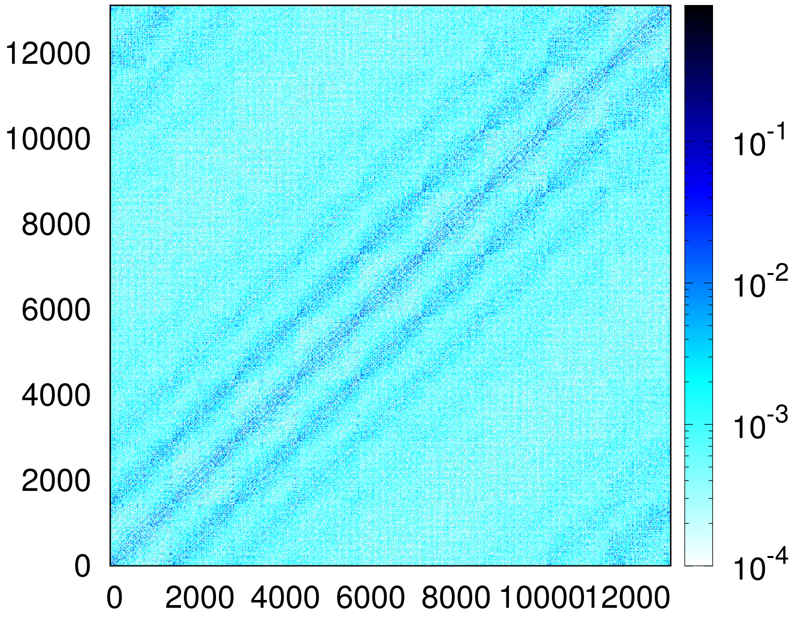


Figure 2.1: Density matrix for a system of 1458 atoms of Tungsten structured in a perfect bcc lattice. Matrix elements in light blue correspond to low values, while elements in dark blue correspond to higher values. The localization of regions in dark blue proves the sparsity of the matrix.

2.3 Real *vs* reciprocal space simulations

There is still, in the general DFT framework, a widely used theoretical (and computational) technique that in some situations can make the computational time of calculations much smaller, while preserving its quality and precision. This technique is known as the k-points method and it is based on a transformation from the real space to the *reciprocal* space, by means of the Fourier Transform. The only condition that the target system must fulfill consist of periodicity in the three spatial dimensions. The method gives such a good performance that it is even used for some special infinite non-periodic systems for which the periodicity is broken only in a specific region of the infinite system. This is done by imposing periodicity in some artificial boundaries located as far as possible from the *defect*. This results in locating its *image* far enough so that there is no interactions between them.

The theoretical background of the method was mainly motivated by the study of pure periodic systems, such as *crystal structures*. This category of solid materials are characterized by being arranged in a highly ordered microscopic structure, forming a crystal lattice extended on three spatial directions – the most iconic example being probably the Diamond crystal structure– and therefore, the positions of the nuclei are repeated periodically in space. One can completely characterize the whole system just by specifying the types and positions of the nuclei in one repeat unit (known as *primitive unit cell*) and the rules that describe the repetition (known as *translations*). The study of such crystal structures is broad, and here we will only comment on the necessary ones to understand how the k-point method works.

Probably, the most basic concept in crystallography is the *lattice of translations*, which defines the set of all possible translations in space. In three dimensions it is defined as

$$\mathbf{T}(\mathbf{n}) \equiv \mathbf{T}(n_1, n_2, n_3) = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3, \tag{2.28}$$

where $\mathbf{a}_i$ are the primitive translation vectors which also define the unit cell dimension, and n_i are integers. However, this construction does not define a unique *unit cell*, for which there are infinite possibilities. The Wigner-Seitz cell is a special one, consisting in the locus of points in space that are closer to that lattice point than to any of the other lattice points.

Although the definitions above may be helpful in order to construct the system, in terms of computational cost it has not given any insight for DFT. The k-points method is a theoretical and computational technique that exploits the periodicity of the system so that one can obtain all the information of the system only by doing the calculation in the Wigner-Seitz of the reciprocal space, which is defined as the *first Brillouin zone* (BZ). The construction of the BZ is as straightforward as the Wigner-Seitz cell once the lattice of translations of the reciprocal space is found. Similarly to the real lattice of translations, the reciprocal one is defined as

$$\mathbf{G}(m_1, m_2, m_3) = m_1\mathbf{b}_1 + m_2\mathbf{b}_2 + m_3\mathbf{b}_3, \tag{2.29}$$

where m_i are integers and $\mathbf{b}_i$ are the reciprocal primitive vectors. The last vectors can be found using the relation

$$\mathbf{b}_i \cdot \mathbf{a}_j = 2\pi\delta_{ij}. \tag{2.30}$$

Arising from the concept of *Bloch waves*, *i.e.* wave functions of the form $\psi_{i,\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}}u_{i,\mathbf{k}}(\mathbf{r})$, where $e^{i\mathbf{k}\cdot\mathbf{r}}$ is a plain wave and $u_{i,\mathbf{k}}(\mathbf{r})$ a periodic function, we also define the set of vectors $\mathbf{K}$ in the reciprocal space (k-vectors) which in our particular case can be expressed as

$$\mathbf{k} = \frac{2\pi}{R_{\text{max}}} \left(\frac{n_1}{|b_1|}\mathbf{b}_1 + \frac{n_2}{|b_2|}\mathbf{b}_2 + \frac{n_3}{|b_3|}\mathbf{b}_3 \right), \tag{2.31}$$

where R_{max} is the maximum distance between two lattice points in the system. It is important to notice that, in (2.31), $\mathbf{k}$ refers to a set of vectors, since we obtain a different one for each possible combination of $\{n_1, n_2, n_3\}$. Moreover, it is worth to notice that there is exactly the same amount of vectors k as cells in the system.

By applying the Bloch Theorem and carefully analyzing and exploiting the periodicity in space for the operators, one may find that these k-points in the BZ specify all possible eigenstates of the system and the Hamiltonian becomes k-dependent. Considering an arbitrary Hamiltonian $\hat{H}$ for the system, one can solve the equations

$$\hat{H}(\mathbf{k}) u_{i,\mathbf{k}}(\mathbf{r}) = \epsilon_{i,\mathbf{k}} u_{i,\mathbf{k}}(\mathbf{r}) \tag{2.32}$$

for each k and obtain the eigenstates and energies, using $\psi_{i,\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} u_{i,\mathbf{k}}(\mathbf{r})$.

For the purpose of this thesis, it is important to note that:

(i) The k-points method assumes periodicity in the system and exploits it by redirecting the problem into the reciprocal space.

(ii) The calculations are done only in the BZ. Since the number of k-points is equal to the number of unit cells in the real space, one may want to increase as much as possible the number of k points in the simulation to increase its accuracy.

(iii) The computational time required to solve the problem is highly reduced, compared to the time one would need to solve the same problem in real space.

(iv) There are numerical methods to find symmetries within the k points and accelerate even more the calculations.

As commented above, probably all DFT codes have the k-points method implemented. It is used for periodic systems and for some specific non-periodic systems. However, there are several questions that may arise regarding the method; what can one do when the system is really big and non-periodic? How can one know if the use of the reciprocal space in periodic systems is introducing errors to the calculations due to the choice of a small unit cell? For most of those big systems, there were no other feasible ways to do ESC. Thus it is not possible to compare the results of the "artificial periodic" system with k-points with the real space ones. This is caused by the high computational cost of latter one. Fortunately, due to the rapidly development of the linear scaling algorithms and the development of new highly parallelized DFT codes, nowadays it is possible reach system sizes that were only possible to reach by means of the reciprocal space, thus providing a new way to perform calculations and compare results. In chapter 4 we will test the linear scaling algorithm of the BigDFT code to perform new real space simulations that are comparable to the corresponding sizes of reciprocal space simulations. It is important to remark that while the k-points sampling is implemented in the traditional cubic version of the BigDFT code, up to now it is not yet implemented in the linear version.

2.3.1 Interests and challenges regarding metal calculations

As commented in chapter 1, this work is motivated by the application of DFT techniques in fusion materials research. The target category of materials to be tested consists mainly of pure metals and metallic alloys. This choice is justified by the fact that materials belonging to those categories generally exhibit good behavior under radiation pressure. However, metallic systems present extra difficulties in DFT calculations compared to other types of solid materials. To understand it we must first introduce what is known as the *HOMO-LUMO gap.* Among all the allowed energy levels where electrons can be (orbitals), there are two categories, *occupied* and *unoccupied* levels. The occupied level with highest energy is know as the *highest occupied molecular orbital* (HUMO) and the unoccupied level with lowest energy as *lowest unoccupied molecular orbital* (LUMO). The energy difference between these particular levels is known as the HOMO-LUMO gap. For a non excited molecule or atom, the HOMO-LUMO gap also correspond to the energy that one would need to give to the system so that an electron in the HOMO level could jump to the LUMO one.

It turns out that in ESC the value of this gap also gives information about the complexity of the calculation. The reason is that the surface that separates the reciprocal space into the occupied and unoccupied levels, the so-called *Fermi surface*, may introduce jump discontinuities to some quantities that need to be integrated numerically in DFT. If the HOMO-LUMO gap is wide enough, it may not introduce much errors, but when the HOMO-LUMO gap is very low or zero, as is the case in metals and alloys, then the Fermi surface is abrupt and difficult to localize and, therefore, the numerical integration may lead to huge numerical errors and make it hard to reach the convergence in the self-consistent cycle [2].

There are different approaches to deal with this problem. For example, increasing the number of k-points so that the integral becomes more accurate, in k-space simulations, or using specific numerical integrating methods in order to obtain more accurate results in the integral. Another technique consists of increasing the temperature of the system, reflected in the Fermi distribution in equation (2.24), which smoothes the functions at the Fermi level. However, this last and effective approach has to be applied very carefully when performing DFT calculations since the price to pay is that one is not anymore using the *real* functional at 0 temperature, but instead another one with a higher electronic temperature, thus introducing an error that must be considered. For these reasons, up to now the linear scaling DFT has been rather limited for metallic systems compared to non-metallic ones, such as organic molecules, and there are still lots of test to perform to check its validity and limitations.

3 | Implementation of DFT in the BigDFT code

In the last chapter the theoretical background of DFT and particular useful methods related to this work have been presented. However, as commented before, each implementation of DFT in a code has very specific characteristics and each of them may be optimized for different types of calculations. The basis set, the numerical solvers, the exchange-correlation potential and the type of pseudopotentials are usual code-dependent characteristics of DFT. In this section we review the specific implementation of DFT in the BigDFT code (see [4, 8, 12–14] for reference papers, web page `http://bigdft.org/`), which will be used in the next chapter to perform the calculations.

3.1 Wavelets

The basis set of a DFT code is used to express the KS orbitals and, therefore, it takes a primary role in the development of the code. Nowadays, the most used basis sets are *plain waves*, *Gaussian-type* orbitals and numerical-like orbitals. Plain waves basis sets are characterized by being non-localized, thus they are optimal for periodic, homogeneous systems. Also, the accuracy increases systematically with the number of basis elements. However, they are not adaptive. Oppositely, Gaussian-type orbitals are optimized for isolated open structures (thanks to their locality), such as molecules, and a small number of basis elements may be enough to reach a moderate accuracy. Their weakness is that they are not systematic. This is the reason why the BigDFT developers decided to work with a new type of basis set, the Daubechies wavelets, which have the potential to take the best properties of both the plane waves and the Gaussian-like basis sets. Daubechies

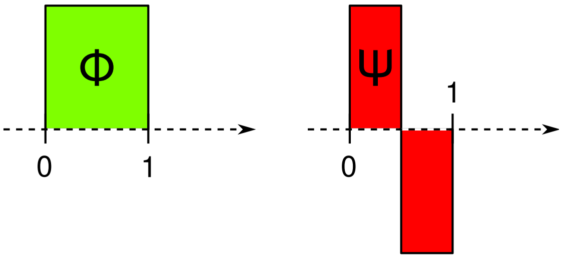


Figure 3.1: Mother scaling and wavelet functions of the Haar family. Image from [8].

wavelets form an orthogonal and systematic basis set, and at the same time are localized and adaptive, along with other useful properties which will be explained in the following.

Among all families of wavelets, BigDFT uses the so-called *least asymmetric Daubechies of order 16* [15]. In the following we introduce general properties of wavelets by using a much easier family, the Haar wavelet family, in order to properly understand and justify the choice of wavelets as a basis set. Then specific features of the Dauchebies family will be mentioned.

Each wavelet family consists of a mother scaling function ϕ and a mother wavelet ψ. The Haar family, which is the simplest possible wavelet, is characterized by the following scaling and wavelet mother functions, respectively (see also figure 3.1),

$$\begin{aligned}\phi(x) &= \begin{cases} 1 & 0 \leq x < 1, \\ 0 & \text{otherwise}, \end{cases} \\ \psi(x) &= \begin{cases} 1 & 0 \leq x \leq 1/2, \\ -1 & 1/2 \leq x < 1, \\ 0 & \text{otherwise}. \end{cases}\end{aligned} \tag{3.1}$$

To generate an orthonormal basis set out of these mother functions, one can use scaling and shifting operations:

$$\begin{aligned}\phi_i^k(x) &\propto \phi(2^k x - i), \\ \psi_i^k(x) &\propto \psi(2^k x - i),\end{aligned} \tag{3.2}$$

where k describes the resolution of the basis – the smaller k is, the thinner func-

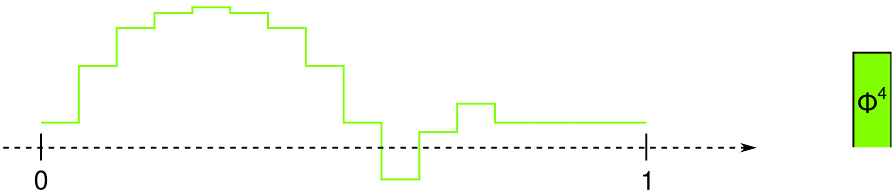


Figure 3.2: Exact representation of the piecewise function by means of only scaling functions. Image from [8].

tions become – and the index i represent the localization in space. Using these scaled and shifted scaling and wavelet functions, one can approximately represent a continuous function.

In order to illustrate how they work, we use the Haar wavelet family to represent a piecewise function in the interval $[0, 1]$. In figure 3.2 the target function is depicted. As a first approach one can write the expansion in terms of the basis set only by means of the scaling functions (without wavelets) as

$$f(x) = \sum_{i=0}^{15} s_i^4 \phi_i^4(x), \tag{3.3}$$

where $s_i^4 = f(i/16)$ (the real value of the function). In this case the representation in the basis set is exact due to its step-wise nature. Oppositely, if the function was continuous this would lead to an approximation, and in the limit $k \to \infty$ the representation would be exact.

The main advantage of the wavelet representation lies in its capability of using different resolution levels to express a function. To this aim, one notices that a scaling function at resolution level k can be written as a linear combination of a scaling function and a wavelet at resolution level $k - 1$. By performing this transformation, equation (3.3) transforms to

$$f(x) = \sum_{i=0}^{7} s_i^3 \phi_i^3(x) + \sum_{i=0}^{7} d_i^3 \psi_i^3(x), \tag{3.4}$$

where the new coefficient can be obtained from those on equation (3.3) by the

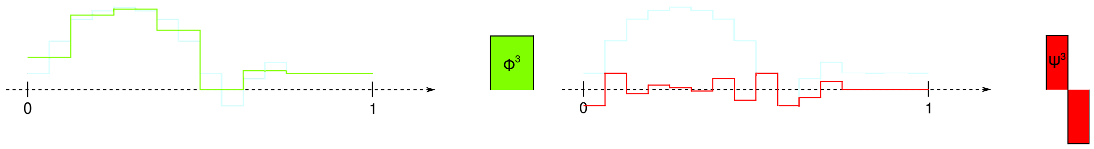


Figure 3.3: Scaling (left) and wavelet (right) functions contribution, as expressed in equation (3.4). The scaling part is "smooth" while the wavelet represent rapid changes. Image from [8].

following relations

$$\begin{aligned} s_i^{k-1} &= \frac{1}{2}s_{2i}^k + \frac{1}{2}s_{2i+1}^k, \\ d_i^{k-1} &= \frac{1}{2}d_{2i}^k - \frac{1}{2}d_{2i+1}^k. \end{aligned} \tag{3.5}$$

Finally, repeating the procedure until level 0 one obtains the last form

$$f(x) = s_0^0\phi_0^0(x) + d_0^0\psi_0^0(x) + \sum_{i=0}^{1} d_i^1\psi_i^1(x) + \sum_{i=0}^{3} d_i^2\psi_i^2(x) + \sum_{i=0}^{7} d_i^3\psi_i^3(x). \tag{3.6}$$

We notice that:

- The total number of expansion coefficients is the same as in equation (3.3), *i.e.* 16.
- The *information* of the function is split into the wavelets and the scaling function. By looking at figure 3.3, which depicts the wavelet and scaling part, one notice that the scaling function represents a *smoothed* version of $f(x)$, while the wavelet represent the rapidly varying corrections to the smoothed function. Note that in this particular case the word *smooth* may lead to confusion, due to the non-continuous nature of the example function. However, in the continuous case this makes much more sense.
- These "rapid variations" in $f(x)$ are more or less localized in some intervals of the domain of $f(x)$ – the function does not have rapid changes everywhere – thus, some of the wavelet coefficient may be zero. This allows to *compress* data, since we will be actually using less than 16 coefficients.

The conclusion is that a mixed scaling function and wavelet expansion allows us to represent a function by a *compressed* basis set.

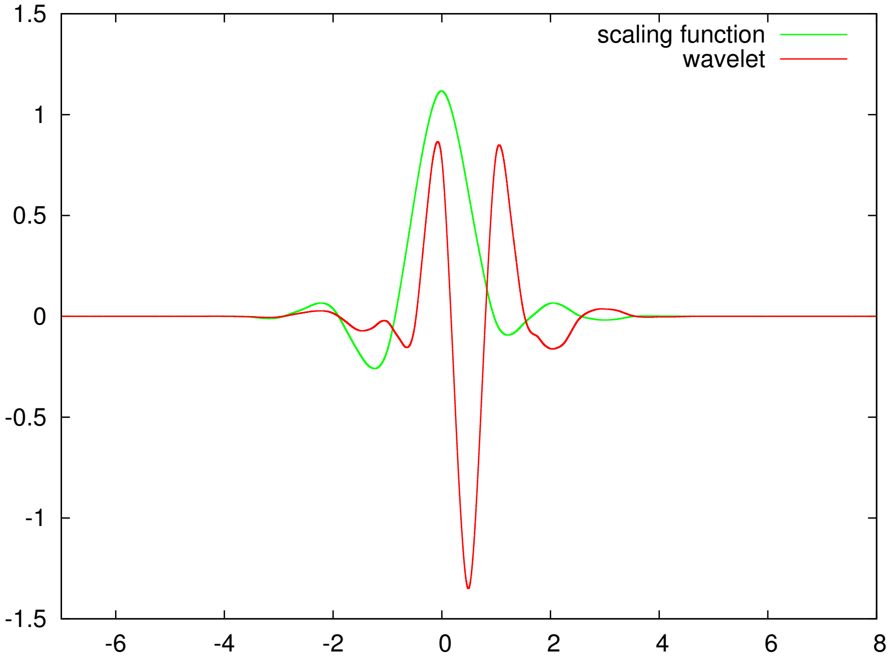


Figure 3.4: Mother scaling and wavelet functions for the least asymmetric Daubechies of order 16. Image from [8].

The wavelet family used by BigDFT, the *least asymmetric Daubechies of order 16*, is orthogonal and has compact support. In figure 3.4 are depicted the scaling and wavelet mother functions for this family. Notice that now the functions are continuous. Apart from the important properties already discussed above, one of the main advantages of using wavelets is that they allow you to choose the resolution level and even separate different regions of space so that you can choose where to increase the resolution and where to decrease it. Such method allows to optimize the resources and computational time, since the parts of the system with more important information can be more carefully represented, while the rest of the system is represented with a simpler resolution. Specifically, BigDFT allows three levels of resolution: (i) In the *fine region*, which is the locus of point in space closer to the nuclei, grid points carry one scaling function and seven wavelets. (ii) In the *coarse region*, grid points carry only one scaling function, because they are further from the nuclei and the associated quantities in this region tend not to change (so) rapidly. Its resolution is half that of the fine region. (iii) Points neither in the coarse or fine regions do not carry any scaling nor wavelet functions, since they do not contribute to any quantity.

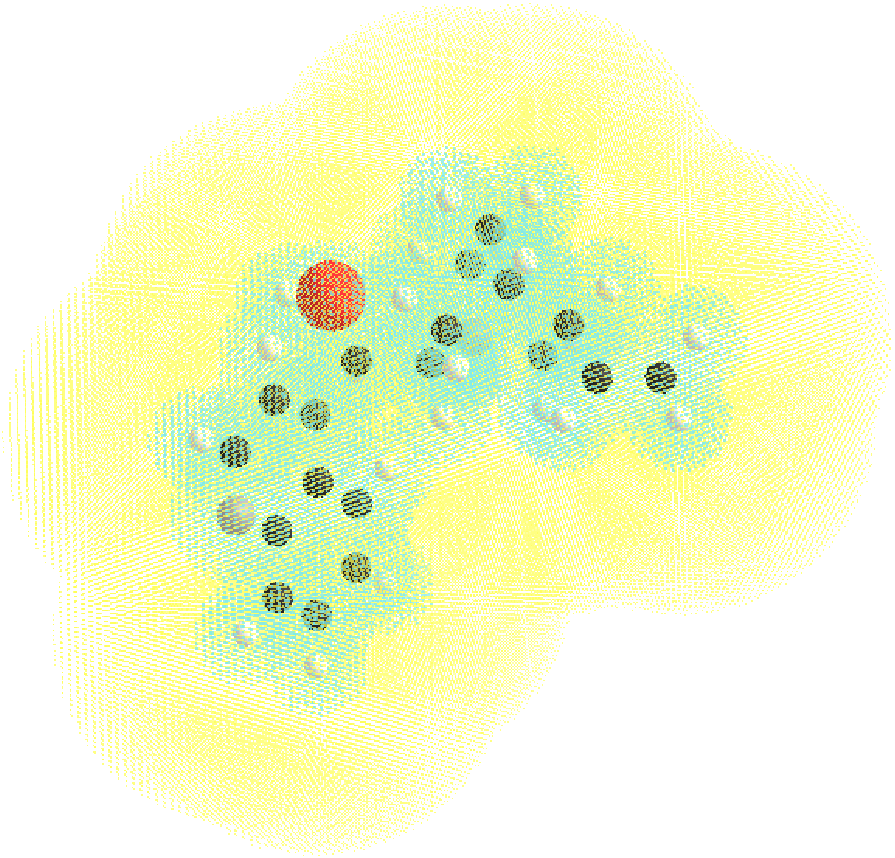


Figure 3.5: Visualization of the fine region (green) and coarse region (yellow) for cinchonidine which has the chemical formula $C_{19}H_{22}N_2O$. We notice that the fine region is closer to the nuclei while the coarse region is further. Image from [8].

Notice that the grid spacing is the same in the entire simulation box, thus the different resolution levels are obtained by associating a number of scaling and wavelet functions to each grid point depending on the region they belong to. The construction of the regions is very straightforward; consider two radii $R_{\text{fine}} < R_{\text{coarse}}$, then the region of space defined by the union of all spheres with radii R_{fine} centered at each of the nuclei is the fine region. The region in space defined by the union of all spheres with radii R_{coarse} centered at each of the nuclei minus the fine region is the coarse region. The value of R_{fine} and R_{coarse} are given by the product of an atom-dependent constant and a user-specified factor. In figure 3.5 one can see both regions for the case of a small molecule.

3.2 Specific linear scaling implementation

In section 2.2 the theoretical background upon which the linear scaling approach is based has been presented. However, there are several ways to exploit the decay properties of the density matrix elements, which at the same time can be combined. Below we present some key points of the linear scaling approach of BigDFT.

Optimal basis and support functions

The *optimal basis* approach is used in BigDFT. This approach works using what is known as support functions, which will be denoted as ϕ_α. These new functions act as a set of auxiliary basis functions, in a higher level than the real basis set, the wavelets. Thus, wavelets are used to express the support functions ϕ_α and then orbitals are expressed via support functions. By means of this approach the dimension of the density matrix is considerably reduced while it still preserves its quality.

The support functions can be written in terms of the wavelets and scaling functions as

$$\phi_\alpha(\mathbf{r}) = \sum_{j_1,j_2,j_3} s^\alpha_{j_1,j_2,j_3} \varphi_{j_1,j_2,j_3}(\mathbf{r}) + \sum_{j_1,j_2,j_3} \sum_{\nu=1}^{7} d^\alpha_{j_1,j_2,j_3;\nu} \psi^{(\nu)}_{j_1,j_2,j_3}(\mathbf{r}), \tag{3.7}$$

where φ are now the scaling functions and ψ, as before, are the wavelets; s and d are their respective coefficients and the notation $\varphi_{j_1,j_2,j_3}(\mathbf{r})$ stands for $\varphi(x-j_1)\varphi(y-j_2)\varphi(z-j_3)$. It is also important to note that the set of support functions used in a particular simulation is optimized in-situ during the calculation. This approach allows one to use a smaller set of support functions and thus optimize both the computational time and the quality of the calculation.

As already stated in chapter 2.2, there is a given radius further from which particles do not interact with each other. This radius, the cornerstone of the linear scaling DFT, will be referred as *cutoff* radius or r_{cut}. Using the wavelet notation, one can write the approximation as

$$\left\{ \begin{array}{l} s^\alpha_{j_1,j_2,j_3} = 0 \\ d^\alpha_{j_1,j_2,j_3;\nu} = 0 \end{array} \right. \quad \text{if} \quad |\mathbf{R}_{j_1,j_2,j_3} - \mathbf{R}_\alpha| > r_{\text{cut}}, \tag{3.8}$$

which is equivalent to the expression given in section 2.2 using the matrix elements directly. Independently from whether considering the fine or coarse region, if condition (3.8) is fulfilled, the value of a grid point turns directly into zero.

Hybrid optimization mode

The hybrid mode is the method used to optimize the support functions. It is build so that a trace minimization and an energy minimization via the Hamiltonian are performed using the same target function and in a smooth way. The transition from the trace to the energy is controlled by a factor included in H_α, which will be referred as *confinement.* H_α is the usual Hamiltonian H plus an extra factor that creates a confining potential. The target function is written as

$$\Omega^{\mathrm{hy}} = \sum_{\alpha} K_{\alpha\alpha} \langle \phi_\alpha | H_\alpha | \phi_\alpha \rangle + \sum_{\beta\neq\alpha} K_{\alpha\beta} \langle \phi_\alpha | H | \phi_\beta \rangle \, , \tag{3.9}$$

where Ω^{hy} is the quantity that needs to be optimized. We note that as the confinement of the potential H_α is decreased, $H_\alpha \to H$, thus leading to the expression of the total energy.

Density kernel

Once the support functions ϕ_α are obtained, the density matrix can also be expressed by means of this new basis as

$$\rho(\mathbf{r}, \mathbf{r}') = \sum_{\alpha,\beta} \phi_\alpha(\mathbf{r}) K_{\alpha\beta} \phi_\beta(\mathbf{r}'), \tag{3.10}$$

where the matrix K is the *density kernel.* It can be easily checked that K is actually the density matrix expressed in the basis of the support functions ϕ_α.

The physical quantities expressed in terms of the orbitals or density matrix in chapter 2, can now be rewritten in terms of the support functions and the density kernel. In particular, BigDFT uses the expressions

$$\begin{aligned} n &= \sum_{\alpha,\beta} K_{\alpha\beta} \langle \phi_\alpha | \phi_\beta \rangle \,, \\ E &= \sum_{\alpha,\beta} K_{\alpha\beta} \langle \phi_\alpha | H | \phi_\beta \rangle \,. \end{aligned} \tag{3.11}$$

to compute the density n and the energy E of the system. Using equation (3.10) and the fact that the density matrix ρ is sparse, one can write a condition analogous to (3.8), in terms of the density kernel, as

$$K_{\alpha\beta} = 0 \quad \text{if} \quad |\mathbf{R}_\alpha - \mathbf{R}_\beta| > r_{\text{cut}}, \tag{3.12}$$

where $\mathbf{R}_\alpha$ is the central position of the support function ϕ_α. Finally, BigDFT uses the *Fermi Operator Expansion* (FOE) method to calculate the density matrix. This method directly calculates the density kernel K as a function of the Hamiltonian H and the particular expansion of the operator in BigDFT uses the Chebyshev polynomials.

FOE and direct diagonalization

The density kernel can be optimized using different methods. The choice of this method is very important, since it controls a large factor of the total computational time of the simulation. Depending on the complexity and system size, one may choose one or another. The linear version of BigDFT allows to choose between a small range of methods. Here we comment the most useful ones available today:

- **Fermi Operator Expansion (FOE):** As briefly commented in the density kernel explanation, the FOE directly calculates the density kernel in the basis of the support functions. Among its different flavors BigDFT uses the Chebyshev FOE. Basically it expresses the density matrix as a function of the Hamiltonian by a Chebyshev polynomial expansion of order n_{pl}. The more important feature of this method in terms of its applicability is that it scales linearly with the system size, *i.e.* if one wants to achieve the pure linear scaling this method has to be chosen. However, its prefactor is big and thus for small system sizes other methods might be preferable.

- **Direct diagonalization:** this is the most straightforward method, since it optimizes the density kernel by performing a direct diagonalization of the Hamiltonian matrix in the basis of the support functions. It is very useful for small system sizes, since it has a small prefactor. However, when the system size increases it becomes prohibitive due to its cubic scaling with the system size.

In figure 3.6 the flowchart of the linear scaling BigDFT method is depicted. We note that there are two inner loops, one for the support functions, being optimized by the hybrid mode, and one for the density kernel, being optimized either by FOE or by direct diagonalization. The outer loop controls the convergence to the final value.

Boundary Conditions

To finish the list of remarkable items of the linear scaling version of BigDFT, we focus on the boundary conditions (BC). BigDFT allows very flexible BC. In particular, one can choose between: (i) an isolated system, characterized by free BC, (ii) surfaces BC, consisting in a 2D periodic plus 1D isolated, (iii) the usual periodic 3D BC and (iv) wires BC, consisting in 1D periodic plus 2D isolated. Actually, these BCs are not only implemented in the linear version but also in the cubic one.

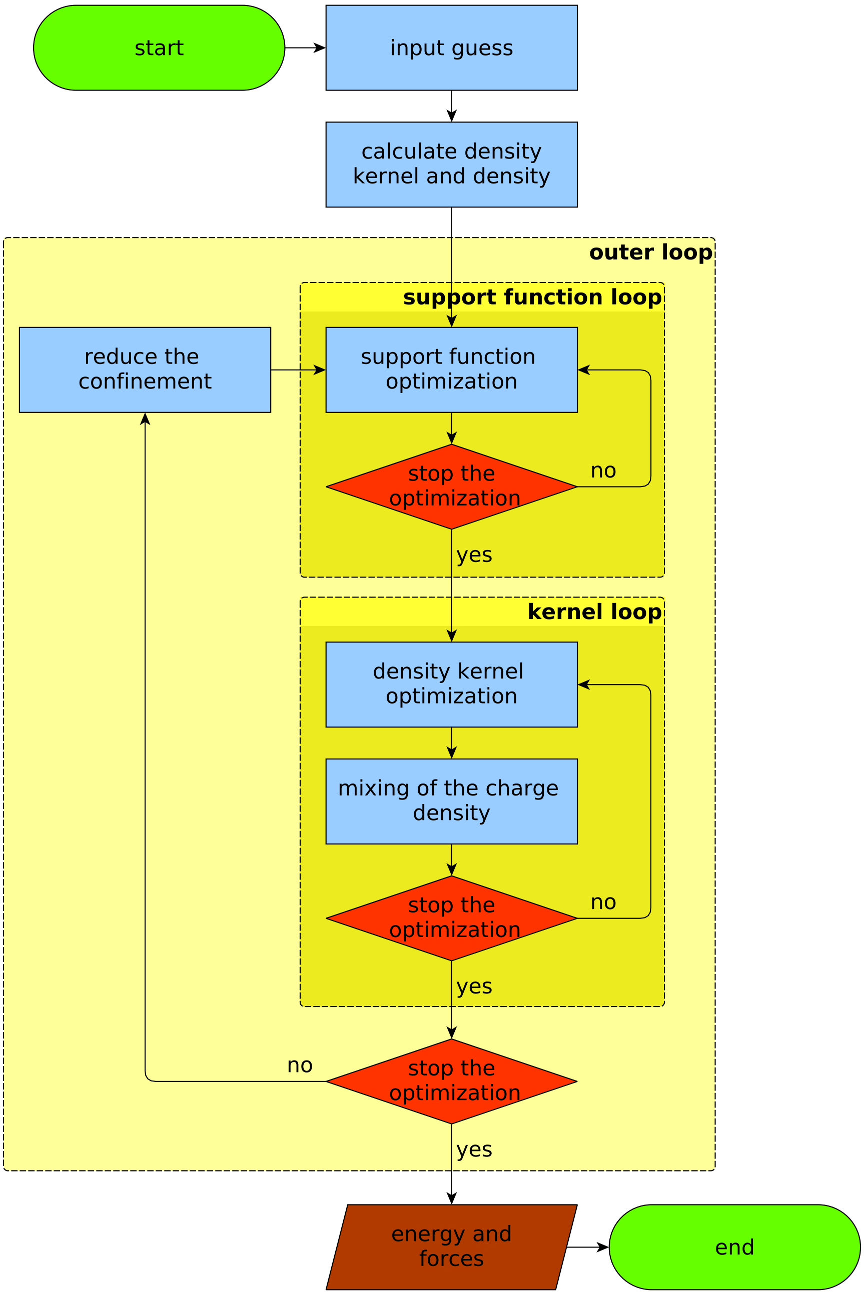

start
input guess
calculate density kernel and density
outer loop
support function loop
reduce the confinement
support function optimization
stop the optimization
no
yes
kernel loop
density kernel optimization
mixing of the charge density
stop the optimization
no
yes
stop the optimization
no
yes
energy and forces
end

Figure 3.6: Flowchart of the linear scaling DFT in BigDFT. The inner loops control minimization of support functions and the density kernel. The outer loop controls the convergence to the real solution of the problem.

4 | Validation of the linear version of BigDFT for large metallic systems

In chapters 2 and 3 the theory upon which the BigDFT code is based has been presented, as well as its specific implementation. A complete understanding of both the approximations and techniques used in DFT is essential in order to properly analyze the results obtained in any simulation. In this chapter we use the BigDFT code to perform electronic structure calculations for different systems and sizes with the aim of thoroughly testing the validity of the linear scaling approach of BigDFT for large metallic systems.

For this purpose, we will perform calculations for exactly the same system and set of parameters using both the cubic and linear scaling version of the BigDFT code. This provides a sufficient test of the validity of the linear scaling approach, since the accuracy and quality of the cubic scaling version have already been shown in different papers (see, for example, [12]). It is important to remark that while the k-points sampling is implemented in the traditional cubic version of the BigDFT code, up to now it is not yet implemented in its linear version, thus all calculations performed with the linear version will not use a k-point grid. In some specific cases, results will also be compared with those obtained with other DFT codes, such as *Abinit* [16] and *SIESTA* [17] and with the force field code (not DFT) *LAMMPS* [18]. *V_sim* has also been used for visualizing the atomic structures (`http://inac.cea.fr/L_Sim/V_Sim/`).

The chosen system to validate the linear scaling version of BigDFT is a perfect Tungsten body centered cubic (bcc) lattice. This structure is shown in figure 4.1 for a system size of 128 atoms (64 minimal cells) from two different perspectives.

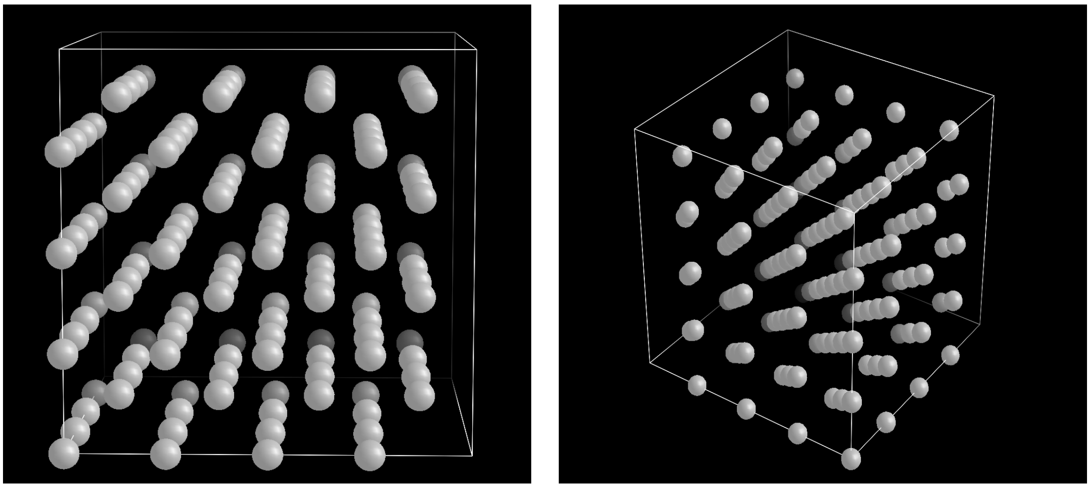


Figure 4.1: Representation of a bcc lattice with 128 atoms from two perspectives. Atoms are represented by white spheres centered at the nuclei positions.

The bcc structure is a type of crystal structure, characterized by the minimal cell shown in figure 4.2. By repeating this cell in space with a translation rule one can build the whole system. The translation vectors of the bcc structure have all the same length and are orthogonal to each other.

The choice of the structure and material is not an arbitrary one. Rather the following reasons have been considered:

- Tungsten is a very promising element for the fusion community, thanks to good properties such as a high threshold for sputtering, a high melting point and a high thermal conductivity [19].
- In terms of convergence of the simulation, Tungsten is a challenging element which requires a large number of basis functions per atom in order to be properly described. Thus, if the code works well with Tungsten it is probable that it will also work well with most other metals.
- Its geometry is rather simple, which makes its visualization and the understanding of the results much easier.

The chapter is structured as follows. In section 4.1 we study the minimal cell of our chosen system. This is required to choose the correct set of parameters so that the simulation becomes as accurate as possible within a reasonable walltime.

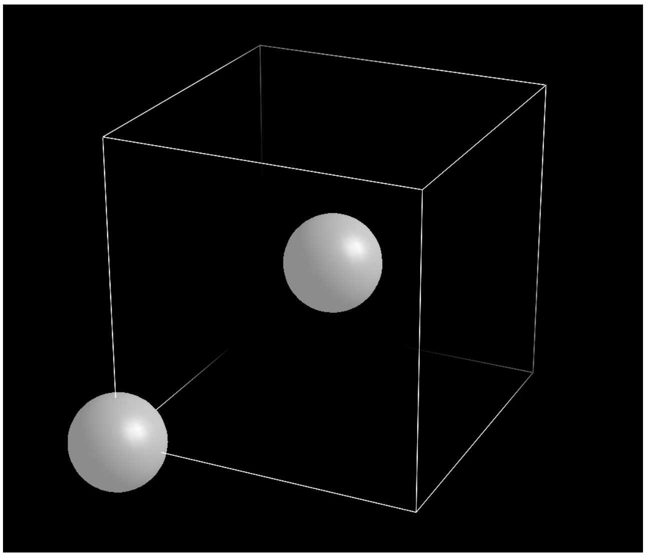


Figure 4.2: Minimal cell of the bcc structure. This minimal cell, along with the translations, carry all the required information to build the whole system.

In section 4.2 the validity of the linear scaling approach is shown, both in terms of quality of the computation and scalability with the system size.

4.1 Minimal cell study and parameter optimization

Every DFT simulation with large complexity requires a pre-analysis of the system. This is done to ensure that:

1. The walltime of the calculation is feasible and agrees with the type of simulation one wants to perform. For instance, if one only cares about a qualitative understanding of the system, the set of parameters should be adjusted so that the computation does not become extremely expensive. Oppositely, if one needs also quantitative results, the parameters have to be tested so that the desired precision is reached and the walltime does not exceed the maximum allowed time.

2. There is no waste of computing resources. For example, it is not worth to set a very tiny grid spacing if the same results can be obtained with a larger

spacing.

3. Physical quantities make sense according to the target system. An example would be, for instance, a calculation of a system in equilibrium that yields a huge pressure. This would lead to a non accurate simulation, since the system would not be stable. In this particular case, the atom spacing would require further tuning.

In this section we show the results regarding this analysis. In particular we show the grid optimization, the lattice parameter optimization and the k-points convergence. Since performing the parameter optimization using a big cell would be extremely expensive and would lead to a long process, we will use in the entire section the minimal cell of the system (figure 4.2). For this purpose, and as explained in section 2.3, we will use the k-points method and the cubic version of BigDFT. This choice is completely justified, since the system is purely periodic. Summarizing, by following this approach we obtain exactly the same results as by using a big cell but saving lots of computational time.

4.1.1 Grid optimization

The grid spacing of the simulation cell plays a very important role in the simulation. A too large value of the parameter could lead to a poor quality of the simulation, since lots of information would be lost in regions between grid points. Choosing a too small value could lead to enormous waste of resources or even to a non-feasible simulation.

In order to find a good guess for the grid spacing we follow the next steps: 1) Fix the whole set of parameters except from the grid spacing. 2) Run simulations of the minimal cell with k-points with different grid spacing. 3) Choose a relevant quantity of the simulation, make a plot and analyze its convergence. 4) If the values of the quantity have not converged yet, choose smaller values for the grid spacing and come back to point 2). If the values converge at some point, pick its corresponding value of the grid spacing.

By performing this methodology, one ensures that the grid spacing is small enough so that the results are good but at the same time there is no waste of resources. In our particular case of Tungsten, the total energy of the system has been the target quantity to check for convergence. In figure 4.3 we show how the

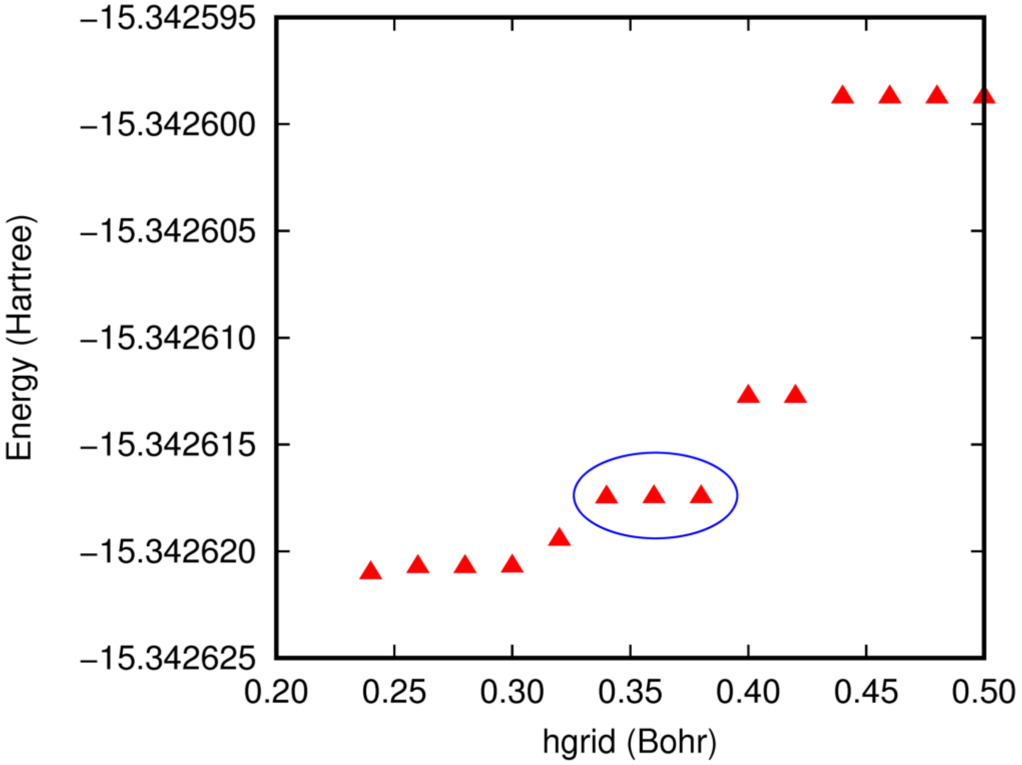


Figure 4.3: Energy convergence with respect to the input parameter `hgrid` that controls the grid spacing. The circle shows the `hgrid` value chosen for the simulations in the rest of the work. Notice that the three circled values lead to the same grid spacing, thus the particular choice is arbitrary.

total energy of the system varies with respect to the `hgrid` parameter, given in Bohr, which controls the grid spacing. We first notice that for some ranges of the `hgrid` parameter the total energy of the system does not change at all. This is caused by the fact that the `hgrid` parameter internally modifies the value of the grid spacing so that it becomes a divisor of the simulation box size. Thus, for the aforementioned ranges the real grid spacing is equal even though the `hgrids` are different. This ensures that the grid is well fitted in the simulation box. We also see that for large values of `hgrid`, corresponding to low resolution, the energy changes are large. However, as we decrease the value of `hgrid` the energy changes also decrease, converging to a given value.

Following this test, we have decided to work with an `hgrid` of 0.38 Bohr. This value allow us to obtain reliable results while maintaining a feasible computational time.

4.1.2 Lattice parameter optimization

The next parameter to optimize is the lattice constant (or lattice parameter). The lattice constant controls the distance between the atoms. More precisely, it corresponds to the length of the translation vectors described in (2.28). Even though its value should be a constant in nature, actually each code use values that slightly differ. This is caused by the fact that each implementation of DFT has small variations in the calculation of quantities in the system, related to the different basis sets, pseudopotentials and others. Although these deviations are often very small, it is very important to take them into account. The emergence of pressure in the system lead to an unstable configuration, thus transforming the target system into another one with a higher ground state energy. To solve this issue, a careful analysis of the pressure in the system needs to be carried out.

In this work we use two different quantities to optimize the lattice constant. Similarly to the `hgrid` optimization, we start by fixing the set of all parameters except for the lattice constant. As an initial guess we can take the experimental value of the lattice parameter. Then we perform simulations with this value and with small variations of it. Once the results are obtained we analyze the following quantities:

- **Pressure P:** this is the most straightforward approach. The pressure of the system is given by the thermodynamic expression

$$P = -\left(\frac{\partial E}{\partial V}\right)_{T,S}, \tag{4.1}$$

  where V the volume, T the temperature, E the energy and S the entropy. If the code one is using computes the pressure of the system, then pick the value closest to zero pressure. If no result is close enough, perform more simulation with improved values of the lattice parameter. Since BigDFT outputs the pressure of the system at the end of the simulation, we have been able to do this analysis. In figure 4.4 the obtained results are presented. We notice that a particular value is just crossing the line $P = 0$, thus leading to an equilibrium system. In brown we have also plotted a linear fit, which agrees with the obtained results in a neighborhood of the equilibrium parameter.

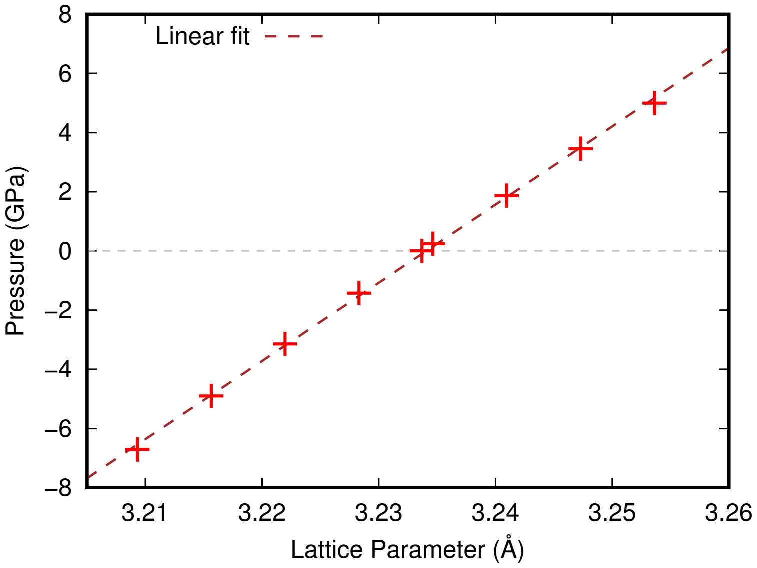


Figure 4.4: Pressure of the system given by BigDFT as a function of the lattice parameter. Crosses in red are the obtained results. The brown dashed line is a linear fit. The dashed grey line corresponds to $P = 0$.

We have obtained the value $\ell_{\text{par}} = 3.2337$ Å.

- **Energy E:** From (4.1) one can deduce that a minimum in the energy corresponds to a zero in the pressure. In figure 4.5 we plot the energy of the system in terms of the lattice parameter. We notice that its minimum indeed corresponds to the same value of the lattice parameter as the one that results to zero pressure in figure 4.4. This proves that the results are consistent.

4.1.3 *k*-points convergence

To finish this section we address the k-points convergence. As already stated in chapter 1, the goal of this work is to evaluate the validity of the linear approach of BigDFT *without* the k-points method. However, in this preparatory state of the work it is very useful to do a convergence test in terms of the k-points. In section 2.3 we have seen that the number of k-points in the reciprocal space is related to the number of unit cells in the real space system. Thus, by performing this study one can extrapolate the ideal size that the system should have in order to obtain accurate results.

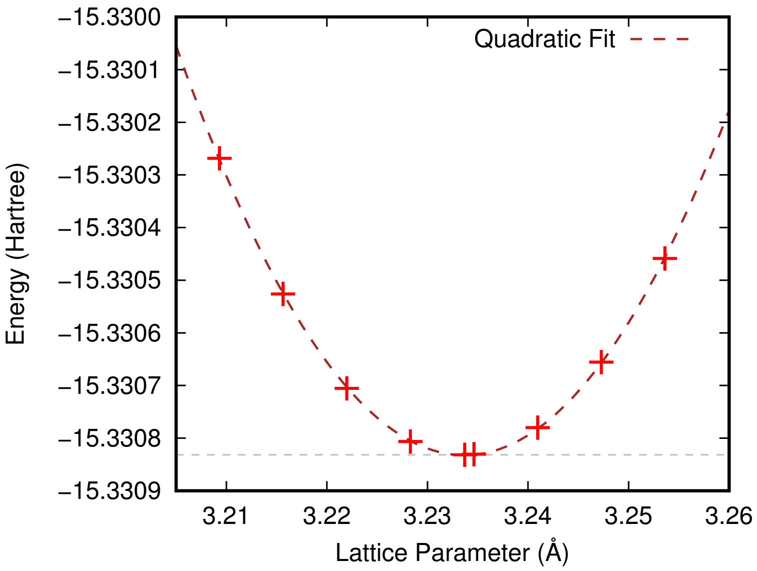


Figure 4.5: Energy of the system given by BigDFT as a function of the lattice parameter. Crosses in red are the obtained results. The brown dashed line is a quadratic fit. The dashed grey line corresponds to the minimum value, with $\ell_{\text{par}} = 3.2337$ Å.

The general idea is very similar to those in section 4.1.1 and 4.1.2: using the values of `hgrid` and ℓ_{lat} already optimized, fix the set of all parameters except for the number of k-points. Then, run simulations with an increasing number of k-points. The obtained values of the energy should converge at some point, thus giving the ideal system size in order to obtain accurate results. In figures 4.6, 4.7 and 4.8 the energy convergence in terms of the number of irreducible k-points, the number of atoms and the box length, respectively, is shown. Notice that the yellow lines correspond to the converged value, the green lines connecting the points are fictitious and EKS refer to the Kohn-Sham energy. All plots correspond to the same simulations, since the number of k-points, the number of atoms and the box length are related among each other. The number of k-points and its corresponding box width in real space are both given by the BigDFT output. With this information, the number of atoms in the system is very easily obtained by

$$n_{\text{atoms}} = 2\, n_{\text{unit cells}} = 2\left(\frac{\text{Box length}}{\ell_{\text{par}}}\right)^3. \tag{4.2}$$

We notice that for a system smaller than approximatively 30 irreducible k-

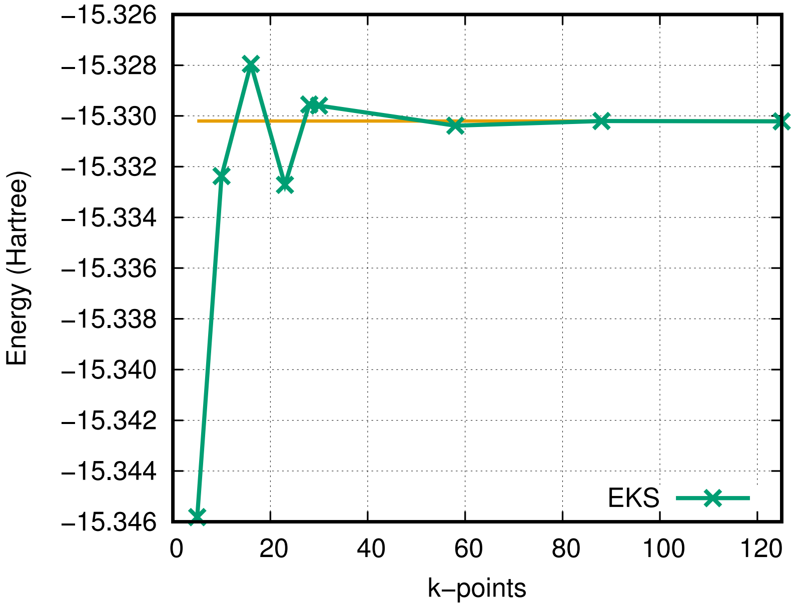


Figure 4.6: Energy of the system given by BigDFT as a function of the number of irreducible *k*-points.

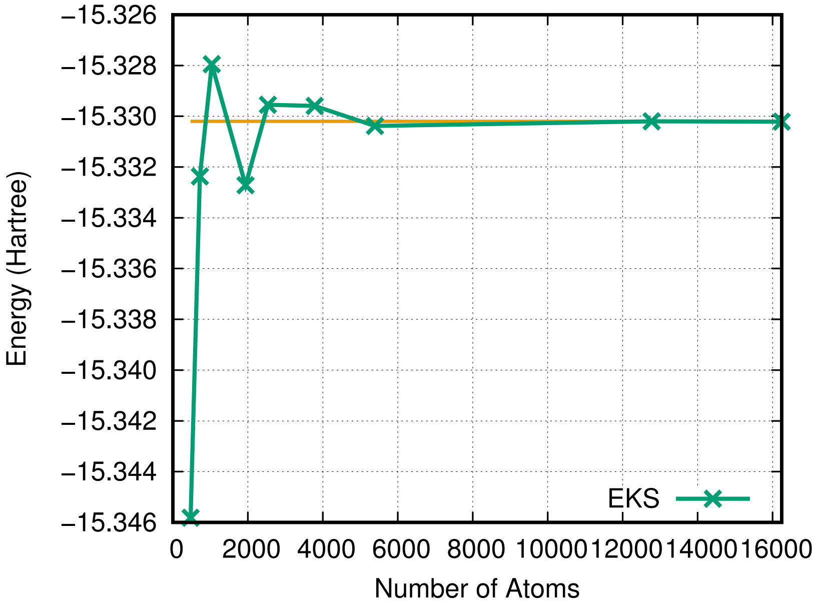


Figure 4.7: Energy of the system given by BigDFT as a function of the number of atoms.

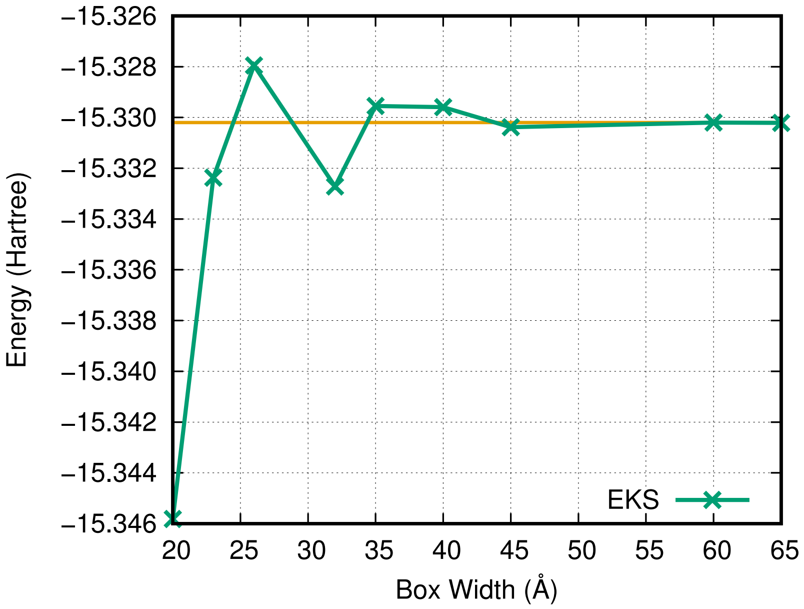


Figure 4.8: Energy of the system given by BigDFT as a function of the box length.

points, 3000 atoms or a box length of 35 Å, the energy values clearly oscillates. For a larger system the energy converges to a given value.

The obtained result seem quite alarming. In chapter 1 we commented that usually DFT simulations range from tens to hundreds of atoms, while figure 4.7 suggest to use around three thousand. This is the main reason why the k-points method is so important in electronic calculations for periodic systems. The reciprocal space approach allows us to reach size limits that would be impossible to perform in real space calculations (from now on, we will refer to real space calculations as Γ-point or simply Γ calculations). However, there are some situations on which the use of the k-points method is not so-well justified and the use of bigger systems would be of great help. This will be discussed in section 5.

To check that these results are not only obtained by BigDFT we have also performed a similar study with the code Abinit. In figure 4.9 we plot the energy convergence as a function of the number of irreducible k-points for both codes. We observe that Abinit's results seem to oscillate more before reaching the convergence. However, at approximately the same number of k-points both codes show a similar convergence.

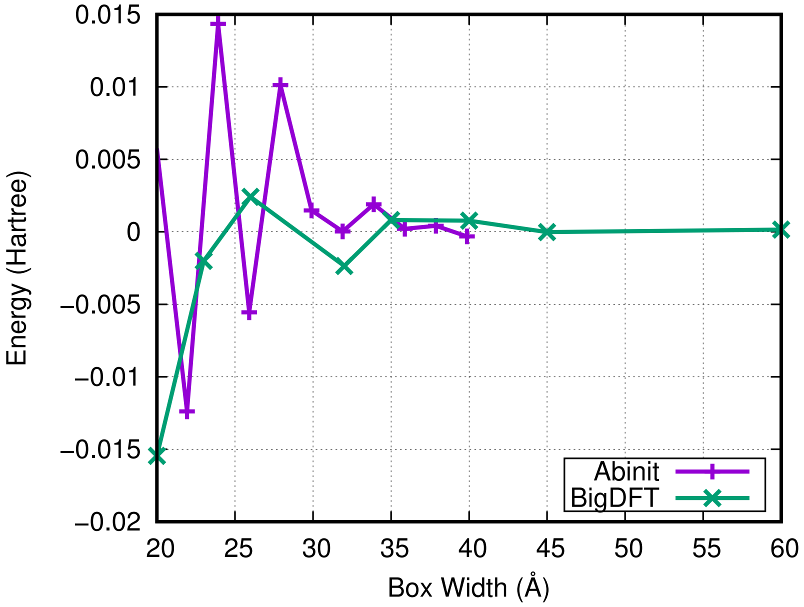


Figure 4.9: Energy convergence as a function of the box width for BigDFT and Abinit codes.

4.2 Towards large systems with the linear scaling approach

Once the set of parameters of the calculations has been optimized, one can proceed to the relevant simulations to test the validity of the linear scaling in BigDFT. In this section we perform Γ-point ESC for the system presented in the beginning of this section with an increasing number of atoms. In particular, the system sizes read $n_{\text{atoms}} = 2\,k \times k \times k$, with $k = 4, 5, \ldots, 10$. k refers to the number of minimal cells in the system and the factor 2 stands to the fact that there are two atoms per minimal cell. Thus, we will show results for a Tungsten bcc lattice with 128, 250, 432, 686, 1024, 1458 and 2000 atoms.

The results will be compared with those obtained with the cubic version of BigDFT. However, the larger simulations can not be compared, since the cubic version would require an unfeasible computational time. It is also important to recall the difficulties regarding metal calculations: a vanishing HUMO-LUMO gap and a slow decay of the density matrix elements. To deal with these problems it is required to use many support functions, namely 9 per atom, and an electronic

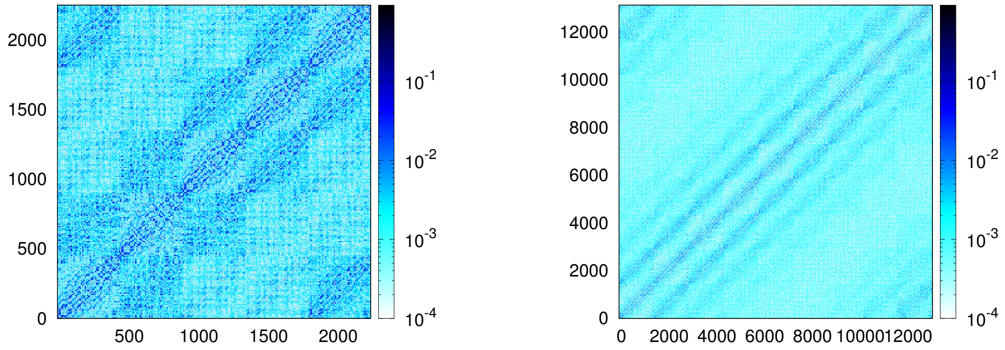


Figure 4.10: Sparse matrix representation for 250 atoms (left) and 1458 atoms (right). Dark blue represents higher values of the matrix elements, while lighter blue represents low values.

temperature different from zero, in our case $1{\cdot}10^{-3}$ Hartree.

In this section we will first, in section 4.2.1, present the results of the linear scaling simulations, including the energy of the systems and other quantities of interest. Secondly, in 4.2.2, we present the comparison with the cubic approach. Finally, in 4.2.3, the performance of both approaches will be analyzed.

4.2.1 Tungsten calculations up to 2000 atoms

In figure 4.10 we show a representation of the locality of the system for two different sizes. Even though the locality is clear in both systems, it increases as the system size grows, as expected. This is due to the fact that for big systems there are more atoms outside the sphere defined by the cutoff radius and thus the number of interactions per atom is highly reduced. We notice that both matrices present a local nature, but the effect is clearly more evident in the system with 1458 atoms. This effect implies that the linear scaling is better exploited in big systems. By using the linear scaling approach we will reach convergence without using the k-points sampling.

Using the parameters obtained in section 4.1, we have prepared the aforementioned structures of a Tungsten bcc lattice using from 128 to 2000 atoms. In figure 4.11 the total energy per atom of each structure is plotted. The energy is normalized to the number of atoms in order to compare the results obtained for different sizes. We observe that the initial oscillations for small systems vanish as

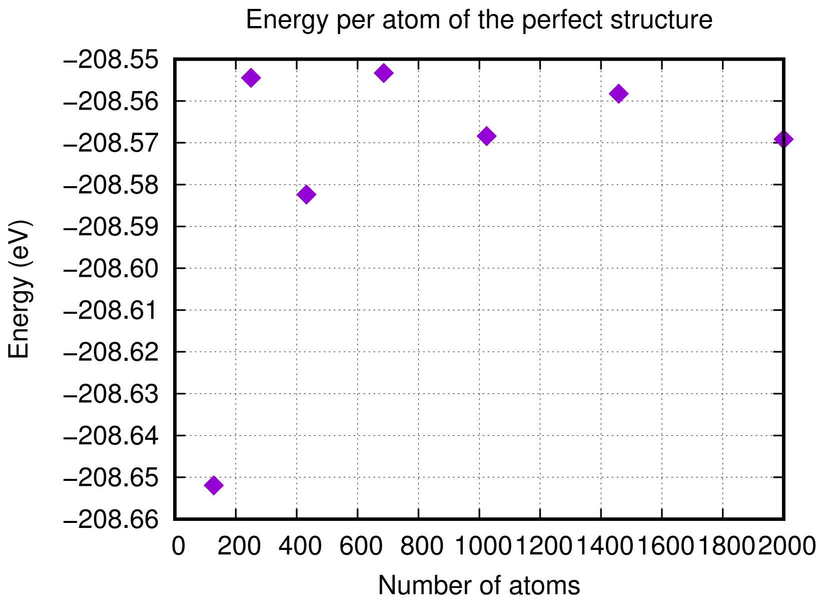


Figure 4.11: Energy of the Tungsten bcc structures as a function of the system size, obtained with the linear version of BigDFT code at Γ.

the system size grows and the energy per atom converges to a particular value. It is important to notice that in order to check for convergence in the values, differences between consecutive energy values need to be analyzed, rather than percentage errors between them. This is a consequence of the fact that the energy as an absolute value is meaningless, since any scalar constant could be added to the values. The energy difference between the first two sizes is about 0.1 eV, while for the last pair it decreases to 0.01 eV, a 10% of the first one. Thus, to reach convergence a large number of atoms is required, as already predicted by the k-points study. Actually, even with 2000 atoms it is not completely clear that the convergence has been reached. Even though an energy difference of 0.01 eV is acceptable to have a qualitative understanding of the system, it is not completely satisfactory if accurate quantitative values need to be derived from the respective energy values.

In order to have a wider understanding of the behavior of the results we have also plotted in figure 4.12 the energy convergence using the two different types of BigDFT simulations. The results obtained with the linear version of BigDFT with Γ-point are represented in green. The results obtained with the cubic version using a minimal cell with 2 atoms and k-points are represented in purple. It is

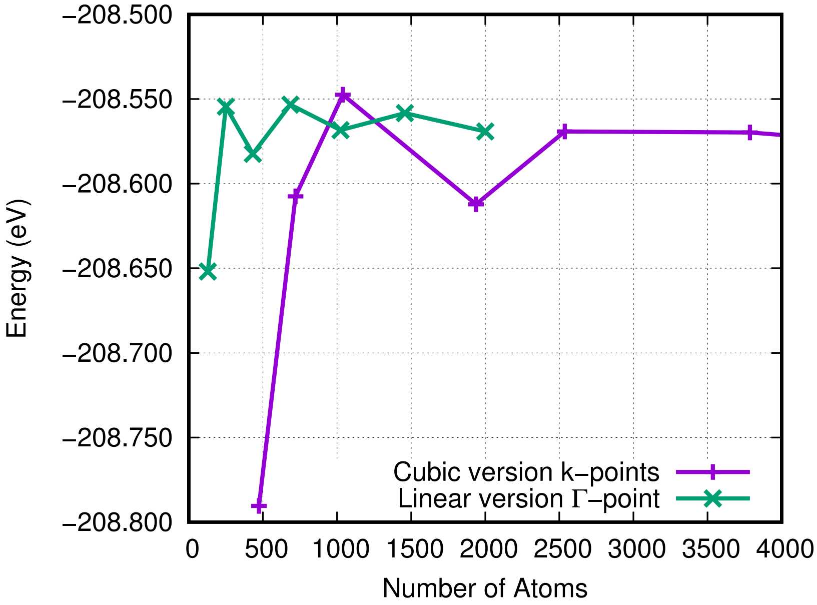


Figure 4.12: Energy of the Tungsten bcc structures as a function of the system size. Data in green obtained with the linear version of the BigDFT code and Γ-point. Data in purple obtained with the cubic version of BigDFT, a minimall cell with two atoms and varying the number of k-points.

important to recall that the number of atoms represented in the k-points approach is determined from the minimal cell and the number of k-points used. We observe that both versions show a similar behavior, although the k-points approach present more oscillations and seems to convergence slower. It is worth pointing out that these differences are due to the fact that for these particular simulations we have only considered the irreducible number of k-points, thus making the conversion from reciprocal to real space not completely exact.

We have also checked that the zero pressure predicted by the minimal cell study is also confirmed by the big systems. The results are plotted in figure 4.13. We note that there are small oscillations but all of them within an acceptable range of approximatively ± 1 GPa. The latter can be justified using the results already presented in figure 4.4: a difference in the lattice parameter of 1% lead to pressure changes of around 12 GPa. Thus, the maximum pressure difference observed in figure 4.13, namely 1,2 GPa, is virtually equivalent to a lattice parameter deviation of 0.1%. Thus, we prove that the system is in equilibrium.

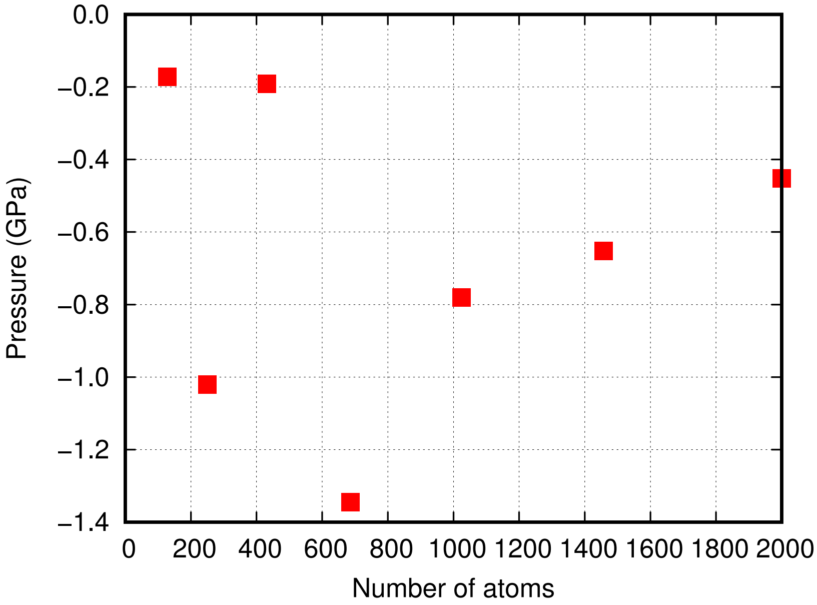


Figure 4.13: System pressure as a function of the system size obtained with the linear version at Γ.

4.2.2 Comparison with the cubic approach

In order to prove that the linear scaling approach is giving accurate results and that the *nearsightedness* concept agrees with this type of metals, one needs to compare the results with those obtained with the traditional cubic version within the same conditions, *i.e.* both performing Γ-point calculations. For this purpose we have run simulations with the cubic approach with system sizes ranging from 128 to 686 atoms. Larger systems have been impossible to simulate due to the large required computational time. In figure 4.14 we present the obtained results. We observe that the values from the different approaches are virtually identical except for the system with 128 atoms. The difference for the latter system may be caused by the fact that a cell containing 128 atoms without k-points is a too small cell for a metallic periodic infinite system. Thus, this may lead to small variations in the final energy depending on which version is used.

Therefore, taking into account these reasons and the fact that other system sizes lead to almost the same results for both approaches, we can conclude that the linear scaling version of BigDFT works for large metallic systems.

Extra tests have been made to check the validity of the linear scaling approach.

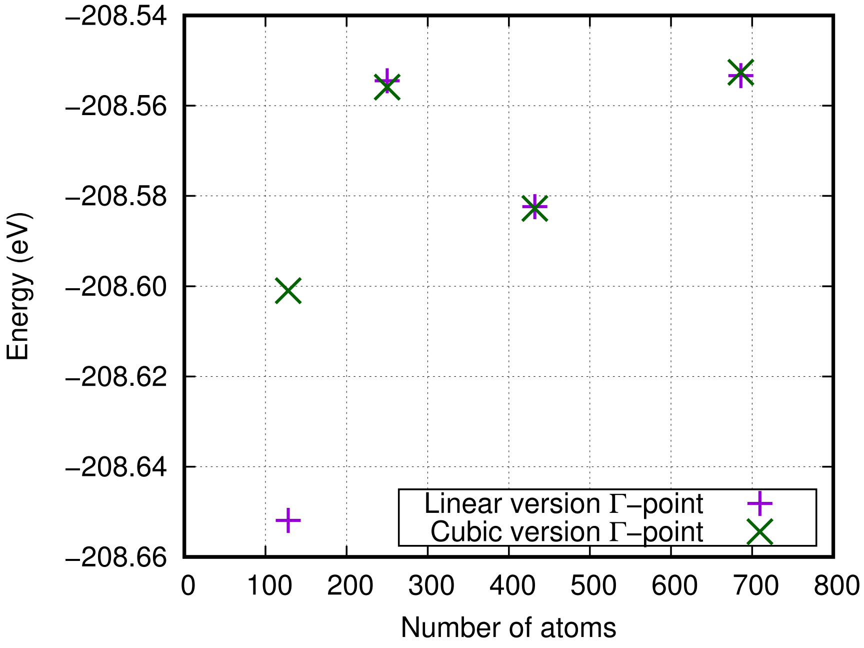


Figure 4.14: Energy of the system for the linear (purple crosses) and cubic (green) approaches and Γ-point.

In this case we compare the linear version at Γ with the cubic version using the unit cell and k-points. We evaluate the energy and pressure of the bcc lattice for different lattice parameters, as in section 4.1.2. The comparison is done using a system of 1458 atoms ($9\times9\times9$ minimal cells) and Γ-point calculation for the linear version and a unit cell with 2 atoms and $9 \times 9 \times 9$ k-points for the cubic version. As stated in section 2.3 these systems should be equivalent. In figure 4.15 the results are plotted. We notice that while the pressure give identical values for both approaches, a small difference is observed for the energy values. This can be easily understood by the discussion in section 4.1.2: energy values should be evaluated by energy differences and not by their absolute value. This explains the small energy deviation from the different approaches, since the difference between the energy predicted by both approaches is almost constant as a function of the volume per unit cell. Oppositely, pressure values are virtually identical since they are computed by means of an energy derivative, thus canceling any constant difference in the energy.

Summarizing, based on the results presented in figures 4.14 and 4.15 we conclude that the linear version of the BigDFT code works for large metallic systems.

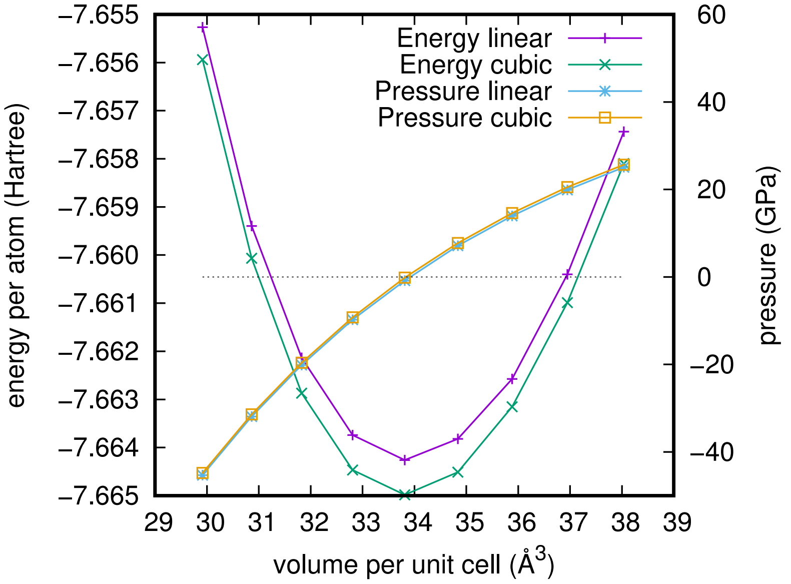


Figure 4.15: Energy per atom and pressure for a Tungsten bcc lattice with 1458 atoms for the linear version at Γ and a unit cell with 9 × 9 × 9 k-points for the cubic version.

In particular, the accuracy of the method has been checked with both the Γ calculations and the k-points method. Our results also demonstrate that it is possible to reach system sizes similar to those achieved with the k-points method without the use of the reciprocal space.

4.2.3 Computing performance and scalability

The reason to develop a linear scaling DFT is to reduce the scalability of the simulations and reach larger system sizes. This is why it is of great importance to analyze the time consumed by the simulations. Furthermore, the use of supercomputers and highly parallelizable codes is compulsory. For the simulations presented earlier in this section we have used a peak number of resources of 7200 cores running in parallel for the same simulation. It is clear that to achieve a good scaling for this large number of resources a very good implementation of MPI and OpenMP is necessary.

In section 3.2 we have seen different specific features of the linear scaling in BigDFT. Among those, the numerical solver is of particular importance regarding the calculations computational time. The most relevant solvers are the direct diagonalization method (`diag`) and the FOE. We recall that for a purely linear scaling approach the FOE should be used. However, due to its large prefactor, for not so large system sizes the `diag` performs better, even though it presents a cubic approach. For the linear version calculations presented earlier in this section we have therefore used the `diag` method. In figure 4.16 we show the total required computational time (CPU-time) for systems from 128 to 686 atoms. It is clear that in the cubic version the cubic scaling becomes already apparent for 432 atoms and makes simulations with a greater number of atoms almost unfeasible. Oppositely, the linear scaling approach scales linearly with the system size and reduces by a factor of 4 the CPU-time of the simulation with 686 atoms. We have also found that the crossover between both versions, *i.e.* the system size at which the linear scaling becomes cheaper than the cubic one, is located at around 150 atoms.

However, due to the choice of the `diag` solver, this linear behavior in the linear version stops with a system size of around 1458 atoms, as shown in figure 4.17. We

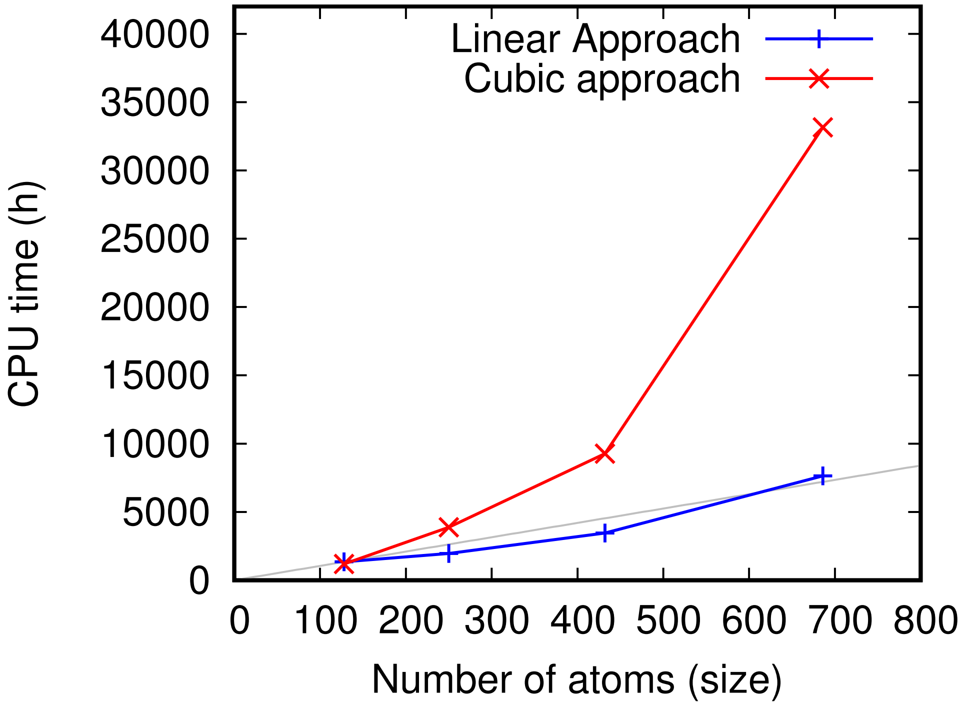


Figure 4.16: CPU-time of the simulations for the linear and cubic versions of BigDFT. The linear version shows a linear CPU-time, represented by the grey line.

must recall that this behavior has nothing to do with the intrinsic linear behavior of the physical description of the system, and it is only related to the `diag` solver. To prove it we have plotted in figure 4.17 as well the CPU-time of the simulations subtracting the time spent in the solver. It is clear that the large increase of CPU-time of the 2000 atoms system is caused by the `diag`. This issue would be completely solved by choosing instead the FOE solver. This would lead to a slightly higher CPU-time for the smaller systems, since the prefactor of FOE is larger. Nevertheless, for bigger system sizes it would scale linearly and, thus, reduce by far the CPU-time required for the `diag` approach.

From the latter reasons it is clear that a repetition of all simulations presented in this section with the FOE solver would be of great interest. From these calculations we could check how computationally expensive would be to perform simulations with even bigger systems and to find the system size at which the crossover between `diag` and FOE is found. We are currently running the aforementioned simulations with FOE but, unfortunately, we still do not have enough results so that they can be presented in the thesis and be properly analyzed. We expect to present an extended study in the upcoming months.

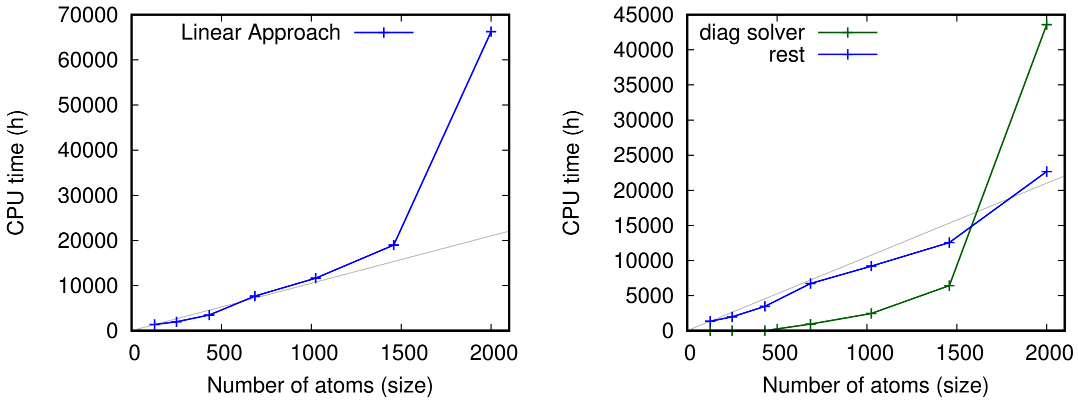


Figure 4.17: Left: total CPU-time for different systems sizes for the linear version of BigDFT with the `diag` solver. Right: the time spent in the `diag` solver has been subtracted from the total CPU-time of the simulations and has been plotted apart. This shows that, for big systems, the intrinsic linear behavior is only broken by the numerical solver.

5 | An application: Tungsten Vacancy Formation Energy

As already stated in chapter 1, this thesis is motivated by the study of particular systems of interest to the fusion community, by means of DFT simulations. Briefly, nuclear reactions produced inside a plasma of a fusion device create highly energetic neutrons and charged particle. When these particles collide into the first wall (or vessel) of the fusion device, important changes in its physical and mechanicals properties can arise. These changes are of importance, since as a result the microstructure of the materials forming the vessel may be changed, affecting its physical behavior.

There are mainly two types of radiation damage in an irradiated material: structural distortions in the atomic structure, the so-called *radiation defects* or simply *defects*, and changes in the chemical composition of the material, known as *transmutation reactions*. In this thesis we focus on the first type. This type of radiation damage needs a multiscale analysis in order to be properly understood. The starting point of defects happens on an atomistic scale, such as a single vacancy or a single self-interstitial atom (SIA). However, these point defects may grow into cascades of defects (clusters) that could even interact over time with each other, requiring other types of models. On the mesoscopic scale the cascades may lead to a particular type of microstructural evolution [20].

A fundamental description of defects on the atomistic scale requires the use of truly *ab initio* models, such as DFT, and also forms the basis upon which models of higher scales are built. Calculations regarding single defects provide key point results such as lowest energy configuration of vacancies and SIAs with high precision. Later on, these values may be used to create higher level models with an accurate parametrization based on the *ab initio* calculations, in this way, providing much lighter models (in terms of computational time) that can still accurately represent atomistic-level systems.

In this chapter we first, in section 5.1, give a short review of point defects and their state-of-the-art. In particular we present the key quantity to analyze, namely the point defect formation energy. Second, in section 5.2, we present the simulations we have performed along with a wide discussion of the results. Those simulations will be done with the linear version of BigDFT at Γ-only. In sections 5.2.1 and 5.2.2 we further analyze the results from two new different perspectives. In section 5.2.1 a new type of simulation is performed, namely a large multi-defect cell. This structure is expected to emulate a k-points grid within the Γ-only approach. In section 5.2.2 we perform a multipole analysis of the defect structures.

5.1 Point defects

Point defects, *i.e.* defects related to a single atom, are divided into two categories, *vacancies* and *self-interstitial atoms* (SIAs). The vacancy occurs when an atom is extracted from its usual place in a perfect lattice (see figure 5.1), while the SIA occurs when an external atom occupies a non-usual place in a perfect lattice (see also figure 5.1). Both can be related, in the sense that the atom that leaves a vacancy may be the one creating a SIA, and are generated by a high-energy particle colliding with it.

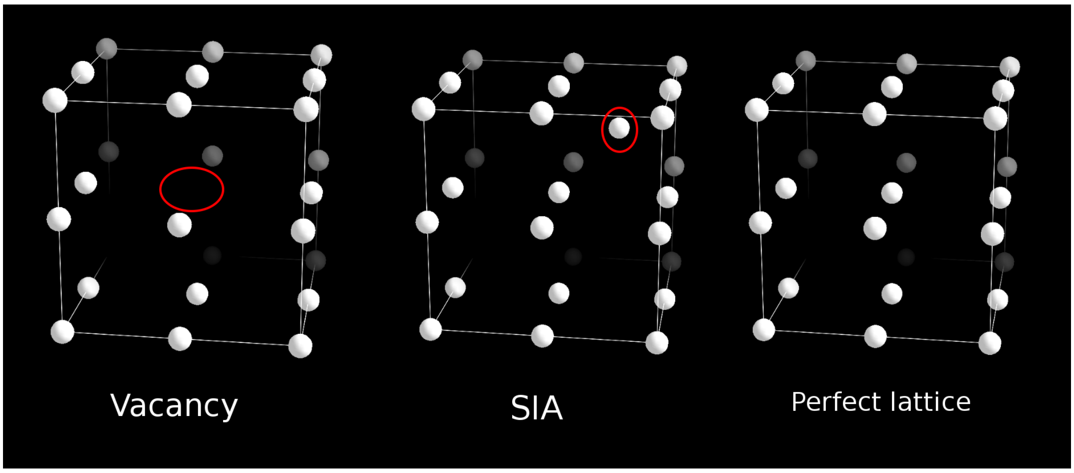


Figure 5.1: Vacancy and SIA point defects for a 27 atoms simple cubic structure.

In this whole section we will analyze the effect of a single vacancy defect in a bcc Tungsten lattice. The vacancy formation energy (VFE) is defined as

$$E_f^v = E_{N-1} - \left(\frac{N-1}{N}\right) E_N, \tag{5.1}$$

and it is a key quantity for fusion materials research. In equation (5.1), E_N refers to the energy of the structure with N particles (perfect structure) and E_{N-1} refers to the energy of the structure with $N-1$ atoms (vacancy or defect structure). It is important to remark that the defect structure is assumed to be relaxed: when an atom is extracted from a perfect system, forces between atoms may arise, leading to a non-stable configuration; a relaxed structure means that this has been taken into account and the affected atoms have been moved so that the system is in equilibrium. This procedure may be very time consuming when performed with DFT codes. Therefore for big structures we will use LAMMPS, a non-*ab initio* force field code, to relax the structures and then use BigDFT to compute its total energy. This procedure may lead to some error; however for many structures, including the vacancy, it is very small, since an accurate parametrization of the LAMMPS potential has been performed using *ab initio* simulations. Oppositely, the consequences of ignoring the relaxation of the structure are of greater importance, leading to VFE differences that can even surpass 1 eV from its true value [21].

Before addressing the calculations it is important to check the state of the art of this topic. VFE values for different elements were already published some decades ago, before any *ab initio* method was computationally feasible (see, for example, [22]). The rapid development of DFT codes at the beginning of this century offered a new and more accurate method to compute VFE and improved the old results (see, for example, [6, 21, 23]). In some cases the energy differences were small but in many other cases the differences were quite big, even leading to different predictions for minimum energy structures. Thus, it became clear that truly *ab initio* calculations were indispensable in order to analyze and quantify point defects.

At present, VFE values are well known and established. Experimentally, the Tungsten VFE is found to be between 3.5 and 4.1 eV [24]. Computationally, Tungsten VFE values from DFT simulations range from approximatively 3.2 to 3.8 eV [6, 25–27]. In both cases the variation represents around a 15%, which is now calculated by dividing the lower by the higher value in the range since the

VFE is an energy difference. This encouraged us to further investigate the VFE and the effects that may cause such high variations.

As already commented in chapter 1, we must point out that in order to have a more complete study, the SIA formation energy would also need to be analyzed. Even though we already have some results regarding the latter point defect, these are not enough so that they can be included in this thesis and to be properly studied. We expect to have them soon and to include them in upcoming publications.

5.2 Simulations and analysis of results

In order to obtain the value of the VFE, two DFT simulations need to be performed: one for the perfect structure and one for the defect structure. The results of the first one have already been presented in section 4.2. Thus, we focus here on the ones for the defect. These days, the standard method to find the total energy of a structure with a point defect works as a "hybrid mode" between the k-points and Γ calculations. To understand why this was proposed, one needs to notice that the defect structure of a single vacancy is not purely periodic. Extracting a single atom from the system breaks the periodicity of the lattice, which makes it impossible to compute the total energy with the unit cell of the system and the k-points. As already commented in section 4.2, using the Γ-point it is also unfeasible to reach the convergence of the system. The solution that was proposed consists of enlarging the unit cell of the system, similar to the Γ case, but still using k-points. The consequences are the following: the unit cell is as large as possible, for example with 128 or 250 atoms, and a smaller number of k-points is used. In [6], for example, a unit cell of 128 Tungsten atoms and a $3 \times 3 \times 3$ k-points grid is used. Within this framework, the vacancy is usually located at the center of the unit cell for convenience and in order to accurately predict the total energy it is required that the vacancy does not interact with its periodic *image* created by the boundary conditions and exploited by the k-points. It is assumed that this is true, since the effect of a single atom is not supposed to reach distances further than the new unit cell. However, since no code has reached still larger system sizes, it has not been strictly proven.

The goal of this section is to reach system sizes close to the 128 unit cell with the $3 \times 3 \times 3$ k-points grid, however within the Γ-point approach. Since a cell of 128 atoms is formed of $4 \times 4 \times 4$ minimal cells, the aforementioned structure would

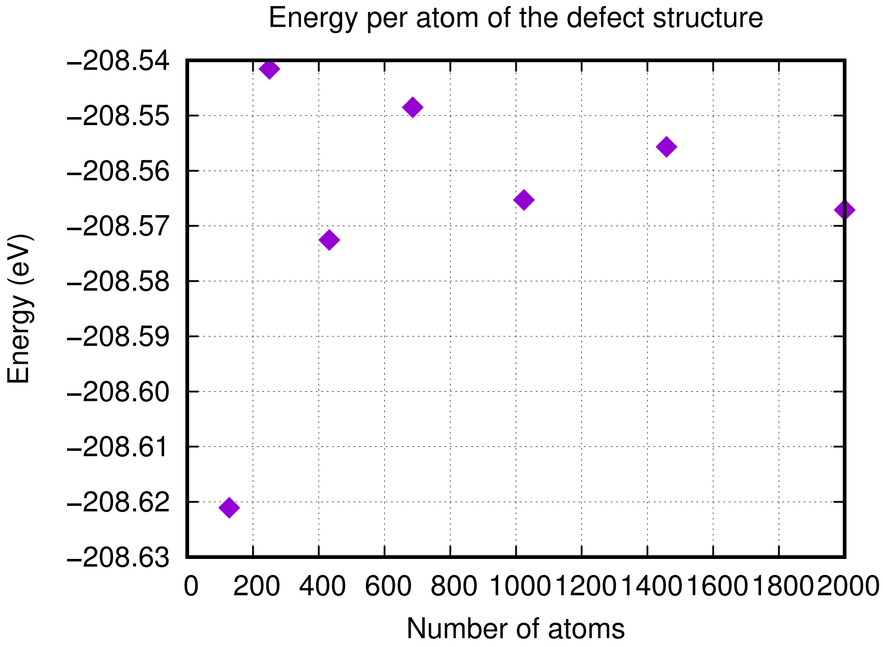


Figure 5.2: Energy per atom of the vacancy structure as a function of the system size, obtained with the linear version at Γ.

be equivalent to a single cell of 3456 atoms. In this thesis we will reach up to a $10 \times 10 \times 10$ unit cell (2000 atoms), which is not as big as desired but will already be meaningful. In figure 5.2 we present the energy per atom of the defect structures from 128 to 2000 relaxed by LAMMPS and calculated with the linear version of BigDFT. We observe that the behavior is very similar to the perfect structures presented in section 4.2, *i.e.* large oscillations for small sizes and convergence for larger ones. The energy difference between the biggest systems, namely 0.01 eV, represents also a 10% of that given by the smallest ones. This fact seems to suggest that the vacancy inside the defect structure does not affect the general behavior of the lattice and only adds a small perturbation, which becomes less apparent for the biggest sizes since the relative effect of the vacancy decreases due to the increasing number of atoms. Thus, it would not be necessary to use even larger system sizes. This can be checked in figure 5.3, where the energies of the perfect and defect structures are plotted together and their similar behavior becomes clear.

Nevertheless, these results are not conclusive. The proper quantity to analyze is the VFE and from equation (5.1) we see that it is calculated from energy differences. Thus, it is possible that even small fluctuations in the energy per atom lead

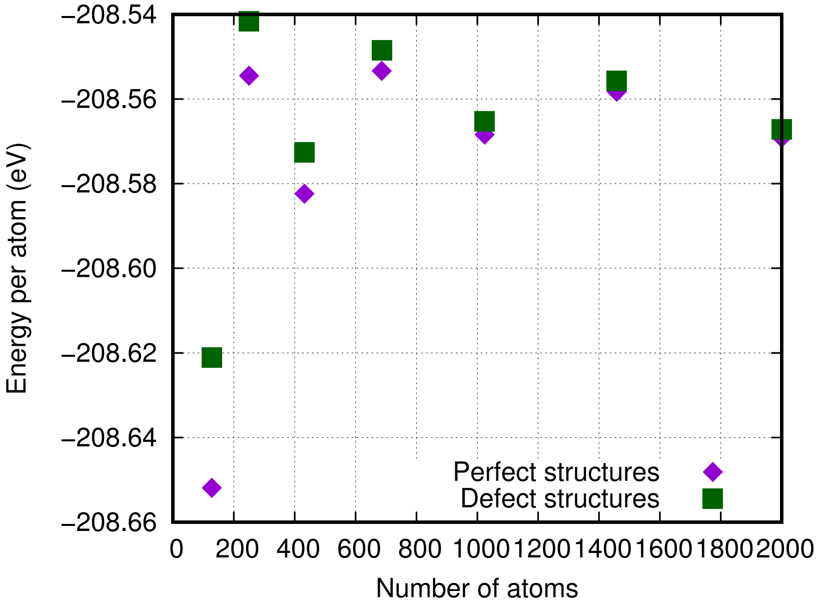


Figure 5.3: Energy per atom of the perfect and vacancy structure as a function of the system size, obtained with the linear version at Γ. The energy difference vanishes as the system size grows.

to relatively big oscillations in the VFE. In figure 5.4 we plot the VFE obtained using the results from sections 4.2 and 5.2. Indeed, VFE values for different sizes seem not to converge even for the 2000 atoms system. Even though most values are inside the accepted range given by reference values in section 5.1, we observe that the maximum variation of our VFE values as a function of the system size is approximatively 23%, while the variation for the biggest systems is 7%.

To have a better comprehension of the results, we plot in figure 5.5 the VFE obtained with the cubic version, LAMMPS and in grey the experimental accepted range. From the figure we can deduce:

- The fact that the cubic and linear version of BigDFT lead to the same values forms another evidence that the linear version works with large metallic systems.
- By analyzing the LAMMPS result, it is clear that non *ab initio* codes can not accurately predict VFE values. This manifest itself in the fact that its results are almost constant, independently of the system size. Thus, if the given value is correct it would only be justified by a previous parametrization

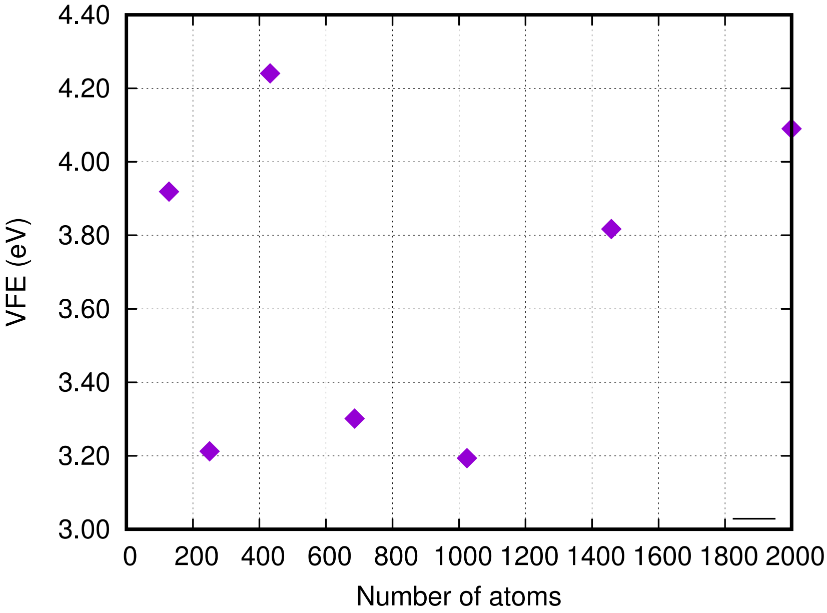


Figure 5.4: VFE of a Tungsten bcc lattice as a function of the system size. Values obtained with the linear version at Γ.

using DFT calculations for a particular system size and setup.

- The VFE for a Tungsten bcc lattice does not reach convergence even for 2000 atoms using Γ-only calculations.

It is clearly of big interest to expand and analyze the latter point, *i.e.* why values have not converged yet. The reason must be, obviously, that the system size is not big enough to converge. This size problem can manifest itself in two ways:

(i) In section 4.2 it has been shown that, when using the k-points method, the convergence is not completely reached until a system size of approximatively 2600-3000 atoms. Thus, it may happen that, although the total energy per atom seems to have converged, actually a bigger system is needed. This would apply not only for the defect structure but also for the perfect one.

(ii) The fact that the VFE has not yet converged for 2000 atoms could also be explained only by the defect structure. The presence of the vacancy in the cell introduces changes in its internal structure, modifying the position

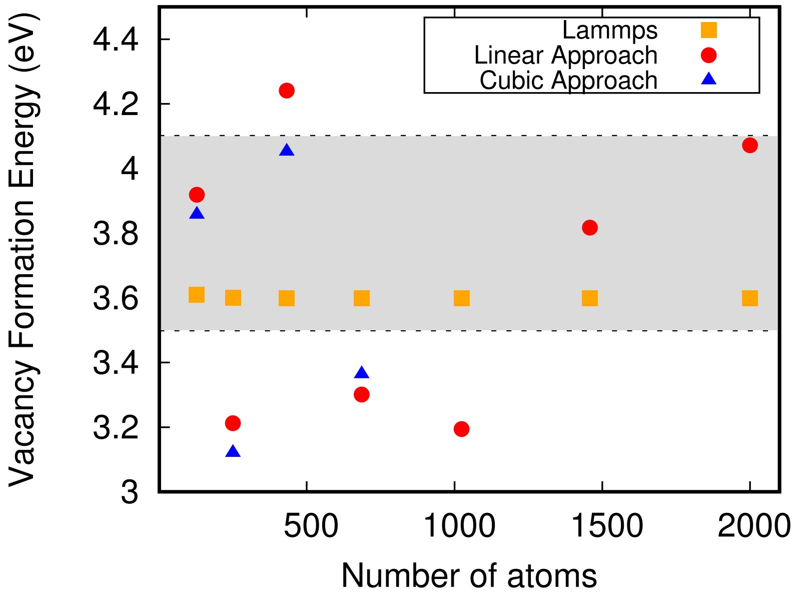


Figure 5.5: VFE of a Tungsten bcc lattice as a function of the system size and for different approaches. The region in grey represents the experimental uncertainty range.

> and charge of near neighbors. Thus, it is possible that the defect structure would inherently need a bigger cell than that used in standard methods. The observed convergence in reference calculations would then be explained by the fact that convergence is reached for a system different from the desired one, since the vacancy would see and interact with its periodic image.

The quickest way to find the answer is to compute the VFE for even bigger cells, at least until 3456 atoms. If convergence was reached and gave a value very similar to the standard method, the use of the latter would be completely justified. Oppositely, if a different value was obtained or convergence was not even reached, the possibility that the vacancy inherently needs bigger cells would gain strength. Unfortunately, due to lack of computational time, it has been impossible to compute those big systems up to now.

The lack of CPU-time needed to compute bigger structures encouraged us to come up with different approaches to gain insights regarding the effect introduced by the vacancy in the final energy value of the simulations. These approaches, briefly described at the beginning of this chapter, will be presented in the following, and whether they require simulations with system sizes similar to those presented up to now, thus feasible in terms of CPU-time, or a postprocessing of particular quantities, such as the atomic monopole, of simulations already presented.

5.2.1 Multidefects analysis

The first alternative methodology is based on the idea of simulating the k-points approach by means of the Γ-only calculation. We recall that the standard method for computing point defects right now consists of taking a large unit cell (usually with 128 atoms) and then add a relatively small k-points grid (usually $3 \times 3 \times 3$). The corresponding real space cell consists of a big cell of 3456 atoms with 27 defects, each of them located at the center of each *repeated* unit cell. Since defects are supposed not to interact with each other, the vacancy formation energy of a single defect would also be valid within this approach, once the total VFE is divided by the total number of vacancies.

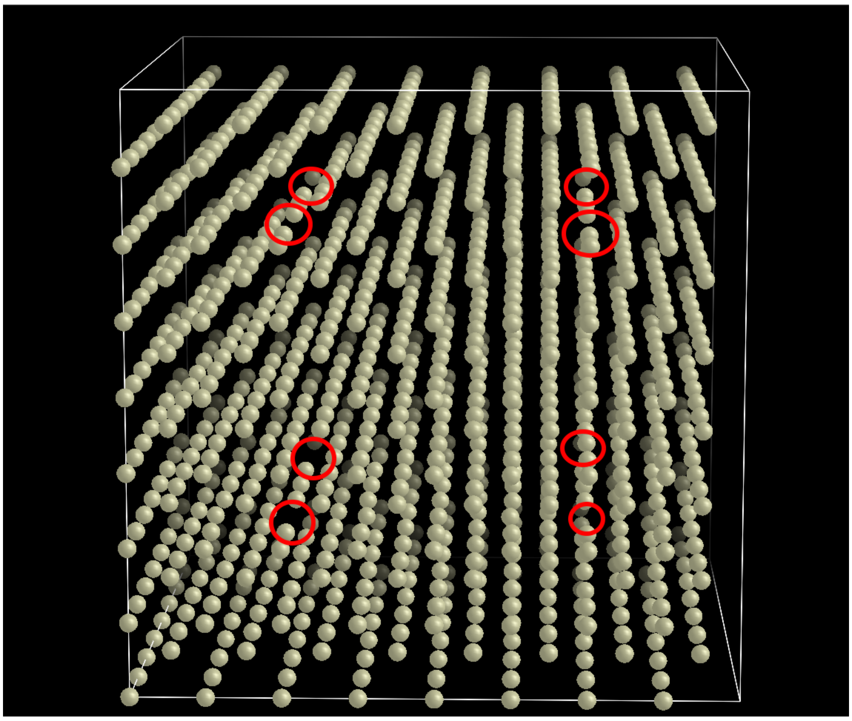


Figure 5.6: Visualization of a Tungsten bcc lattice with 1016 (1024-8) atoms corresponding to 8 vacancies. This system is equivalent to a unit cell containing 128 atoms and a single vacancy with $2 \times 2 \times 2$ k-points.

To check this fact, we have performed a new analysis, consisting of building the aforementioned equivalent structure in real space and performing the calculations at Γ. Due to the already commented limitations with respect to the CPU-time, we have restricted our simulation to a unit cell of 128 atoms and a single vacancy with a grid of $2 \times 2 \times 2$ k-points. In figure 5.6 we show the corresponding Γ-only structure, *i.e.* 8 repeated 128 single-defect cells. It is worth recalling the three types of simulations that we will analyze in the following,

(i) Standard method: cell of 128 atoms and a single vacancy with a grid of $2 \times 2 \times 2$ k-points.

(ii) Γ-point and multiple defects: big cell containing 1024 atoms and 8 vacancies created to mimic the standard method and check for agreement or disagreement.

(iii) Γ-point and single defect: big cell containing 1024 atoms and a single vacancy. This has been the method used up to now in our calculations.

The results are presented in table 5.1. We notice that the standard method and the multi-defect approach give similar results, with a variation of approximatively 2%. This is as expected, since the standard method and the multi-defect are supposed to be equivalent. However, the Γ-only simulation with the single defect has a larger difference, with a variation of 15%, which is a quite surprising result. The fact that the 1024 atoms system with a single vacancy differs by approximatively 0.5 eV seems to suggest that there is indeed some interaction between the vacancies.

	vacancy formation energy per defect (eV)
128 atoms with a single vacancy and a $2 \times 2 \times 2$ k-points grid	3.673
1024 atoms at Γ-only and 8 vacancies	3.750
1024 atoms at Γ-only and single vacancy	3.194

Table 5.1: Vacancy formation energy per defect for the three different methods.

Nevertheless, we must emphasize that these are preliminary results that must be further validated. First, we are only dealing with one system size, which could be not representative. Second, as already commented earlier, a system size of 1024 is not large enough to be in the converged range, thus some fluctuations between different methods may arise. Therefore, we went one step further and performed the same simulations but changing the 128 atoms unit cell by a 250 atoms unit cell. This transformation leads to the following systems:

(i) Standard method: cell of 250 atoms and a single vacancy with a grid of $2 \times 2 \times 2$ k-points.

(ii) Γ-point and multiple defects: big cell containing 2000 atoms and 8 vacancies.

(iii) Γ-point and single defect: big cell containing 2000 atoms and a single vacancy.

The results are presented in table 5.2. This the time results are not so easy to interpret. We notice that for the Γ calculations, there exists a difference of approximatively 0.3 eV between the single defect and multi-defect, corresponding to a variation of 9%. This still seems to suggest that there exists an interaction between the vacancies, leading to different values of the VFE. Furthermore, its difference is smaller than for the 128 atoms system, which agrees with the fact that vacancies in bigger cells are farther away from each other and therefore their interaction is smaller. However, the k-points method seems to be closer to the Γ calculation with a single vacancy than to the multi-defects, even though it should be equivalent to the latter. This last results, which seem contradictory, force us to interpret the results with great care. In particular, these somehow contradictory results could again explained by the fact that the system size is not big enough to be converged.

	vacancy formation energy per defect (eV)
250 atoms with a single vacancy and a $2 \times 2 \times 2$ k-points grid	3.972
2000 atoms at Γ-only and 8 vacancies	3.734
2000 atoms at Γ-only and single vacancy	4.090

Table 5.2: Vacancy formation energy per defect for the three different methods.

5.2.2 Charge analysis

The lack of a clear evidence of the interaction between vacancies regarding the approach just presented made us come up with another one. The second approach

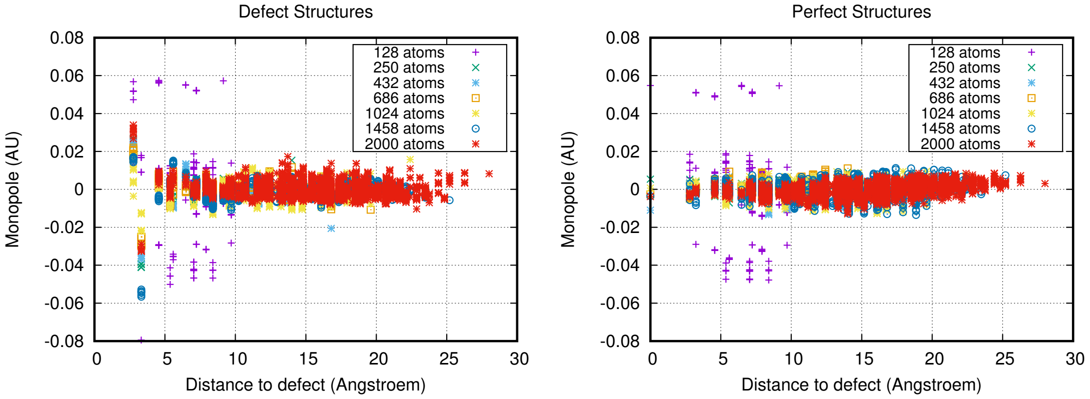


Figure 5.7: Monopole distribution as a function of the distance to the vacancy.

consists of studying how the presence of the vacancy affects the charge of particles around it. If the difference of the charge distribution between the perfect and defect structures is appreciable and reaches distances of the order of the whole cell, our hypothesis that vacancies impose the use of large cells would be reinforced.

For this purpose the output files of the simulations at Γ with the single vacancy and the linear version are used. Although no extra DFT simulations need to be performed, a considerable work of postprocessing the data files need to be carried out. In particular, we perform an atomic multipole analysis. Even though there exist different atomic multipole moments related to the charge density in an atom, here we will focus only on the monopole. Further work could include the study of higher order moments, such as the dipole or quadrupole. In order to study the charge distribution around the defect, the monopole of each atom has been plotted as a function of its distance to the vacancy position, obtaining a distribution of the monopole moment. In figure 5.7 the distribution has been plotted for defect and perfect structures of all system sizes. In figure 5.8 the same plot is shown only for the system with 2000 atoms, in order to have a better visualization.

From both figures it is clear that the charge distribution is influenced by the presence of the vacancy. However, it is not clear to which extent this may affect the final values of the calculation. In order to have a rigorous proof of the vacancy interactions, it would be required that a difference in the charge distribution is found for the largest distances. Looking again at figures 5.7 and 5.8 we notice that there is a very tiny difference at the largest distances. Thus, a more detailed analysis would be needed.

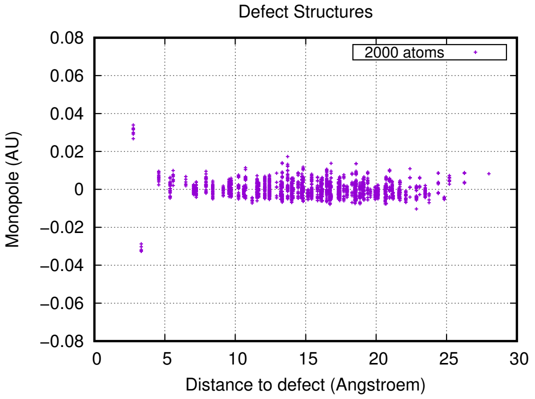


Figure 5.8: Monopole distribution as a function of the distance to the vacancy only for the system with 2000 atoms.

For this purpose, a direct visualization of the system could be helpful. Instead of a distribution of the monopole as a function of the distance to the defect, we show in figures 5.9 and 5.10 a representation of the system, where each atom is colored as a function of its monopole charge. Atoms in blue represent positive charges, atoms in red represent negative charges, while atoms in white represent a zero charge.

If we focus on the system with 128 atoms (figure 5.9), we observe that the perfect structure present a regular distribution. It is worth noticing that in the perfect periodic system the monopole should be neutral for all atoms. Due to the small system size, this effect is not observed, leading to small fluctuations in the monopoles charge. In the defect structure, the presence of the vacancy breaks the regularity and modifies the charge of some atoms in the cell, especially in the nearest neighbors of the defect. From this, we can infer that

(i) In a 128 atoms cell the vacancy modifies the internal monopole distribution appreciably. Thus, its presence could be affecting the interaction with neighboring periodic cells.

(ii) The fact that the perfect structure's atomic monopoles are not neutral show that the calculation is not in the converged range, which agrees with the fluctuations in the energy for systems of this size presented in section 4.2.

If we focus on the system with 2000 atoms (figure 5.10), we observe that the increase of the system size has allowed the atomic monopole of the perfect structure

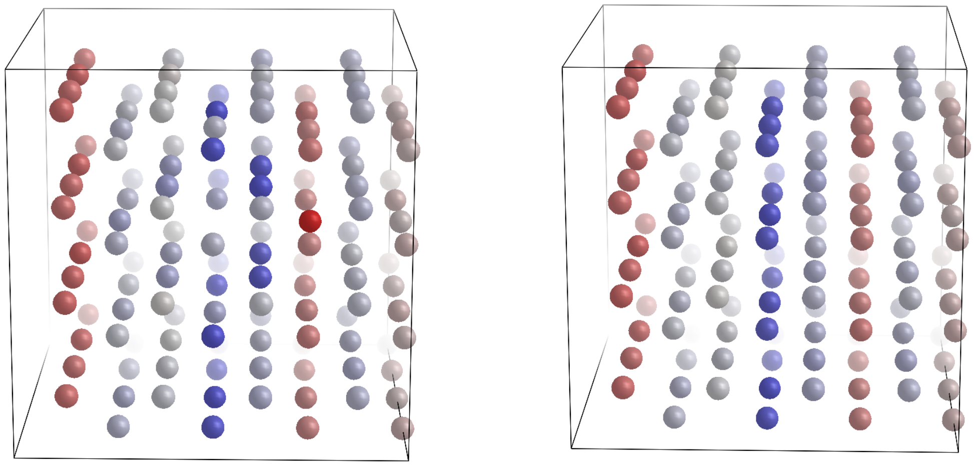


Figure 5.9: Visualization of the cell with 128 atoms with a colored representation of the charges: blue represents large positive charges, red represents large negative charges and withe stands for zero charges. Left: defect structure. Right: perfect structure.

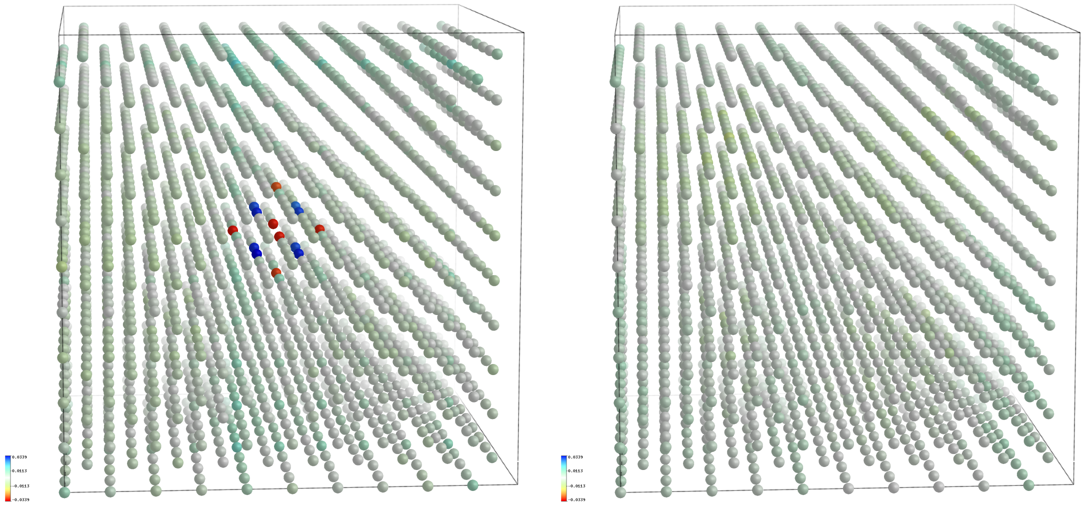


Figure 5.10: Visualization of the cell with 2000 atoms with a colored representation of the charges: blue represents large positive charges, red represents large negative charges and withe stands for zero charges. Left: defect structure. Right: perfect structure.

to become neutral, as expected. Also the defect structure shows now the expected result: only the nearest neighbors of the vacancy are affected, creating a structured monopole pattern. From this we can infer:

(i) In a 2000 atoms cell the defect only modifies the monopole of the vacancy's nearest neighbors. Thus, its presence should not interact with neighboring periodic cells.

(ii) The fact that the perfect structure's atomic monopoles are neutral, agrees with the better convergence of the energy for the biggest structures.

Summarizing, the results obtained with the charge analysis seem to suggest that indeed the presence of the vacancy may affect the accuracy of the simulations using unit cells of 128 atoms. This can be observed in figures 5.9 and 5.10: the smallest cell shows an unstructured monopole pattern, modifying the charge not only of its nearest neighbors but also of farther ones. Oppositely, the largest cell show a very structured monopole pattern around the vacancy, only affecting its nearest neighbors.

6 Conclusions

The main goals of this work were: (*i*) providing a sufficient test to validate the linear scaling version of the BigDFT code in large metallic systems and (*ii*) use the latter version of the code to perform electronic structure calculations on fusion materials with relevance in the field. For both purposes we have chosen Tungsten, mainly due to its big importance in the fusion materials area and possible future usage in upcoming nuclear fusion reactors. In particular, the Tungsten structure consisted in a pristine bcc lattice.

In order to validate the linear scaling version of BigDFT in large metallic systems, we have first performed a wide analysis of the simulation parameters set. In particular we have obtained the following optimal parameter values: $\ell_{\mathrm{lat}} = 3.2337$ Å and `hgrid` = 0.38 Bohr. To test the performance of the linear scaling DFT we have focused on the accuracy and the scalability of the code. Regarding the accuracy, in section 4.2.2, we have first presented the results obtained in the energy calculations of the bcc Tungsten structure for both the traditional cubic and linear version at Γ-only, as a function of the system size. The results have clearly shown that both version agree, providing virtually identical values. As an exception, the simulation with the smallest system size, namely 128 atoms, has led to a small difference between both approaches. Nevertheless, this is only due to a too small system size and not to a wrong physical behavior in any of the two approaches. The accuracy of the linear scaling version at Γ-only has also been compared with the cubic version with a Monkhorst-Pack k-points grid. To perform this comparison, the linear Γ-only system has been built with a fixed number of unit cells, equivalent to the number of k-points used in the cubic one. Then, the energy of both systems as a function of the lattice parameter has been compared, resulting in another proof of the validity of the linear version thanks to the agreement at the values.

Concerning the scalability of the code, we have restricted ourselves to the usage of the `diag` solver only. The latter solver has a very low prefactor with the price that it actually scales cubicly with the system size. It is important to recall that the solver scalability is completely independent to the scalability of the physical description of the system, which is completely linear in the linear version of BigDFT. The CPU-time spent in each simulation as a function of the system size has been presented and discussed in section 4.2.3. We have been able to check that (*i*) the linear description of the system agrees with the linear CPU-time of the simulations for the `diag` solver until approximatively 1024, where the linear scalability is lost due to the solver contribution. (*ii*) If the CPU-time consumed by the `diag` solver is subtracted from the total CPU-time we obtain an almost linear scalability up to 2000 atoms. (*iii*) The crossover between the cubic and linear version with the `diag` solver and for a Tungsten bcc lattice takes place at a system size of around 150 atoms. We remark that in order to be in a completely linear scaling regime, the `FOE` solver must be used. The latter solver presents a completely linear scaling with the price of a large prefactor.

Once the scalability of the system has been validated, the thesis has focused on the study of the formation energy of a particular point defect, namely the single vacancy, in Tungsten. In order to obtain values for the VFE we have first simulated the defect structures. The relaxation of the structures have been performed using the force field code LAMMPS, while the *ab initio* energy calculation has been obtained with the linear version of BigDFT at Γ-only. The results show that the VFE values do not reach convergences in our system size range, *i.e.* from 128 to 2000 atoms. This fact was not surprising, since bigger system sizes are required to reach convergence. This was already predicted by the k-points study in section 4.1.1 and also stated in the literature. Due to the cubic scalability of `diag`, structures larger that the 2000 atoms one have been unfeasible. Thus two alternative approaches to evaluate the effect that the vacancy introduces to the simulations have been performed.

The standard method to compute the VFE uses relatively large unit cells containing a single vacancy, along with a k-point grid to reach the aforementioned big system sizes. The usage of the Monkhorst-Pack k-points grid is actually computing a big system made up with copies of the unit cell, thus containing multiple defects. Our aim was then to check if there was any interaction between vacancies that could lead to a not so accurate results of the VFE. In section 5.2.1 we have performed simulations equivalent to those using the k-point grid within the

Γ-only approach. The results have suggested that indeed there seems to be an interaction between vacancies when the 128 unit cell system with a k-point grid is used. Nevertheless, further studies need to be carried out, since the small system size could be causing the difference in the VFE for the different approaches. In section 5.2.2 an atomic monopole analysis of the systems has been performed. The results seems to agree with those in section 5.2.1: a single vacancy in the center of a 128 unit cell may change monopole values of atoms at the edge of the unit cell, which could lead to an interaction between vacancies in neighboring unit cells. This effect vanishes with the 2000 atoms unit cell.

In summary, we have proven that the linear scaling version of BigDFT may become a really powerful tool to be used in large non-periodic system, as well as periodic ones that present specific defects, such as point defects or even larger defects such as vacancy clusters, where the usage of large unit cells would be compulsory.

Finally, we reiterate that we are already working with simulations using the FOE solver, as well as with structures presenting the SIA point defect. Unfortunately, the results obtained up to now are not enough so that they could be presented in this thesis. We hope that we obtain those results briefly and present them in upcoming publications.

7 Bibliography

[1] Martin R., *Electronic Structure: Basic Theory and Practical Methods*. Cambridge, UK: Cambridge University Press (2004).

[2] Marzari N., *Ab-initio Molecular Dynamics for Metallic Systems*. PhD Thesis (1996).

[3] Kohn W., *Density functional and density matrix method scaling linearly with the number of atoms*, Phys. Rev. Lett. **76** 3168 (1996).

[4] BigDFT web page, `http://bigdft.org/`.

[5] Hirai T. *et al.*, *Use of tungsten material for the ITER divertor*, Nuclear Materials and Energy, **9**, 616 (2016).

[6] Nguyen-Manh D.,Horsfield A. P. and Dudarev S., *Self-interstitial atom defects in bcc transition metals: Group-specific trends*, Phys. Rev. B **73** 020101(R) (2006).

[7] Fock V., *Näherungsmethode zur Lösung des quanten-mechanischen Mehrkörperprobleme*, Z. Phys. **61** 126 (1930).

[8] Mohr S., *Fast and accurate electronic structure methods: large systems and applications to boron-carbon heterofullerenes*. PhD Thesis (2013).

[9] Kohn W. and Hohenberg P., *Inhomogeneous electron gas*, Phys. Rev., **136**, B864 (1964).

[10] Kohn W. and Sham L. J., *Self-consistent equations including exchange and correlation effects*, Phys. Rev. **140** A113 (1965).

[11] Jacques Des Cloizeaux, *Energy Bands and Projection Operators in a Crystal: Analytic and Asymptotic Properties*, Phys. Rev. **135** A685, (1964).

[12] Genovese L. *et al.*, *Daubechies wavelets as a basis set for density functional pseudopotential calculations*, The Journal of Chem. Phys. **129** 014109 (2008).

[13] Mohr S. *et al.*, *Daubechies wavelets for linear scaling density functional theory*, The Journal of Chem. Phys. **140** 204110 (2014).

[14] Mohr S. *et al.*, *Accurate and efficient linear scaling DFT calculations with universal applicability*, Phys. Chem. Chem. Phys. **17** 31360 (2015).

[15] Daubechies I. *et al.*, *Ten lectures on wavelets*, volume 61, SIAM (1992).

[16] Gonze X. *et al.*, *Recent developments in the ABINIT software package*, Comp. Phys. Comm. **205**, 106 (2016).
`http://www.abinit.org`.

[17] Soler J. M. *et al.*, *The SIESTA method for ab-initio order-N materials simulation*, J. Phys.: Condens. Matt. **14**, 2745 (2002).
`https://www.icmab.es/siesta`.

[18] S. Plimpton, *Fast Parallel Algorithms for Short-Range Molecular Dynamics*, J. Comp. Phys., **117**, 1 (1995).
`http://lammps.sandia.gov`.

[19] Causey R. and Venhaus T., *The use of tungsten in fusion reactors: a review of the hydrogen retention and migration properties*, Physica Scripta, **2001** T94 (2001).

[20] Dudarev S., *Density Functional Theory Models for Radiation Damage*, Annu. Rev. Mater. Res., **43**, 35 (2013).

[21] Soderlind P. *et al.*, *First-principles formation energies of monovacancies in bcc transition metals*, Phys. Rev. B **61** 2579 (2000).

[22] Gillan M., *Calculation of the vacancy formation energy in Aluminium*, J. Phys.: Condens. Matt. **1**, 689 (1989).

[23] Carling K. *et al.*, *Vacancies in Metals: From First-Principles Calculations to Experimental Data*, Phys. Rev. Lett. **85** 3862 (2000).

[24] Ehrhart P. *et al.*, *Atomic Defects in Metals*, edited by Ullmaier H., Landolt-Börnstein, New Series, Group III, Vol. 25 (Springer-Verlag, Berlin, 1991).

[25] Satta A., Willaime F. and Stefano de Gironcoli, *Vacancy self-diffusion parameters in tungsten: Finite electron-temperature LDA calculations*, Phys. Rev. B **57** 18 (1998).

[26] Söderlind P., Yang L. and Moriarty J., *First-principles formation energies of monovacancies in bcc transition metals*, Phys. Rev. B **61** 4 (2000).

[27] Ventelon L. *et al.*, *Ab initio investigation of radiation defects in tungsten: Structure of self-interstitials and specificity of di-vacancies compared to other bcc transition metals* Journal of Nuclear Materials, **425** 16 (2012).

[28] Knaster J., Moeslang A. and Muroga T., *Materials research for fusion*, Nature Physics, **12** 424 (2016).

[29] Dudarev S. *et al.*, *The EU programme for modelling radiation effects in fusion reactor materials: An overview of recent advances and future goals*, Journal of Nuclear Materials **386** 1 (2009).

[30] Derlet P., Nguyen-Manh D. and Dudarev S., *Multiscale modeling of crowdion and vacancy defects in body-centered-cubic transition metals* Phys. Rev. B **76**, 054107 (2007).

[31] Becquart C. *et al.*, *Microstructural evolution of irradiated tungsten: Ab initio parameterisation of an OKMC model*, Journal of Nuclear Materials **403** 75 (2010).