UNIVERSITY OF CALIFORNIA, SAN DIEGO

Electronic Structure Models: Solution Theory, Linear Scaling Methods, and Stability Analysis

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Physics with a Specialization in Computational Science

by

Houdong Hu

Committee in charge:

Professor Michael Holst, Chair
Professor Randolph Bank
Professor Julius Kuti
Professor Michael Norman
Professor John Weare

2014

The dissertation of Houdong Hu is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____

Chair

University of California, San Diego

2014

DEDICATION

I dedicate to my wife and my parents, whose love, encouragement and prays of day and night make me able to pursue my dream.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my thesis advisor, Professor Michael Holst for giving me the opportunity to conduct cutting edge research in his group. None of my accomplishment would be possible without his guidance, scrutiny, and encouragement. He has provided a research environment that is both intellectually challenging and uplifting, and I have become a truly independent scientist because of that.

Special acknowledgment must be given to Professor John Weare for sharing his insight and experience, and the numerical part of my thesis can't get finished without his help. I would like to thank Duo Song and Ying Chen for sharing their expertise on the computational chemistry. The assistance of Dr. Yunrong Zhu is indispensable for helping me on the part of mathematical proofs. I would also like to thank the other members of my theis committee, Professor Julius Kuti, Professor Michael L. Norman and Professor Randolph E. Bank for their continuing guidance and support.

Chapter 3, is currently begin prepared for submission for publication of the material. Holst, M., Hu, H., and Zhu, Y. The dissertation author was the primary investigator and author of this material.

Chapter 4, is currently begin prepared for submission for publication of the material. Hu, H., Song, D., Weare, J., and Holst, M. The dissertation author was the primary investigator and author of this material.

Chapter 5, is currently begin prepared for submission for publication of the material. Hu, H., Marzuola, J., Lu, J., Song, D., Weare, J., and Holst, M. The dissertation author was the primary investigator and author of this material.

Appendix A, in full, is a reprint of the material as it appears in Applied Physics Letter 2010. Hu, H., Ma, C., and Liu Z. The dissertation author was the primary investigator and author of this material.

VITA

2008	B. S. in Physics, University of Science and Technology of China
2010	M. S. in Physics, University of California, San Diego
2014	Ph. D. in Physics, University of California, San Diego

PUBLICATIONS

Holst, M., Hu, H., and Zhu, Y. "A mathematical framework of models in calculating electronic structures of materials", to be submitted, 2014 [67].

Hu, H., Song, D., Weare, J., and Holst, M. "A linear scaling finite element solver for Hartree-Fock", to be submitted, 2014 [71].

Hu, H., Marzuola, J., Lu, J., Song, D., Weare, J., and Holst, M. "A systematic way to study stability conditions of models in calculating electronic structures of materials", in preparation, 2014 [72].

Hu, H., Ma, C., and Liu Z. "Plasmonic dark field microscopy", *Applied Physics Letters*, 2010, 96, 11 [70].

ABSTRACT OF THE DISSERTATION

Electronic Structure Models: Solution Theory, Linear Scaling Methods, and Stability Analysis

by

Houdong Hu

Doctor of Philosophy in Physics with a Specialization in Computational Science

University of California, San Diego, 2014

Professor Michael Holst, Chair

Various approaches have been developed to investigate materials in the last century. Quantum methods, solving the Schrödinger equation with different types of approximations, have best accuracy comparing with other methods. The most commonly used quantum methods are wavefunction method and density functional method. Wavefunction method is based on obtaining the wavefunction of the system. The simplest wavefunction method is Hartree-Fock (HF) theory. Density functional theory (DFT) considers electronic density as the basic variable instead of the wave function. Density functional theory offers a practical computational scheme, which is similar to Hartree-Fock method.

Although Hartree-Fock and density functional theory have been studied many

years, only a small number of papers discuss the mathematical properties of these models. We generalize the work of Suryanarayana et al. on local spin density approximation model, and developed finite-element formulations for Hartree-Fock and density functional theory models, including restricted Hartree-Fock, unrestricted Hartree-Fock, local spin density approximation, generalized gradient approximation, meta generalized gradient approximation and more generalized density functional theory models. We prove the well-posedness, and the existence of minimizers for these models in a finite calculation domain for both the all-electron problem and pseudopotential approximations. We also established the convergence of the finite-element approximation, and the convergence of finite element approximation with numerical quadratures by Γ-convergence for both the all-electron problem and pseudopotential approximations. It will be useful for the development of studying Hartree-Fock and density functional theory models numerically by finite element method. In Chapter 3, we present the mathematical proofs of the mathematical properties of these models.

After well-defined models are proposed, numerical methods are needed to discretize and solve these problems. Much effort has been devoted to develop scalable real-space methods over the past decade , such as finite difference method, wavelet method, and finite element method. Difficulties arise when finite element method is used to solve a model with exact exchange energy, such as Hartree-Fock model. In order to handle exact exchange operator, we successfully proposed and numerically implemented a linear time cost and memory cost method with finite element bases. Multigrid method is used in our algorithm in order to achieve linear scaling. A variety of numerical experiments for atoms and molecules demonstrate reliable precision and speed. Our method could be generalized to density functional theory and hybrid models in a straightforward way. Each step in our self-consistent solver is well defined, and this makes the parallelization of our solver is feasible. In Chapter 4, we present the algorithms and numerical results of our linear scaling method.

In Hartree-Fock and density functional theory solvers, molecular orbitals and orbital wavefunctions are solved by the Euler-Lagrange equation of the total energy functional. In order to establish whether that extremum is a maximizer, minimizer, or saddle point, the second functional derivative must be analyzed. We successfully proposed

and numerically implemented a systematic way to study stability condition of Hartree-Fock and density functional theory models. We generated a few bifurcation figures for Hartree-Fock and density functional theory. Hessian analysis have also been conducted on extremums of Hartree-Fock and density functional theory models. A weakly bounded state is found when the nuclear charge of the two-electron atom gets really close to 1. It is a powerful tool to understand the stability of different solutions of Hartree-Fock and density functional theory models. In Chapter 5, we show a systematic way of stability study by Hessian matrix analysis.

Chapter 1

Introduction

Interdisciplinary research is research activities that applies knowledge and methodology from more than one academic discipline to solve an essential problem. In recent decades, the growth of science and technology has prompted scientists to address complex problems and create new methods by crossing boundaries with deep knowledge from different perspectives. A glance across today's hot topics reveals how many are interdisciplinary: neuroscience, bioinformatics, nanotechnology, quantum information, etc. All those require the contributions from two or more disciplines [108].

Computational science is an interdisciplinary approach to solve scientific problems that use knowledge and methods from applied discipline, mathematics and computer science. Computational science takes advantage of not only the improvements in computer hardware, but probably more importantly, the improvements in computer algorithms and mathematical techniques. Computational models in an applied discipline are constructed based on theory and experimental results. Computation is conducted on theses models and allows us to make predictions of experimental results. We could use computational science to predict what cannot be tested in laboratory, such as models to explain the origin of our universe. Computational science can also be used to give people some initial predictions to reduce the number of experiments, such as computational drug design [108].

Figure 1.1 is straightforward for people to understand the concept of computational science. An applied discipline in Figure 1.1 is physics. Many fundamental physics models, such as heat conduction, hydrodynamics, electrodynamics and quan-

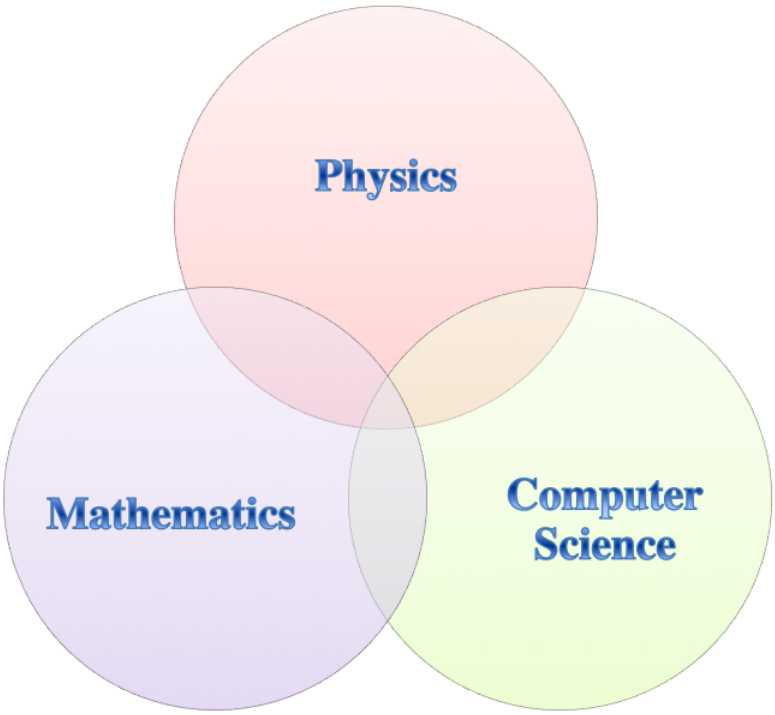

Figure 1.1: Discription of Computational Science

tum physics, are partial differential equations. After deriving the fundamental models, people need mathematical techniques to analyze the properties of them. Numerical approximations and algorithms are provided to discrete and solve partial differential equations, where we need to conduct approximation analysis and convergence analysis. We need to keep comparing the predicted results with current experiment results, which is a motivation to revise an old model or develop a new model if predicted results are not consistent with experimental results.

Computational physics is broad and there is a corresponding computational branch for almost every major field in physics, for example computational astrophysics, computational fluid dynamics, computational solid mechanics, computational particle physics, computational electrodynamics, etc. We feel interested in computational solid state physics, which is a very important field related to material science. People also call computational solid state physics as computational chemistry.

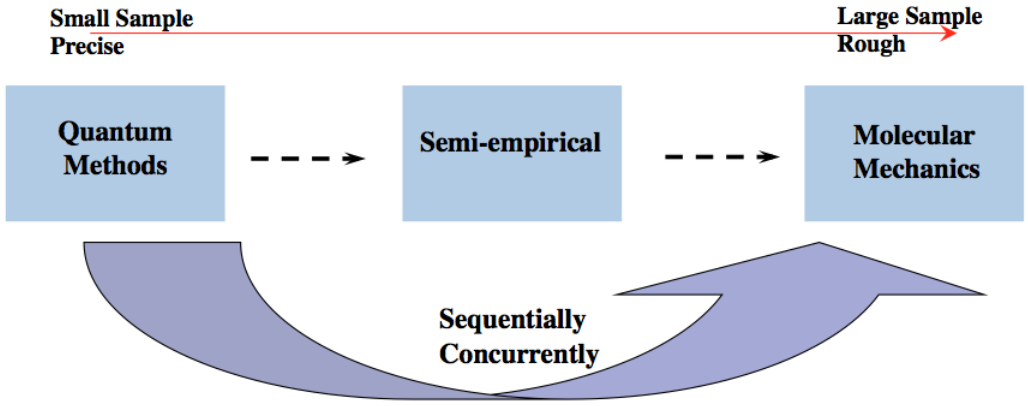

Figure 1.2: Methods of Computational Chemsitry

Currently, there are three major ways to investigate materials as shown in Figure 1.2. Quantum methods, derived from the quantum mechanics, use different types of approximations to solve the Schrödinger equation. Quantum methods can only solve small-sized systems as a result of expensive computational cost; however, it has best accuracy comparing with other methods. In order to allow for the simulations of systems of larger size, semi-empirical methods are proposed with further approximations, where some parameters are neglected or given by experimental results. Semi-empirical methods demand less computational cost than quantum methods, while they are less rigorous than quantum methods. Finally, molecular mechanics uses classical physics to explain and interpret the behaviors of molecules. Molecular mechanics can solve very large systems with the least computational cost [85].

In order to compute large-scale systems with precise results, we could combine these three methods sequentially or in parallel. The parameters of a coarse model could be precomputed on the basis of quantum methods applied on subsystems, such as, fitting of pseudopotentials for large-scale calculations. Also, we can divide the system into pieces that are modeled at different levels in parallel, as for quantum mechanics/molecular mechanics calculations. Our research is mainly on the basis of quantum methods.

Figure 1.3 gives us an overview of quantum methods. The most commonly used quantum methods are the wavefunction method [129] and density functional method [66, 81]. Wavefunction method is based on obtaining the wavefunction of the system.

Density functional method studies the properties of the system through its electronic density.

The simplest wavefunction method is Hartree-Fock (HF) theory [129]. There are two types of Hartree-Fock theory: restricted Hartree-Fock (RHF) and unrestricted Hartree-Fock (UHF). Restricted Hartree-Fock is used for closed shell atoms or molecules, which have an even number of electrons. Each spatial orbital is occupied by two electrons, one spin-up and one spin-down. Unrestricted Hartree-Fock is used for atoms or molecules with an even or odd number of electrons. It assumes spatial orbitals for spin-up and spin-down electrons are different.

There have been a large number of methods developed to improve Hartree-Fock results, which are called post-Hartree-Fock (Post-HF) methods. Several common post-Hartree-Fock methods are configuration interaction (CI) [139], coupled cluster (CC) [29] and the second order Moller-Plesset perturbation theory (MP2) [103].

Configuration interaction is a variational method that accounts for the correlation energy using a variational wavefunction, which is a linear combination of determinants or configuration state functions built from spin orbitals. If the expansion includes all possible configurations of the appropriate symmetry, it is a full configuration interaction (FCI) [44] procedure.

Coupled cluster takes the molecular orbitals from Hartree-Fock theory to construct the multi-electron wavefunction, then uses the exponential cluster to account for electron correlation energy missing in Hartree-Fock calculation. The cluster operator is written in the form of a summation, where the first term is the operator of all single excitations, the second term is the operator of all double excitations and so forth. The coupled method including only the first and the second term is CCSD [138]. The most famous application of CCSD that provides an estimate of connected triples is CCSDT [138].

Moller-Plesset perturbation theory (MP) is a special case of Rayleigh-Schrödinger (RS) perturbation theory. In Rayleigh-Schrödinger perturbation theory, one considers an unperturbed Hamiltonian operator with a small perturbation added, which accounts for the missing electron correlation energy. Usually we use first order MP theory , second order MP theory, third order MP theory and fourth order MP theory.

Density functional theory (DFT) considers electronic density as the basic variable instead of the wave function. Density functional theory offers a practical computational scheme, which is similar to Hartree-Fock method. The most widely used approximations are the local density approximation (LDA) [81] and local spin density approximation (LSDA) [81], where the functionals depend only on the density at the coordinate where the functional is evaluated. Generalized gradient approximation (GGA) [118, 83, 10, 86, 123, 117, 119] are proposed, which take into account of the gradient of the density. Potentially more accurate meta generalized gradient approximation (Meta-GGA) [116, 156, 9, 33, 127, 121, 149] includes higher order derivative of the electron density, and it is a natural development from generalized gradient approximation.

Hybrid models [12, 120, 78, 145, 10, 86, 157, 11, 65] are combinations of density functional theory and Hartree-Fock theory, which generate accurate results and have been widely applied in computational chemistry.

We include semi-empirical quantum methods [85] in Figure 1.3, because semi-empirical quantum methods represent a bridge between quantum methods and molecular mechanics, and it tries to overcome the obstacles of both slow speed and low accuracy.

Figure 1.4 shows all the approximations made on quantum methods in order to solve the problem theoretically and numerically, and our research areas are marked as purple paint. The first approximation is made to avoid calculations on many-electron Hamiltonian. The approximate models, including Hartree-Fock method, density functional method and hybrid method, are proposed, which solve the effective one-electron Hamiltonian instead of many-electron Hamiltonian.

The singular nuclear charges could be approximated by Gaussian charges [38, 87, 38]. The nuclear potential could be smoothed by the pseudopotential approximation [63, 56, 58, 124]. The pseudopotetial approximation replaces the collective system of nuclei and core electrons with an effective, smoother potential, and reproduces well the true potential of the valence wavefunction in the interstitial region. This method significantly reduces the number of bases set needed in the calculation.

Approximations are also made during the discretization of partial differential equations with chosen basis sets. There is a trade-off between accuracy and calculation time, which depends on the size of basis sets. For a chosen basis set, we also need to

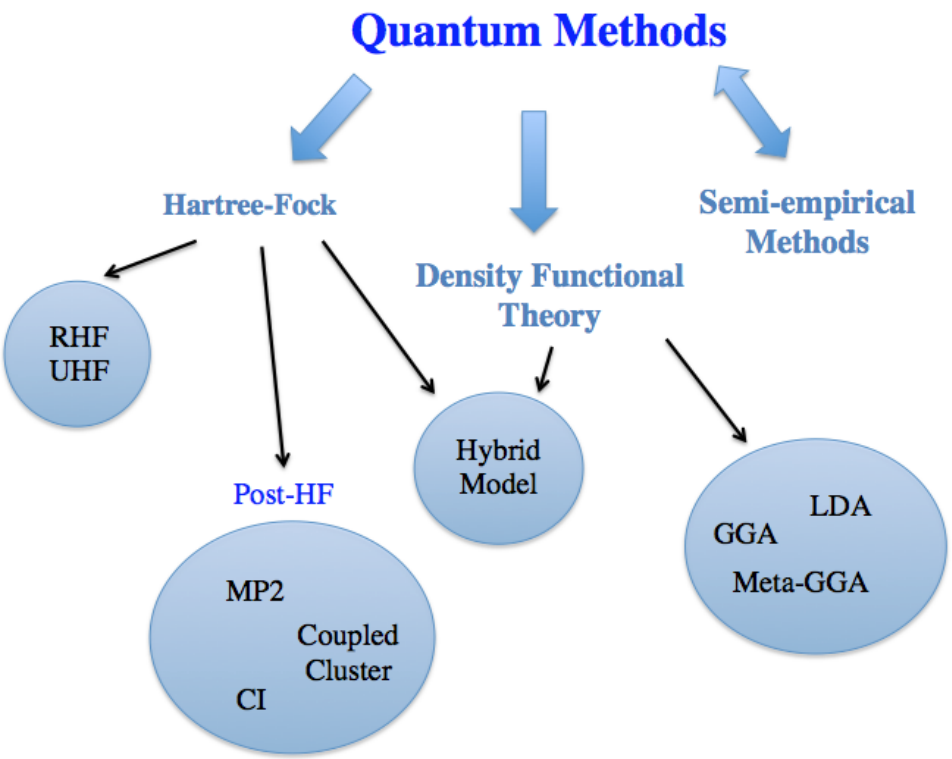

Figure 1.3: Quantum Methods

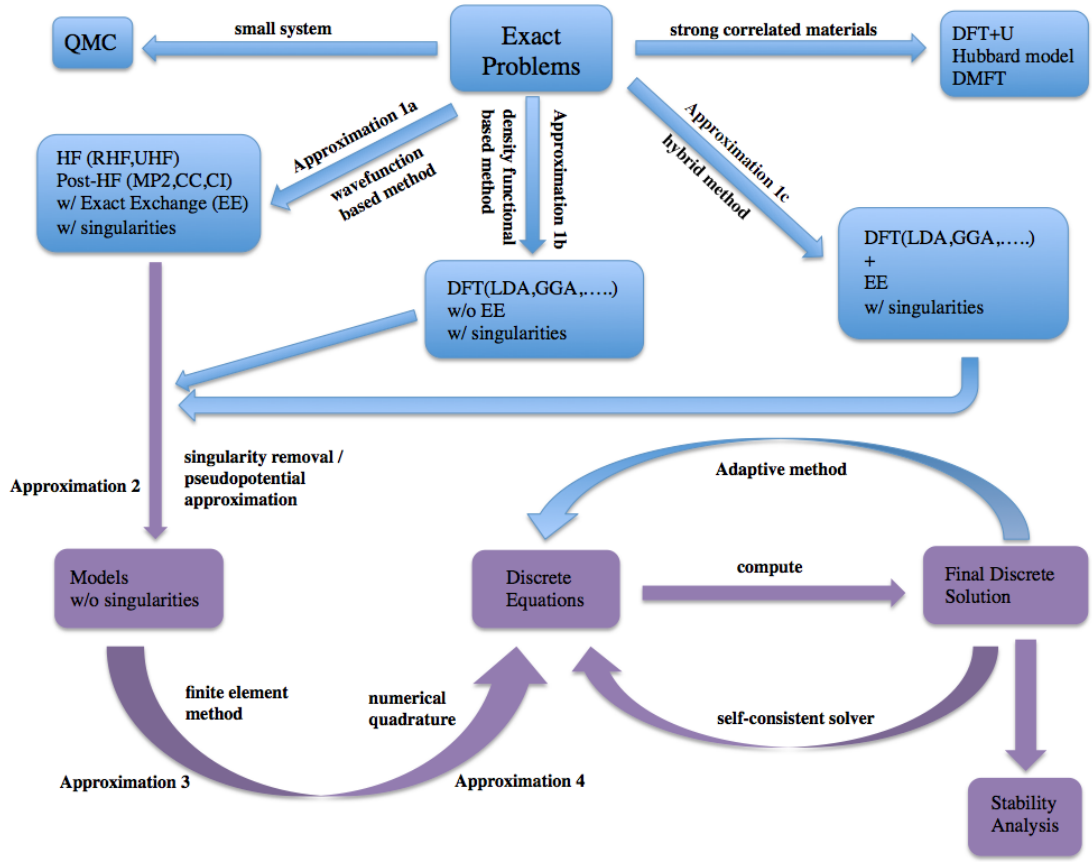

Figure 1.4: Theoretical and Numerical Approximations

choose numerical quadrature rules to assembly the stiffness matrix of partial differential equations. Certain order of numerical quadrature rules are needed in order to maintain the accuracy.

Although Hartree-Fock and density functional theory have been studied many years, only a small number of papers [92, 90, 3, 46, 84, 146] discuss the mathematical properties of these models, such as the existence of a minimizer, the convergence of finite element approximation, the convergence of finite element approximation with numerical quadratures and the convergence with pseudopotential approximation. In Chapter 3, we present the mathematical proofs of the mathematical properties of these models [67].

After well-defined physics models are proposed, numerical methods are needed to discretize and solve these problems. The development of systematically improvable

and scalable real-space techniques for electronic structure calculations has received significant attention over the past decade [8, 143, 26, 23, 48, 15, 141, 102]. Among the real-space methods, finite element method [160, 151, 152, 153, 113, 114, 115, 166, 22, 87, 146, 91, 36, 6, 105, 104, 96, 132, 27] has the advantages of significantly greater flexibility in the construction of the mesh and the scalability on parallel computing platforms. Difficulties arise when finite element method is used to solve a model with exact exchange energy, such as Hartree-Fock method. Because exact exchange energy operator is non-local and finite element method needs a large number of bases comparing with atomic-orbital type basis sets and Gaussian basis sets. The stiffness matrix of the exact exchange energy operator will be dense, and the time cost of solving the eigenvalue problem will become huge.

We developed a general linear scaling method with finite element bases to calculate electronic properties of materials, which could be applied to all kinds of Hartree-Fock, density functional theory and hybrid models [71]. Our linear scaling method will guarantee both the linear time cost and linear memory cost. In Chapter 4, we present the algorithms and numerical results of our linear scaling method.

In Hartree-Fock and density functional theory solvers, molecular orbitals and orbital wavefunctions are solved by the Euler-Lagrange equation of the total energy functional. In order to establish whether that extremum is a maximizer, minimizer, or saddle point, the second functional derivative must be analyzed. Few papers [74, 136, 110, 140] discussed the stability of Hartree-Fock and density functional theory results. In Chapter 5, we show a systematic way of stability study by Hessian matrix analysis [72].

The remainder of the thesis is organized as follows. In Chapter 2, the descriptions of the Hartree-Fock and density functional theory formulations are presented. In Chapter 3, we present the mathematical proofs on the existence of a minimizer, convergence of finite element approximation, convergence of finite element approximation with numerical quadratures and convergence of pseudopotential approximation for Hartree-Fock and density functional theory [67]. In Chapter 4, we present algorithms and numerical examples of a linear scaling finite element method to solve Hartree-Fock model with exact exchange [71]. In Chapter 5, we present formulations and numerical

results of stability analysis on numerical solutions of Hartree-Fock and density functional theory [72]. In Chapter 6, a summary of my Ph. D. work is presented. In appendix A, my Ph. D. work [70] in Dr. Zhaowei Liu group during 2009-2010 is attached.

Chapter 2

Mathematical Models of Electronic Structure

A lot of models have been proposed to study the properties of materials in the last century. These approaches include precise quantum methods, semi-empirical methods and rough molecular mechanics [85]. In quantum methods, the calculation of complex many electrons-nucleus Hamiltonian is very important , but extremely hard and time-consuming. Effective one-electron Hamiltonian models have been proposed to overcome this problem. Two famous models are Hartree-Fock theory [129] and density functional theory [66, 81]. The only difference between Hartree-Fock theory and density functional theory is the exchange and correlation energy term. Hartree-Fock theory has exact exchange energy, but ignores correlation energy completely due to the mean field approximation. Thus it works well for some weakly correlated systems. Density functional theory considers exchange and correlation energy as a whole term, and makes assumptions on this term as a functional of electron density.

Although Hartree-Fock and density functional theory models have been studied many years, but few papers [3, 46, 84, 146] discuss the mathematical properties of these models, such as the existence of a minimizer, the convergence of finite element approximation, the convergence of finite element approximation with numerical quadratures and the convergence with pseudopotential approximation. In this chapter, we will give a clear and precise description of the Hartree-Fock and density functional theory formulations. In the next chapter, we will present the mathematical proofs of the mathematical

properties of these models.

This chapter is organized as follows. Section 2.1 gives an introduction to quantum mechanics. Section 2.2 describes the mathematical formulations for Hartree-Fock theory. Section 2.3 describes the mathematical formulations for density functional theory. Section 2.4 introduces the Self-consistent methods, which are used to solve Hartree-Fock and density functional theory numerically.

2.1 Elementary Quantum Mechanics

The section follows Chapter 1 and Chapter 2 of Parr and Yang's book [112]. We will introduce the Schrödinger equation in Section 2.1.1, which is the most fundamental knowledge in quantum mechanics. Density matrices will be discussed in Section 2.1.2. Since electronic density is the only basic variable in density functional theory, that's why density matrices are very important.

2.1.1 The Schrödinger Equation

This section reviews elementary quantum theory [112, 165, 131, 88, 111, 98, 148]. Schrödinger equation covers all the problems in the electronic structure of matter including the time. In most cases, people are interested in atoms and molecules without time dependent interactions. Thus we will focus on the time independent Schrödinger's equation. For an isolated N-electron system, this is given by

$$\widehat{H}\Upsilon = E\Upsilon, \tag{2.1.1}$$

where E is the electronic energy, $\Upsilon = \Upsilon(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N)$ is the wave function, and the $\widehat{H}$ is the Hamiltonian operator,

$$\widehat{H} = \sum_{\mu=1}^{N} -\frac{1}{2}\nabla_\mu^2 + \sum_{\mu=1}^{N} V_{ext}(\vec{r}^{\mu}) + \frac{1}{2}\sum_{\mu\neq\nu,\mu\nu=1}^{N} \frac{1}{|\vec{x}^{\mu} - \vec{x}^{\nu}|}., \tag{2.1.2}$$

where $\vec{r}$ comprise $\vec{x}$ and s stands for space coordinates and spin coordinates respectively, and the superscript μ labels for the μth electron.

(2.1.1) must be solved subject to appropriate boundary conditions. Zero at infinity is an appropriate boundary condition for an atom or a molecule. Periodic boundary

conditions might be good for a regular infinite solid. $|\Upsilon|^2$ is a probability distribution function in the sense that$|\Upsilon(\vec{r}_N)|^2 d\vec{r}_N$ is the probability of finding electrons with position coordinates between $\vec{x}_N$ and $\vec{x}_N + d\vec{x}_N$ and spin coordinates equal to s_N. Here are some coordinate notations:

$$d\vec{r}_N = d\vec{r}^1 d\vec{r}^2 \dots d\vec{r}^N, \tag{2.1.3}$$

$$d\vec{r}^i = d\vec{x}^i ds^i \tag{2.1.4}$$

$$d\vec{x}_N = d\vec{x}^1 d\vec{x}^2 \dots d\vec{x}^N, \tag{2.1.5}$$

$$\vec{x}_N = \left(\vec{x}^1, \vec{x}^2, \dots, \vec{x}^N\right), \tag{2.1.6}$$

$$s_N = \left(s^1, s^2, \dots, s^N\right). \tag{2.1.7}$$

The spatial coordinates $\vec{x}_N$ are continuous. The spin coordinates s_N are discrete. Because of Pauli exclusion principle, Υ needs to be antisymmetric with respect to interchange of any two electrons' coordinates.

For a given system, there are multiple solutions of (2.1.1). Each solution will include eigenfunctions Υ_k and their corresponding energy eigenvalues E_k. Because the set Υ_k is complete and the physics properties are unchanged under a unitary transformation, we could always choose orthogonal and normalized Υ_k,

$$\int \Upsilon_k^* \Upsilon_l d\vec{r}_N = \langle \Upsilon_k | \Upsilon_l \rangle = \delta_{kl}. \tag{2.1.8}$$

For a Hermitian operator $\widehat{A}$, its expectation value is given by

$$\langle \widehat{A} \rangle = \frac{\int \Upsilon^* \widehat{A} \Upsilon d\vec{r}_N}{\int \Upsilon^* \Upsilon d\vec{r}_N} = \frac{\langle \Upsilon | \widehat{A} | \Upsilon \rangle}{\langle \Upsilon | \Upsilon \rangle}. \tag{2.1.9}$$

The Hamiltonian operator $\widehat{H}$ is hermitian, and expectation values of the electronic energy is given by

$$E(\Upsilon) = \langle \widehat{H} \rangle = \frac{\int \Upsilon^* \widehat{H} \Upsilon d\vec{r}_N}{\int \Upsilon^* \Upsilon d\vec{r}_N} = \frac{\langle \Upsilon | \widehat{H} | \Upsilon \rangle}{\langle \Upsilon | \Upsilon \rangle}. \tag{2.1.10}$$

We denote the ground-state wave function and energy by Υ_0 and E_0, and we have

$$E(\Upsilon) \geq E_0. \tag{2.1.11}$$

The ground state energy E_0 is a lower bound to the energy computed from a guessed Υ. The real ground state Υ_0 and energy $E(\Upsilon_0) = E_0$ is given by the minimization of the functional $E(\Upsilon)$,

$$E_0 = \min_{\Upsilon} E(\Upsilon). \tag{2.1.12}$$

Every eigenstate Υ is an extremum of the functional $E(\Upsilon)$. By variational principle, we could replace the Schrödinger's equation (2.1.1) by

$$\delta E(\Upsilon) = 0. \tag{2.1.13}$$

It is convenient to restate (2.1.13) in a way that will guarantee that the final Υ will be normalized. By the method of Lagrange multipliers, extremization of $\langle\Upsilon|\widehat{H}|\Upsilon\rangle$ subject to the constant $\langle\Upsilon|\Upsilon\rangle = 1$ is equivalent to making $\Big[\langle\Upsilon|\widehat{H}|\Upsilon\rangle - E\langle\Upsilon|\Upsilon\rangle\Big]$ stationary without constraint, where E is the Lagrange multiplier. This gives

$$\delta\Big[\langle\Upsilon|\widehat{H}|\Upsilon\rangle - E\langle\Upsilon|\Upsilon\rangle\Big] = 0 \tag{2.1.14}$$

The electron density is defined by

$$\rho(\vec{x}) = N\int\cdots\int|\Upsilon(\vec{r},\vec{r^2},\ldots\vec{r^N})|^2 ds d\vec{r^2}\cdots d\vec{r^N}, \tag{2.1.15}$$

2.1.2 Density Matrices

We will introduce density matrices in this section [112, 32, 99, 148, 159]. A general description of a quantum state could be described as

$$\Upsilon(\vec{r^1},\vec{r^2},\ldots,\vec{r^N})\Upsilon^*(\vec{r^1},\vec{r^2},\ldots,\vec{r^N}) \tag{2.1.16}$$

is the probability distribution associated with a solution of the Schrödinger equation (2.1.1).

We can define a more general quantity as

$$\gamma_N(\vec{r^{1'}},\vec{r^{2'}},\ldots,\vec{r^{N'}},\vec{r^1},\vec{r^2},\ldots,\vec{r^N}) = \Upsilon(\vec{r^{1'}},\vec{r^{2'}},\ldots,\vec{r^{N'}})\Upsilon^*(\vec{r^1},\vec{r^2},\ldots,\vec{r^N}). \tag{2.1.17}$$

The two sets of coordinates $\vec{r^{1'}},\vec{r^{2'}},\ldots\vec{r^{N'}}$ and $\vec{r^1},\vec{r^2},\ldots\vec{r^N}$ are two sets of indices that give (2.1.17) a numerical value. We therefore might consider (2.1.17) as an element of

a matrix, and we call this matrix density matrix. If we set $\vec{r}^{\,i} = \vec{r}^{\,i'}$ for all i, we get a diagonal element of the density matrix as the original (2.1.16). We can also consider (2.1.17) as the coordinate representation of the density operator,

$$\widehat{\gamma_N} = |\Upsilon\rangle\langle\Upsilon|. \tag{2.1.18}$$

For an operator $\widehat{A}$, we can also get

$$\langle\widehat{A}\rangle = tr(\widehat{\gamma_N}\widehat{A}) = tr(\widehat{A}\widehat{\gamma_N}). \tag{2.1.19}$$

A pure state could be described by a wave function, and a mixed state cannot be described by a wave function. We only discussed the pure state above, and we will discuss the mixed state below. A mixed state can be characterized by a probability distribution over all the possible pure states, described by the ensemble density operator as a generalization of the density operator of (2.1.18). Here is the definition of the ensemble density operator:

$$\widehat{\Gamma_N} = \sum_i p_i|\Upsilon_i\rangle\langle\Upsilon_i|, \tag{2.1.20}$$

where p_i is the probability being found in the state $|\Upsilon_i\rangle$, and the sum is over all possible pure states. Since the $|\Upsilon_i\rangle$ are orthonormal and the probability p_i are positive, we have

$$p_i \geq 0,\ \sum_i p_i = 1. \tag{2.1.21}$$

For a mixed state, the expectation value for the observable $\widehat{A}$ can be defined similarly as (2.1.19)

$$\langle\widehat{A}\rangle = tr(\widehat{\Gamma_N}\widehat{A}) = \sum_i p_i\langle\Upsilon_i|\widehat{A}|\Upsilon_i\rangle. \tag{2.1.22}$$

The Hamiltonian operator of (2.1.2) is the sum of two symmetric one-electron operators and a symmetric two-electron operator, which does not depend on spin. Similarly, operators corresponding to other physical observables are always one-electron or two-electron type and often are spin free. We therefore could simplify the expectation value formulas (2.1.19) and (2.1.22) by integrating over $N-2$ of its variables, and the concepts of reduced density matrix will be introduced.

We call (2.1.17) the density matrix of order N for a pure state of an N-electron system. Then we define the pth order reduced density matrix by

$$\begin{aligned} & \gamma_p(\vec{r}^{1'}, \ldots \vec{r}^{p'}, \vec{r}^{1}, \ldots \vec{r}^{p}) \\ = \quad & C_N^p \int \cdots \int \gamma_N(\vec{r}^{1'}, \ldots, \vec{r}^{p'}, \vec{r}^{p+1}, \ldots, \vec{r}^{N}, \vec{r}^{1}, \ldots, \vec{r}^{p}, \vec{r}^{p+1}, \ldots, \vec{r}^{N}) d\vec{r}^{p+1} \cdots d\vec{x}^{N}, \end{aligned} \tag{2.1.23}$$

where C_N^p is the binomial coefficient. In particular, the second order reduced density matrix is

$$\begin{aligned} & \gamma_2(\vec{r}^{1'}, \vec{r}^{2'}, \vec{r}^{1}, \vec{r}^{2}) \\ = \quad & \frac{N(N-1)}{2} \int \cdots \int \Upsilon(\vec{r}^{1'}, \vec{r}^{2'}, \vec{r}^{3}, \ldots, \vec{r}^{N}) \Upsilon^*(\vec{r}^{1}, \vec{r}^{2}, \vec{r}^{3}, \ldots, \vec{r}^{N}) d\vec{r}^{3} \cdots d\vec{r}^{N}, \end{aligned} \tag{2.1.24}$$

and the first order reduced density matrix is

$$\gamma_1(\vec{r}^{1'}, \vec{r}^{1}) = N \int \cdots \int \Upsilon(\vec{r}^{1'}, \vec{r}^{2}, \ldots, \vec{r}^{N}) \Upsilon^*(\vec{r}^{1}, \vec{r}^{2}, \ldots, \vec{r}^{N}) d\vec{r}^{2} \cdots d\vec{r}^{N}. \tag{2.1.25}$$

For a mixed state, the reduced density matrices and operators could be defined correspondingly, and the same properties all hold. We denote (2.1.20) as the Nth-order density operator. Then the pth-order mixed state density matrix is defined as

$$\begin{aligned} & \Gamma_p(\vec{r}^{1'}, \ldots \vec{r}^{p'}, \vec{r}^{1}, \ldots \vec{r}^{p}) \\ = \quad & C_N^p \int \cdots \int \Gamma_N(\vec{r}^{1'}, \ldots, \vec{r}^{p'}, \vec{r}^{p+1}, \ldots, \vec{r}^{N}, \vec{r}^{1}, \ldots, \vec{r}^{p}, \vec{r}^{p+1}, \ldots, \vec{r}^{N}) d\vec{r}^{p+1} \cdots d\vec{r}^{N}. \end{aligned} \tag{2.1.26}$$

Most of formulas below hold for mixed states as well as pure states, but we will not discuss them separately.

For an antisymmetric N-body wavefunction Υ, the expectation value of a one-electron operator

$$\widehat{A_1} = \sum_{i=1}^{N} A_1(\vec{r}^{i}, \vec{r}^{i'}). \tag{2.1.27}$$

We have

$$\langle \widehat{A_1} \rangle = tr(\widehat{A_1}\widehat{\gamma_N}) = \int A_1(\vec{x}^{1}, \vec{x}^{1'}) \gamma_1(\vec{r}^{1'}, \vec{r}^{1}) d\vec{r}^{1} d\vec{r}^{1'}. \tag{2.1.28}$$

All one-electron operators of interest are local, conventionally we write down only the diagonal part as

$$\widehat{A_1} = \sum_{i=1}^{N} A_1(\vec{r}^i), \tag{2.1.29}$$

and the corresponding expectation-value could be described as

$$\langle\widehat{A_1}\rangle = \int [A_1(\vec{r}^1)\gamma_1(\vec{r}^{1'}, \vec{r}^1)]_{\vec{r}^{1'}=\vec{r}^1} d\vec{r}^1. \tag{2.1.30}$$

As most two-electron operators in molecular physics are local, so we may also denote the operators by their diagonal part, which is

$$\widehat{A_2} = \frac{1}{2} \sum_{i\neq j, ij=1}^{N} A_2(\vec{r}^i, \vec{r}^j), \tag{2.1.31}$$

and its corresponding expectation value is

$$\langle\widehat{A_2}\rangle = \int\int [A_2(\vec{r}^1, \vec{r}^2)\gamma_2(\vec{r}^{1'}, \vec{r}^{2'}, \vec{r}^1, \vec{r}^2)]_{\vec{r}^{1'}=\vec{r}^1, \vec{r}^{2'}=\vec{r}^2} d\vec{r}^1 d\vec{r}^2. \tag{2.1.32}$$

Many operators that concern us do not involve spin coordinates, for instance the Hamiltonian operators for atoms or molecules. We therefore could make further reduction of the density matrices by summation over the spin coordinates. We denote the pth order spinless density matrices by

$$\rho_p(\vec{x}^1 \dots \vec{x}^p) = \int \cdots \int \gamma_p(\vec{r}^{1'}, \dots, \vec{r}^{p'}, \vec{r}^1, \dots, \vec{r}^p)|_{\vec{r}^{1'}=\vec{r}^1, \dots, \vec{x}^{p'}=\vec{x}^p} ds^1 \cdots ds^p. \tag{2.1.33}$$

and

$$\rho(\vec{x}^1) = \rho_1(\vec{x}^1) \tag{2.1.34}$$

The total energy (2.1.16) becomes

$$E = \int [(-\frac{1}{2}\nabla_1^2 + V_{ext}(\vec{x}^1))\rho_1(\vec{x}^{1'}, \vec{x}^1)]_{\vec{x}^{1'}=\vec{x}^1} d\vec{x}^1 + \int\int \frac{1}{|\vec{x}^1 - \vec{x}^2|}\rho_2(\vec{x}^1, \vec{x}^2) d\vec{x}^1 d\vec{x}^2. \tag{2.1.35}$$

We can rewrite $\rho_2(\vec{x}^1, \vec{x}^2)$ as

$$\rho_2(\vec{x}^1, \vec{x}^2) = \frac{1}{2}\rho(\vec{x}^1)\rho(\vec{x}^2)[1 + h(\vec{x}^1, \vec{x}^2)]. \tag{2.1.36}$$

Integrate $\vec{x}^2$, we find

$$\frac{N-1}{2}\rho(\vec{x}^1) = \frac{1}{2}\rho(\vec{x}^1)[N + \int \rho(\vec{x}^2)h(\vec{x}^1, \vec{x}^2) d\vec{x}^2]. \tag{2.1.37}$$

We have

$$\int \rho(\vec{x}^2)h(\vec{x}^1,\vec{x}^2)d\vec{x}^2 = -1, \tag{2.1.38}$$

which holds for all $\vec{x}^1$. If we denote the exchange-correlation hole of an electron at x^1 by

$$\rho_{xc}(\vec{x}^1,\vec{x}^2) = \rho(\vec{x}^2)h(\vec{x}^1,\vec{x}^2), \tag{2.1.39}$$

then we have

$$\int \rho_{xc}(\vec{x}^1,\vec{x}^2)d\vec{x}^2 = -1, \tag{2.1.40}$$

which is a unit charge opposite to that of the electron. The electron interaction term could be expressed in terms of ρ_{xc},

$$\begin{aligned}\int\int \frac{1}{|\vec{x}^1-\vec{x}^2|}\rho_2(\vec{x}^1,\vec{x}^2)d\vec{x}^1d\vec{x}^2 &= \frac{1}{2}\int\int \frac{1}{|\vec{x}^1-\vec{x}^2|}\rho(\vec{x}^1)\rho(\vec{x}^2)d\vec{x}^1d\vec{x}^2 \\ &+\frac{1}{2}\int\int \frac{1}{|\vec{x}^1-\vec{x}^2|}\rho(\vec{x}^1)\rho_{xc}(\vec{x}^1,\vec{x}^2)d\vec{x}^1d\vec{x}^2\end{aligned} \tag{2.1.41}$$

2.1.3 A Table of Models

We will pick up a few most important models, and study the mathematical properties of these models in Chapter 2 and Chapter 3. For your convenience, Tabel (2.1) is a list of these models, which will be discussed carefully later.

As shown in Tabel 2.1, Hartree-Fock method includes restricted Hartree-Fock method and unrestricted Hartree-Fock method, and density functional theory includes local spin density approximation, generalized gradient approximation, meta generalized gradient approximation and generalized density functional theory models.

2.2 Hartree-Fock

The section follows Roothaan's paper [129], and introduces Hartree-Fock method . Given an N-electron system, Hartree-Fock method assigns each electron a wavefunction depending on the space coordinates and spin coordinates, called a molecular spinorbital (MSO):

$$\psi_\kappa^\mu = \psi_\kappa(\vec{x}^\mu, s^\mu), \tag{2.2.1}$$

Table 2.1: A List of Models in Chapter 2 and Chapter 3

Category	Model
Hartree-Fock (HF)	Restricted Hartree-Fock (RHF)
	Unrestricted Hartree-Fock (UHF)
Density Functional Theory (DFT)	Local Spin Density Approximation (LSDA)
	Generalized Gradient Approximation (GGA)
	Meta Generalized Gradient Approximation (Meta-GGA)
	Generalized Density Functional Theory Models

where $\vec{x}$ and s stands for space coordinates and spin coordinates respectively, the superscript μ labels for the μth electron, and the subscript κ labels different molecular spinorbitals.

The total N-electron wave function can be approximated by a single Slater determinant of N molecular spinorbitals [50].

$$\Upsilon = (N!)^{-\frac{1}{2}} \begin{vmatrix} \psi_1^1 & \psi_2^1 & \dots & \psi_N^1 \\ \psi_1^2 & \psi_2^2 & \dots & \psi_N^2 \\ \dots & \dots & \dots & \dots \\ \psi_1^N & \psi_2^N & \dots & \psi_N^N \end{vmatrix}. \tag{2.2.2}$$

Since we neglect spin-orbital interaction, each molecular spinorbital factors into a molecular orbital (MO) and a spin function:

$$\psi_\kappa^\mu = \phi_{i(\kappa)}(\vec{x}^\mu)\eta_\kappa(s^\mu) = \phi_{i(\kappa)}^\mu \eta_\kappa^\mu \tag{2.2.3}$$

where $\phi_{i(\kappa)}^\mu$ is a molecular orbital and η_κ^μ is the spin factor. We denote the spin factors by:

$$\eta_\kappa^\mu = \begin{cases} \alpha^\mu \\ \beta^\mu \end{cases}, \tag{2.2.4}$$

where α stands for spin-up and β stands for spin-down.

We denote the set of molecular spinorbitals ψ_κ by a row vector Ψ:

$$\Psi = (\psi_1 \psi_2 \dots \psi_N)\,. \tag{2.2.5}$$

The Pauli exclusion principle guarantees that all the molecular spinorbitals must be linearly independent.The physical property of Ψ is same under a unitary transformation:

$$\Psi^{'} = \Psi Det(A), \tag{2.2.6}$$

with $AA^* = A^*A = \mathbf{I}$, $\mathbf{I}$ being the identity matrix.

We could choose the transformation matrix A such that molecular spinorbitals Ψ' are orthonomal. Thus we might always assume that molecular spinorbitals are orthonormal:

$$\int \psi^*_\kappa(\vec{x})\psi_\lambda(\vec{x})d\vec{x} = \delta_{\kappa\lambda}, \quad \kappa, \lambda = 1, \dots, N, \tag{2.2.7}$$

without loss of generality.

It is straightforward to show N-electron wave function (2.2.2) is normalized, that is:

$$\int \Upsilon^*(\vec{x}_N)\Upsilon(\vec{x}_N)d\vec{x}_N = 1. \tag{2.2.8}$$

When and electronic state is approximated by a Slater determinant Υ (2.2.2), the total electronic energy is given by:

$$\mathcal{E} = \int \Upsilon^*(\vec{x}_N)\widehat{H}\Upsilon(\vec{x}_N)d\vec{x}_N, \tag{2.2.9}$$

where the total Hamiltonian operator $\widehat{H}$ is defined by

$$\widehat{H} = \sum_{\mu=1}^{N} H^\mu + \frac{1}{2}\sum_{\mu\neq\nu,\mu\nu=1}^{N} \frac{1}{|\vec{x}^\mu - \vec{x}^\nu|}, \tag{2.2.10}$$

$$= \sum_{\mu=1}^{N}\left[-\frac{1}{2}\nabla^2_\mu + V_{ext}(\vec{x}^\mu)\right] + \frac{1}{2}\sum_{\mu\neq\nu,\mu\nu=1}^{N} \frac{1}{|\vec{x}^\mu - \vec{x}^\nu|}, \tag{2.2.11}$$

where H^μ is the Hamiltonian operator for the μth electron moving in the external field.

The exact wave function for a system of many interacting electrons is never a single determinant or a simple combination of a few determinants. The calculation of the error in energy by Hartree-Fock is called correlation energy. There has been a vast amount of work and much progress has been made on calculating correlation

energy. The methods employed include the linear mixing of many determinants called configuration interaction [148] and many-body perturbation techniques [148]. You can also find more details in comprehensive reviews [73, 162].

We will introduce restricted Hartree-Fock method and unrestricted Hartree-Fock method in Section 2.2.1 and Section 2.2.2 separately.

2.2.1 Restricted Hartree-Fock

Restricted Hartree-Fock method assumes each occupied molecular spinorbital is occupied by two electrons with opposite spins. Suppose the electronic system has $2n$ electrons and n fully occupied molecular orbitals, we can denote the molecular spinorbitals by:

$$\Psi = \{\psi_1 \psi_2 \dots \psi_{2n}\}, \tag{2.2.12}$$

and molecular orbitals by:

$$\Phi = \{\phi_1 \phi_2 \dots \phi_n\}. \tag{2.2.13}$$

The relationship among molecular spinorbitals and molecular orbital are given by:

$$\psi_{2i-1} = \phi_i \alpha, \quad i = 1, \dots, n, \tag{2.2.14}$$

$$\psi_{2i} = \phi_i \beta, \quad i = 1, \dots, n. \tag{2.2.15}$$

Because (2.2.7) and the orthonormality of spin factors, we can obtain:

$$\int \phi_i^*(\vec{x}) \phi_j(\vec{x}) d\vec{x} = \delta_{ij}, \quad i, j = 1, \dots, n. \tag{2.2.16}$$

The $2n$-electron wave function is given by

$$\begin{aligned}\Upsilon &= [(2n)!]^{-\frac{1}{2}} \begin{vmatrix} \psi_1^1 & \psi_2^1 & \dots & \psi_{2n}^1 \\ \psi_1^2 & \psi_2^2 & \dots & \psi_{2n}^2 \\ \dots & \dots & \dots & \dots \\ \psi_1^{2n} & \psi_2^{2n} & \dots & \psi_{2n}^{2n} \end{vmatrix}, \\ &= [(2n)!]^{-\frac{1}{2}} \begin{vmatrix} (\phi_1\alpha)^1 & (\phi_1\beta)^1 & \dots & (\phi_n\alpha)^1 & (\phi_n\beta)^1 \\ (\phi_1\alpha)^2 & (\phi_1\beta)^2 & \dots & (\phi_n\alpha)^2 & (\phi_n\beta)^2 \\ \dots & \dots & \dots & \dots & \dots \\ (\phi_1\alpha)^{2n-1} & (\phi_1\beta)^{2n-1} & \dots & (\phi_n\alpha)^{2n-1} & (\phi_n\beta)^{2n-1} \\ (\phi_1\alpha)^{2n} & (\phi_1\beta)^{2n} & \dots & (\phi_n\alpha)^{2n} & (\phi_n\beta)^{2n} \end{vmatrix}.\end{aligned}$$

Substituting this wavefunction into the expression (2.2.9), we find the total energy is:

$$\mathcal{E} = T(\Psi) + V(\Psi) + J(\Psi) + K(\Psi). \tag{2.2.17}$$

The first term is the kinetic energy of the non-interacting electrons,

$$T(\Psi) = 2\sum_{i=1}^{n} T_i = -\sum_{i=1}^{n} \int \phi_i^*(\vec{x})\nabla^2\phi_i(\vec{x})d\vec{x}. \tag{2.2.18}$$

The second term is the electronic energy between electrons and the external field $V_{ext}(\vec{x})$,

$$V(\Psi) = 2\sum_{i=1}^{n} V_i = 2\sum_{i=1}^{n} \int \phi_i^*(\vec{x})V_{ext}(\vec{x})\phi_i(\vec{x})d\vec{x}. \tag{2.2.19}$$

The third term is the electrostatic interaction energy among electrons,

$$J(\Psi) = 2\sum_{i,j=1}^{n} J_{ij} = 2\sum_{i,j=1}^{n} \int\int \frac{\phi_i^*(\vec{x})\phi_j^*(\vec{x}')\phi_i(\vec{x})\phi_j(\vec{x}')}{|\vec{x}-\vec{x}'|}d\vec{x}d\vec{x}'. \tag{2.2.20}$$

The last term is the exact exchange energy of electrons,

$$K(\Psi) = -\sum_{i,j=1}^{n} K_{ij} = -\sum_{i,j=1}^{n} \int\int \frac{\phi_i^*(\vec{x})\phi_j^*(\vec{x}')\phi_j(\vec{x})\phi_i(\vec{x}')}{|\vec{x}-\vec{x}'|}d\vec{x}d\vec{x}'. \tag{2.2.21}$$

When each molecular orbital ϕ_i is varied by an infinitesimal amount $\delta\phi_i$, the

variation of the energy becomes

$$\begin{aligned}\delta\mathcal{E} &= 2\sum_{i=1}^{n}(\delta T_i+\delta V_i)+\sum_{i,j=1}(2J_{ij}-K_{ij}),\\ &= 2\sum_{i=1}^{n}(\delta\phi_i^*)\left\{\widehat{T}_i+\widehat{V}_i+\sum_{j=1}^{n}(2\widehat{J}_j-\widehat{K_j})\right\}\phi_i d\vec{x}\\ &\quad +2\sum_{i=1}^{n}(\delta\phi_i)\left\{\widehat{T}_i^{\,*}+\widehat{V}_i^{\,*}+\sum_{j=1}^{n}(2\widehat{J}_j^{\,*}-\widehat{K_j}^{\,*})\right\}\phi_i^* d\vec{x}, \qquad (2.2.22)\end{aligned}$$

where $\widehat{T}_i, \widehat{V}_i, \widehat{J}_j, \widehat{K_j}$ are defined as

$$\begin{aligned}&\widehat{T}_i=-\frac{1}{2}\nabla^2,\\ &\widehat{V}_i=V_{ext}(\vec{x}),\\ &\widehat{J}_j\phi_i(\vec{x})=\int\frac{\phi_j^*(\vec{x}')\phi_j(\vec{x}')}{|\vec{x}-\vec{x}'|}d\vec{x}'\phi_i(\vec{x}),\\ &\widehat{K_j}\phi_i(\vec{x})=\int\frac{\phi_j^*(\vec{x}')\phi_i(\vec{x}')}{|\vec{x}-\vec{x}'|}d\vec{x}'\phi_j(\vec{x}). \qquad (2.2.23)\end{aligned}$$

The resulting restrictions on the variation $\delta\phi_i$, obtained by varying (2.2.16), are as follows:

$$\int(\delta(\phi_i)^*(\vec{x}))\phi_j(\vec{x})d\vec{x}+\int(\delta\phi_j)(\vec{x}))\phi_i^*(\vec{x})d\vec{x}=0. \qquad (2.2.24)$$

By the method of the Lagrange multiplier, we multiply each (2.2.24) by the Lagrange multiplier $-2\epsilon_{ji}$ and add the resulting equations together, we obtain

$$\begin{aligned}\delta\mathcal{E} &= 2\sum_{i=1}^{n}(\delta\phi_i^*)\left[\{\widehat{T}_i+\widehat{V}_i+\sum_{j=1}^{n}(2\widehat{J}_j-\widehat{K_j})\}\phi_i-\sum_{j=1}^{n}\phi_j\epsilon_{ji}\right]d\vec{x}\\ &\quad +2\sum_{i=1}^{n}(\delta\phi_i)\left[\{\widehat{T}_i^{\,*}+\widehat{V}_i^{\,*}+\sum_{j=1}^{n}(2\widehat{J}_j^{\,*}-\widehat{K_j}^{\,*})\}\phi_i^*-\sum_{j=1}^{n}\phi_j^*\epsilon_{ij}\right]d\vec{x}, \qquad (2.2.25)\end{aligned}$$

The conditions for $\delta\mathcal{E}=0$ will give

$$\{\widehat{T}_i+\widehat{V}_i+\sum_{j=1}^{n}(2\widehat{J}_j-\widehat{K_j})\}\phi_i=\sum_{j=1}^{n}\phi_j\epsilon_{ji}, \qquad (2.2.26)$$

$$\{\widehat{T}_i^{\,*}+\widehat{V}_i^{\,*}+\sum_{j=1}^{n}(2\widehat{J}_j^{\,*}-\widehat{K_j}^{\,*})\}\phi_i^*=\sum_{j=1}^{n}\phi_j^*\epsilon_{ij}. \qquad (2.2.27)$$

Take the complex conjugate of (2.2.27), and subtract it from the (2.2.26). We get

$$\sum_{j=1}^{n} \phi_j(\epsilon_{ji} - \epsilon_{ij}^*) = 0. \tag{2.2.28}$$

Since ϕ_j are linearly independent, it follows that

$$\epsilon_{ji} = \epsilon_{ij}^*, \tag{2.2.29}$$

which shows the matrix ϵ is Hermitian. Thus (2.2.26) and (2.2.27) are equivalent, and (2.2.26) could be written as

$$\widehat{F}\phi_i = \sum_{j=1}^{n} \phi_j \epsilon_{ji}, \tag{2.2.30}$$

where

$$\widehat{F} = \widehat{T}_i + \widehat{V}_i + \sum_{j=1}^{n} (2\widehat{J}_j - \widehat{K_j}). \tag{2.2.31}$$

we could express (2.2.30) in matrix form as

$$\widehat{F}\phi = \phi\epsilon. \tag{2.2.32}$$

Now let us subject molecular orbitals and their corresponding matrix ϵ to a unitary transformation

$$\begin{aligned} \phi^{'} &= \phi B, \\ \epsilon^{'} &= B^* \epsilon B, \end{aligned} \tag{2.2.33}$$

where B is a unitary matrix. It follows that

$$\widehat{F'}\phi = \phi\epsilon. \tag{2.2.34}$$

We have

$$\sum_{i=1}^{n} \phi_i^{*'} \phi_i^{'} = \sum_{j,k=1}^{n} \phi_j^* \phi_k \sum_{i=1}^{n} B_{ji}^* B_{ki} = \sum_{j,k=1}^{n} \phi_j^* \phi_k \delta_{jk} = \sum_{j=1}^{n} \phi_j^* \phi_j, \tag{2.2.35}$$

from this equality we could get that

$$\sum_{i=1}^{n} \widehat{J_j^{'}} = \sum_{i=1}^{n} \widehat{J}_j, \tag{2.2.36}$$

$$\sum_{i=1}^{n} \widehat{K_j^{'}} = \sum_{i=1}^{n} \widehat{K_j}. \tag{2.2.37}$$

Thus, we have

$$\widehat{F'} = \widehat{F}. \tag{2.2.38}$$

Since the matrix ϵ is Hermitian, there exists an unitary matrix B such that $\epsilon^{'}$ is a real diagonal matrix. Therefore, we can replace (2.2.30) by the eigenvalue problem

$$\widehat{F}\phi_i = \epsilon_i\phi_i, i = 1, 2, \ldots, n, \tag{2.2.39}$$

without any loss of generality.

2.2.2 Unrestricted Hartree-Fock

Unrestricted Hartree-Fock method is the most common molecular orbital method. You could consider restricted Hartree-Fock as a special case of unrestricted Hartree-Fock. In unrestricted Hartree-Fock method, it assumes the spatial orbitals for spin-up α and spin-down electrons β are different. Suppose the electronic system has n_α spin-up electrons and n_β spin-down electrons, we denote the molecular spinorbitals by:

$$\Psi = \{\psi_1\psi_2 \ldots \psi_{n_\alpha+n_\beta}\}, \tag{2.2.40}$$

and molecular orbitals by

$$\Phi = \{\phi_1\phi_2 \ldots \phi_{n_\alpha+n_\beta}\}. \tag{2.2.41}$$

The $(n_\alpha + n_\beta)$-electron wave function is given by

$$\begin{aligned}
\Upsilon &= C \begin{vmatrix} \psi_1^1 & \psi_2^1 & \ldots & \psi_{n_\alpha+n_\beta}^1 \\ \psi_1^2 & \psi_2^2 & \ldots & \psi_{n_\alpha+n_\beta}^2 \\ \ldots & \ldots & \ldots & \ldots \\ \psi_1^{n_\alpha+n_\beta} & \psi_2^{n_\alpha+n_\beta} & \ldots & \psi_{n_\alpha+n_\beta}^{n_\alpha+n_\beta} \end{vmatrix}, \\
&= C \begin{vmatrix} (\phi_1\alpha)^1 & \ldots & (\phi_{n_\alpha}\alpha)^1 & (\phi_{n_\alpha+1}\beta)^1 & \ldots & (\phi_{n_\alpha+n_\beta}\beta)^1 \\ (\phi_1\alpha)^2 & \ldots & (\phi_{n_\alpha}\alpha)^2 & (\phi_{n_\alpha+1}\beta)^2 & \ldots & (\phi_{n_\alpha+n_\beta}\beta)^2 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ (\phi_1\alpha)^{n_\alpha} & \ldots & (\phi_{n_\alpha}\alpha)^{n_\alpha} & (\phi_{n_\alpha+1}\beta)^{n_\alpha} & \ldots & (\phi_{n_\alpha+n_\beta}\beta)^{n_\alpha} \\ (\phi_1\alpha)^{n_\alpha+1} & \ldots & (\phi_{n_\alpha}\alpha)^{n_\alpha+1} & (\phi_{n_\alpha+1}\beta)^{n_\alpha+1} & \ldots & (\phi_{n_\alpha+n_\beta}\beta)^{n_\alpha+1} \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ (\phi_1\alpha)^{n_\alpha+n_\beta} & \ldots & (\phi_{n_\alpha}\alpha)^{n_\alpha+n_\beta} & (\phi_{n_\alpha+1}\beta)^{n_\alpha+n_\beta} & \ldots & (\phi_{n_\alpha+n_\beta}\beta)^{n_\alpha+n_\beta} \end{vmatrix},
\end{aligned} \tag{2.2.42}$$

where C is $[(n_\alpha + n_\beta)!]^{-\frac{1}{2}}$.

Substituting this wavefunction into the expression (2.2.9), we find the total energy is:

$$\begin{aligned}\mathcal{E} &= T(\Psi) + V(\Psi) + J(\Psi) + K(\Psi), \\ &= \sum_{i=1}^{n_\alpha+n_\beta} T_i + \sum_{i=1}^{n_\alpha+n_\beta} V_i + \frac{1}{2}\sum_{i,j=1}^{n_\alpha+n_\beta} J_{ij} - \frac{1}{2}\left[\sum_{i,j=1}^{n_\alpha} + \sum_{i,j=n_\alpha+1}^{n_\alpha+n_\beta}\right] K_{ij}, \quad (2.2.43)\end{aligned}$$

where the kinetic energy of the non-interacting electrons T_i, the electronic energy between electrons and the external field V_i, the electrostatic interaction energy among electrons J_{ij} and the exact exchange energy of electrons K_{ij} are defined by (2.2.18), (2.2.19), (2.2.20) and (2.2.21).

By the method of the Lagrange multiplier and similar process in restricted Hartree-Fock, we obtain eigenvalue equation for spin-up electrons

$$\begin{aligned}\widehat{F_\alpha}\phi_i &= \{\widehat{T_i} + \widehat{V_i} + \sum_{j=1}^{n_\alpha+n_\beta} \widehat{J_j} - \sum_{j=1}^{n_\alpha} \widehat{K_j}\}\phi_i, \\ &= \epsilon_i\phi_i, i = 1, \ldots, n_\alpha, \quad (2.2.44)\end{aligned}$$

and eigenvalue equation for spin-down electrons

$$\begin{aligned}\widehat{F_\beta}\phi_i &= \{\widehat{T_i} + \widehat{V_i} + \sum_{j=1}^{n_\alpha+n_\beta} \widehat{J_j} - \sum_{j=n_\alpha+1}^{n_\alpha+n_\beta} \widehat{K_j}\}\phi_i, \\ &= \epsilon_i\phi_i, i = n_\alpha + 1, \ldots, n_\alpha + n_\beta. \quad (2.2.45)\end{aligned}$$

2.3 Density Functional Theory

This section follows Chapter 3 of Parr and Yang's book [112], and reviews density functional theory methods. Density functional theory replaces the complicated N-electron wave function by the much simpler electron density ρ. It has a long history. In the 1920s, Thomas and Fermi [150, 39, 40, 41, 95] realize that the distribution of electrons in an atom could be approximated by statistical methods. The assumptions stated by Thomas [150] are that "Electrons are distributed uniformly in the six-dimensional phase space for the motion of an electron at the rate of two for each h^3 of volume," and

that there is an effective potential field that "is itself determined by the nuclear charge and this distribution of electrons." The formulations of Thomas-Fermi model for electron density could be derived from these assumptions. We will give a derivation of Thomas-Fermi theory below.

The space could be divided into many small cubes. Each cube with side l and volume $\triangle V = l^3$ contains $\triangle N$ electrons. We assume that the electrons in each cell are fermions independent of electrons in the other cells, at the temperature 0K.

The energy levels of a particle in a three-dimensional infinite well are given by

$$\varepsilon(n_x, n_y, n_z) = \frac{1}{8l^2}(n_x^2 + n_y^2 + n_z^2) = \frac{1}{8l^2}R^2, \tag{2.3.1}$$

where n_x, n_y, n_z are positive integers. Then we could get the number of energy levels between ε and $\varepsilon + \delta\varepsilon$, that is:

$$g(\varepsilon)\triangle\varepsilon = \frac{\pi}{4}(8l^2)^{3/2}\varepsilon^{1/2}\delta\varepsilon + O((\delta\varepsilon)^2). \tag{2.3.2}$$

The probability for the occupied state with energy ε is the Fermi-Dirac distribution $f(\varepsilon)$,

$$f(\varepsilon) = \frac{1}{1 + e^{\beta(\varepsilon-\mu)}}. \tag{2.3.3}$$

$f(\varepsilon)$ will become a step function at 0 K:

$$f(\varepsilon) = \begin{cases} 1, & \varepsilon < \varepsilon_F \\ 0, & \varepsilon > \varepsilon_F \end{cases}, \tag{2.3.4}$$

where ε_F is the so-called Fermi energy. The states with energy smaller than ε_F are occupied. The states with energy greater than ε_F are unoccupied. The Fermi energy ε_F is the chemical potential μ at the temperature 0K.

The total energy of the electrons could be calculated by summing the contributions from all the energy states:

$$\triangle E = 2\int \varepsilon f(\varepsilon)g(\varepsilon)d\varepsilon = \frac{8\pi}{5}(2l^2)^{3/2}\varepsilon_F^{5/2}. \tag{2.3.5}$$

Because each energy level is occupied by two electrons, we have the factor 2 in (2.3.5). We could integrate the product of energy level density $g(\varepsilon)$ and the Fermi-Dirac distribution $f(\varepsilon)$ to get the number of electrons $\triangle N$ in the cell, that is:

$$\triangle N = 2\int f(\varepsilon)g(\varepsilon)d\varepsilon = \frac{8\pi}{3}(2l^2)^{3/2}\varepsilon_F^{3/2}. \tag{2.3.6}$$

From (2.3.5) and (2.3.6), we obtain

$$\triangle E = \frac{3}{5}\triangle N\varepsilon_F = \frac{3}{10}\left(\frac{3l^2}{8\pi}^{2/3}\right)\left(\frac{\triangle N}{l^3}\right)^{5/3} \tag{2.3.7}$$

(2.3.7) gives us a relationship between total kinetic energy and the electron density $\rho = \triangle N/l^3 = \triangle N/\triangle V$ for each cell. We could get the total kinetic energy by summing the contributions from all cells,

$$T(\rho) = C_F \int \rho^{5/3}(\vec{x})d\vec{x}, \tag{2.3.8}$$

where

$$C_F = \frac{3}{10}(3\pi^2)^{2/3}. \tag{2.3.9}$$

If we neglect the exchange and correlation energy, the total energy functional can be written as:

$$\mathcal{E} = C_F \int \rho^{5/3}(\vec{x})d\vec{x} + \int \rho(\vec{x})V_{ext}(\vec{x})d\vec{x} + \frac{1}{2}\int\int \frac{\rho(\vec{x})\rho(\vec{x}')}{|\vec{x}-\vec{x}'|}d\vec{x}d\vec{x}', \tag{2.3.10}$$

which is in terms of electron density ρ alone, and this is the energy functional of the Thomas-Fermi theory.

People try to modify and improve the Thomas-Fermi theory over the years, but it cannot give accurate enough results as other methods. Thus Thomas-Fermi theory is viewed as an oversimplified model.

However, the situation changed as the paper by Hohenberg and Kohn [66] got published. They showed that the Thomas-Fermi model can be considered as an approximation to an exact theory, density functional theory. Density functional theory includes precise kinetic energy and an approximate exchange and correlation energy term.

For an N-electron Hamiltonian (2.1.2), the minimization of the energy functional (2.1.16) will determine both the ground-state energy and the ground-state wave functions. The external potential V_{ext}, which defines the whole nuclear frame, completely fixes the Hamiltonian. Thus the number of electrons N and the external potential V_{ext} will fix all properties for the ground state.

Theorem 2.3.1 will show a one-to-one relationship between the ground state electron density ρ and external potential V_{ext}. Also the number of electrons N is determined by simple quadrature of ρ. It follows that we could use of electron density ρ

as basic variable in place of N and V_{ext}. Theorem 2.3.2 shows the energy variational principle, which uses electron density ρ as basic variable.

Theorem 2.3.1 ([66]). *(The first Hohemberg-Kohn theorem) The external potential V_{ext} is determined, within a trivial additive constant, by the electron density ρ.*

Proof. Consider the electron density ρ for for the ground state electron density of some N-electron system. If there are two external potentials V_{ext} and $V_{ext}^{'}$, and they are different by more than a constant. Suppose two external potentials give the same ρ for its ground sate, then we have two Hamiltonians $\widehat{H}$ and $\widehat{H'}$ with the same ground-state densities, but different normalized N-electron wave functions Υ and $\Upsilon^{'}$. We could take $\Upsilon^{'}$ as a trial function for the $\widehat{H}$, that is:

$$\begin{aligned} E_0 < \langle \Upsilon^{'} | \widehat{H} | \Upsilon^{'} \rangle &= \langle \Upsilon^{'} | \widehat{H'} | \Upsilon^{'} \rangle + \langle \Upsilon^{'} | \widehat{H} - \widehat{H'} | \Upsilon^{'} \rangle \\ &= E_0^{'} + \int \rho(\vec{x})[V_{ext}(\vec{x}) - V_{ext}^{'}(\vec{x})]d\vec{x}, \end{aligned} \tag{2.3.11}$$

where E_0 and $E_0^{'}$ are the ground-state energies for $\widehat{H}$ and $\widehat{H'}$ respectively. We can also take Υ as a trial function for the $\widehat{H'}$ problem, that is:

$$\begin{aligned} E_0^{'} < \langle \Upsilon | \widehat{H'} | \Upsilon \rangle &= \langle \Upsilon | \widehat{H} | \Upsilon \rangle + \langle \Upsilon | \widehat{H'} - \widehat{H} | \Upsilon \rangle \\ &= E_0 - \int \rho(\vec{x})[V_{ext}(\vec{x}) - V_{ext}^{'}(\vec{x})]d\vec{x}. \end{aligned} \tag{2.3.12}$$

Adding (2.3.11) and (2.3.12), we get the contradiction $E_0 + E_0^{'} < E_0 + E_0^{'}$, so there cannot be two different V_{ext} that give the same ρ for their ground states.

It follows that ρ also determines V_{ext}, within a trivial additive constant. □

Theorem 2.3.2 ([66]). *(The second Hohemberg-Kohn theorem) For a trial density $\rho^{'}$, such that $\rho^{'} \geq 0$ and $\int \rho^{'}(\vec{x})d\vec{x} = N$, then*

$$E_0 \leq E[\rho^{'}]. \tag{2.3.13}$$

Proof. The Hamiltonian $\widehat{H}$ has its own ground-state density ρ and ground-state wave function Υ. By the first Hohemberg-Kohn theorem, $\rho^{'}$ determines its own $V_{ext}^{'}$, Hamiltonian $\widehat{H'}$ and ground-state wave function $\Upsilon^{'}$. Thus, $\rho^{'}$ will be just a trivial function for the $\widehat{H}$ and we could take it as a trial function for the $\widehat{H}$, that is:

$$E[\rho^{'}] = \langle \Upsilon^{'} | \widehat{H} | \Upsilon^{'} \rangle \geq \langle \Upsilon | \widehat{H} | \Upsilon \rangle = E_0, \tag{2.3.14}$$

which is (2.3.13). □

Theorem 2.3.1 and Theorem 2.3.2 show the existence of a one-to-one relationship between the ground state electron density and the ground state wavefunction. Density functional theory replaced the complicated N-electron wave function Υ (2.2.2) by the much simper electron densityρ, and electron densityρ determines all properties of the ground energy state.

We will discuss the formulations of density functional theory below. For the sake of clarity and notational simplicity, we denote electrons' spinorbital wavefunctions (SOW) and orbital wavefunctions (OW) by ψ and ϕ, which are similar notations as molecular spinorbitals and molecular orbitals in Hartree-Fock method. Suppose the electronic system has n_α spin-up electrons and n_β spin-down electrons, the vector of spinorbital wavefunctions and orbital wavefunctions are given by:

$$\Psi = \{\psi_1\psi_2 \dots \psi_{n_\alpha+n_\beta}\}, \tag{2.3.15}$$

$$\Phi = \{\phi_1\phi_2 \dots \phi_{n_\alpha+n_\beta}\}, \tag{2.3.16}$$

where the subscripts label different electrons. The relationship among spinorbital wavefunctions and orbital wavefunctions are given by:

$$\psi_i = \phi_i\alpha, \qquad i = 1, \dots, n_\alpha, \tag{2.3.17}$$

$$\psi_i = \phi_i\beta, \qquad i = n_\alpha + 1, \dots, n_\alpha + n_\beta. \tag{2.3.18}$$

The electron density are given by:

$$\rho_\alpha(\vec{x}) = \sum_{i=1}^{n_\alpha} \phi_i^*(\vec{x})\phi_i(\vec{x}) = \sum_{i=1}^{n_\alpha} |\phi_i(\vec{x})|^2, \tag{2.3.19}$$

$$\rho_\beta(\vec{x}) = \sum_{i=n_\alpha+1}^{n_\alpha+n_\beta} \phi_i^*(\vec{x})\phi_i(\vec{x}) = \sum_{i=n_\alpha+1}^{n_\alpha+n_\beta} |\phi_i(\vec{x})|^2, \tag{2.3.20}$$

$$\rho(\vec{x}) = \rho_\alpha(\vec{x}) + \rho_\beta(\vec{x}). \tag{2.3.21}$$

Spinorbital wavefunctions and orbital wavefunctions are orthonormal, thereby satisfying the relation:

$$\int \psi_i^*(\vec{x})\psi_j(\vec{x})d\vec{x} = \delta_{ij}, \qquad i, j = 1, \dots, n_\alpha + n_\beta, \tag{2.3.22}$$

$$\int \phi_i^*(\vec{x})\phi_j(\vec{x})d\vec{x} = \delta_{ij}, \qquad i, j = 1, \dots, n_\alpha, \tag{2.3.23}$$

$$\int \phi_i^*(\vec{x})\phi_j(\vec{x})d\vec{x} = \delta_{ij}, \qquad i, j = n_\alpha + 1, \dots, n_\alpha + n_\beta. \tag{2.3.24}$$

Also the system requires the conservation of charge:

$$\int \rho_\alpha(\vec{x})d\vec{x} = n_\alpha, \tag{2.3.25}$$

$$\int \rho_\beta(\vec{x})d\vec{x} = n_\beta, \tag{2.3.26}$$

$$\int \rho(\vec{x})d\vec{x} = n_\alpha + n_\beta. \tag{2.3.27}$$

The total electronic energy of an $(n_\alpha + n_\beta)$-electron system is given by:

$$\mathcal{E} = T(\rho) + V(\rho) + J(\rho) + K(\rho). \tag{2.3.28}$$

The first term is the kinetic energy of the non-interacting electrons,

$$T(\rho) = -\frac{1}{2}\sum_{i=1}^{n_\alpha+n_\beta}\int \phi_i^*(\vec{x})\nabla^2\phi_i(\vec{x})d\vec{x}. \tag{2.3.29}$$

The second term is the elctronic energy between electrons and external field, where $V_{ext}(\vec{x})$ is the external field,

$$V(\rho) = \int \rho(\vec{x})V_{ext}(\vec{x})d\vec{x}. \tag{2.3.30}$$

The third term is the electrostatic interaction energy among the electrons,

$$J(\rho) = \frac{1}{2}\int\int \frac{\rho(\vec{x})\rho(\vec{x}')}{|\vec{x}-\vec{x}'|}d\vec{x}d\vec{x}'. \tag{2.3.31}$$

The last term is the electron exchange and correlation energy term,

$$K(\rho) = \int \mathbb{k}(\rho(\vec{x}))d\vec{x}. \tag{2.3.32}$$

The first three terms in (2.3.28) are always fixed. The only difference among different density functional theory models is $K(\rho)$. Numerous approaches have been proposed to approximate electron exchange and correlation energy term $K(\rho)$.

Section 2.3.1 introduces local spin density approximation. Section 2.3.2 introduces generalized gradient approximation. Meta generalized gradient approximation and generalized density functional theory models are discussed in Section 2.3.3 and Section 2.3.4.

2.3.1 Local Spin Density Approximation

If we assume $K(\rho)$ is only a function of ρ itself, which is right if the electron gas is uniform, (2.3.28) becomes local spin density approximation model [81]. The exchange and correlation energy $K(\rho)$ of local spin density approximation can be separated into individual contributions from the exchange part $K_x(\rho)$ and the correlation part $K_c(\rho)$,

$$K(\rho) = K_x(\rho) + K_c(\rho), \tag{2.3.33}$$

where

$$K_x(\rho) = -\frac{3}{4}\left(\frac{6}{\pi}\right)^{1/3} \int (\rho_\alpha^{3/4}(\vec{x}) + \rho_\beta^{3/4})(\vec{x}) d\vec{x}, \tag{2.3.34}$$

and

$$K_c(\rho) = \int \mathbb{k}_c(\rho(\vec{x})) d\vec{x}. \tag{2.3.35}$$

People always use the parameterization of [122] fitted to accurate Monte Carlo simulations carried out by [24] to approximate $K_c(\rho)$.

Local spin density approximation is very accurate for solids, and it is still widely used in condensed matter physics. It is less useful for atoms and molecules, which have a nonuniform electron gas.

By the method of the Lagrange multiplier, we multiply each (2.3.22) by the Lagrangian multiplier $-\epsilon_{ji}$ and add to the equation (2.3.28). By the similar process in restricted Hartree-Fock, we obtain eigenvalue equation for spin-up electrons

$$\begin{aligned} \widehat{F_\alpha}\phi_i &= \{\widehat{T}_i + \widehat{V}_i + \sum_{j=1}^{n_\alpha+n_\beta} \widehat{J}_j + \widehat{K_{i,\alpha}}\}\phi_i, \\ &= \epsilon_i\phi_i, i = 1, \ldots, n_\alpha, \end{aligned} \tag{2.3.36}$$

and eigenvalue equation for spin-down electrons

$$\begin{aligned} \widehat{F_\beta}\phi_i &= \{\widehat{T}_i + \widehat{V}_i + \sum_{j=1}^{n_\alpha+n_\beta} \widehat{J}_j + \widehat{K_{i,\beta}}\}\phi_i, \\ &= \epsilon_i\phi_i, i = n_\alpha + 1, \ldots, n_\alpha + n_\beta, \end{aligned} \tag{2.3.37}$$

where $\widehat{T}_i, \widehat{V}_i, \widehat{J}_j$ have been defined in (2.2.23) and

$$\widehat{K_{i,\alpha}} = \frac{\partial \mathbb{k}(\rho)}{\partial \rho_\alpha},$$
$$\widehat{K_{i,\beta}} = \frac{\partial \mathbb{k}(\rho)}{\partial \rho_\beta}, \tag{2.3.38}$$

2.3.2 Generalized Gradient Approximation

Although many generalizations of the local spin density approximation were proposed, the first practical one was the generalized gradient approximation [118, 83, 10, 86, 123, 117, 119], which introduces $\bigtriangledown\rho$ as additional local ingredients of $K(\rho)$. Thus,

$$K(\rho, \bigtriangledown\rho) = \int_\Omega \mathbb{k}(\rho, \bigtriangledown\rho) d\vec{x}. \tag{2.3.39}$$

The original motivation of generalized gradient approximation is the second-order gradient expansion approximation (GEA). The expression of the second-order gradient expansion approximation is:

$$\begin{aligned} K(\rho, \bigtriangledown\rho) = \quad & \int \left[\rho_\alpha^{4/3}(\vec{x}) + \rho_\beta^{4/3}(\vec{x}) + C_{\alpha\alpha} \frac{(\bigtriangledown\rho_\alpha(\vec{x}))^2}{\rho_\alpha^{4/3}(\vec{x})} \right. \\ & \left. + C_{\beta\beta} \frac{(\bigtriangledown\rho_\beta(\vec{x}))^2}{\rho_\beta^{4/3}(\vec{x})} + C_{\alpha\beta} \frac{\bigtriangledown\rho_\alpha(\vec{x}) \bigtriangledown\rho_\beta(\vec{x})}{\rho_\alpha^{2/3}(\vec{x}) \rho_\beta^{2/3}(\vec{x})} \right] d\vec{x}. \end{aligned} \tag{2.3.40}$$

The second-order gradient expansion approximation is valid for slowly varying densities. The coefficients are derived in the hope that gradient expansion approximation could improve local spin density approximation for real solids and even for molecules.

We can decompose exchange-correlation hole ρ_{xc} in (2.1.39) into exchange and correlation contributions, then

$$\rho_{xc}(\vec{x}_1, \vec{x}_2) = \rho_x(\vec{x}_1, \vec{x}_2) + \rho_c(\vec{x}_1, \vec{x}_2). \tag{2.3.41}$$

From the definition of the exact exchange energy functional (2.2.21), we can define the exchange hole as

$$\rho_x(\vec{x}_1, \vec{x}_2) = -\frac{1}{2} \frac{|\rho_1(\vec{x}_1, \vec{x}_1)|^2}{\rho(\vec{x}_1)}. \tag{2.3.42}$$

Thus, the exchange hole satisfies

$$\int \rho_x(\vec{x}_1, \vec{x}_2) d\vec{x}_2 = -1, \tag{2.3.43}$$

and the correlation hold satisfies

$$\int \rho_c(\vec{x}_1, \vec{x}_2) d\vec{x}_2 = 0, \tag{2.3.44}$$

PBE-GGA [118] is another widely used non-empirical generalized gradient approximation model, which satisfies (2.3.43) and (2.3.44). $K(\rho, \triangledown\rho)$ includes exchange energy part $K_x(\rho, \triangledown\rho)$ and correlation energy part $K_c(\rho, \triangledown\rho)$. The expressions are:

$$K_x(\rho, \triangledown\rho) = \int \rho \varepsilon_X^{unif}(\rho) F_X(s) d\vec{x}, \tag{2.3.45}$$

where

$$F_X(s) = 1 + \kappa - \frac{\kappa}{1 + \mu s^2/\kappa}. \tag{2.3.46}$$

$$K_c(\rho, \triangledown\rho) = \int \rho[\varepsilon_C^{unif}(r_s, \zeta) + H(r_s, \zeta, t)] d\vec{x}, \tag{2.3.47}$$

where

$$H = \frac{\gamma\phi^3}{a_0} \ln\left\{1 + \frac{\beta}{\gamma} t^2 \left[\frac{1 + At^2}{1 + At^2 + A^2t^4}\right]\right\}. \tag{2.3.48}$$

All the parameters and explicit expressions could be found in [118].

By the method of the Lagrange multiplier and similar process in local spin density approximation, we obtain eigenvalue equation for spin-up electrons

$$\begin{aligned} \widehat{F_\alpha}\phi_i &= \{\widehat{T}_i + \widehat{V}_i + \sum_{j=1}^{n_\alpha+n_\beta} \widehat{J}_j + \widehat{K_{i,\alpha}}\}\phi_i, \\ &= \epsilon_i\phi_i, i = 1, \ldots, n_\alpha, \end{aligned} \tag{2.3.49}$$

and eigenvalue equation for spin-down electrons

$$\begin{aligned} \widehat{F_\beta}\phi_i &= \{\widehat{T}_i + \widehat{V}_i + \sum_{j=1}^{n_\alpha+n_\beta} \widehat{J}_j + \widehat{K_{i,\beta}}\}\phi_i, \\ &= \epsilon_i\phi_i, i = n_\alpha + 1, \ldots, n_\alpha + n_\beta, \end{aligned} \tag{2.3.50}$$

where $\widehat{T}_i, \widehat{V}_i, \widehat{J}_j$ have been defined in (2.2.23) and

$$\begin{aligned} \widehat{K_{i,\alpha}} &= \frac{\partial \mathbb{k}(\rho, \triangledown\rho)}{\partial\rho_\alpha} - \triangledown\left(\frac{\partial \mathbb{k}(\rho, \triangledown\rho)}{\partial\triangledown\rho_\alpha}\right), \\ \widehat{K_{i,\beta}} &= \frac{\partial \mathbb{k}(\rho, \triangledown\rho)}{\partial\rho_\beta} - \triangledown\left(\frac{\partial \mathbb{k}(\rho, \triangledown\rho)}{\partial\triangledown\rho_\beta}\right). \end{aligned} \tag{2.3.51}$$

2.3.3 Meta Generalized Gradient Approximation

Meta generalized gradient approximation [116, 156, 9, 33, 127, 121, 149] assumes $K(\rho)$ is a function of ρ, $\nabla\rho$, $\nabla^2\rho$ and $\tau = \frac{1}{2}\sum|\nabla\psi_i|^2$. Thus,

$$K(\rho, \nabla\rho, \nabla^2\rho, \tau) = \int_\Omega \Bbbk(\rho, \nabla\rho, \nabla^2\rho, \tau)d\vec{x}, \tag{2.3.52}$$

The Laplacians $\nabla^2\rho$ seem like the more natural next step, since they appear in the fourth-order gradient expansion. τ appears in the Taylor expansion of the exchange hole density about $|\vec{x}' - \vec{x}| = 0$.

The most famous meta generalized gradient approximation model is TPSS-Meta-GGA [149], where the exchange energy part is:

$$K_x(\rho, \nabla\rho, \nabla^2\rho, \tau) = \int_\Omega \rho\varepsilon_X^{unif}(\rho)F_X(p, z)d\vec{x}, \tag{2.3.53}$$

and the correlation energy part is:

$$K_c(\rho, \nabla\rho, \nabla^2\rho, \tau) = \int_\Omega \rho\varepsilon_c^{revPKZB}[1 + d\varepsilon_c^{revPKZB}(\tau^W/\tau)^3]d\vec{x}. \tag{2.3.54}$$

All the parameters and explicit expressions could be found in [149].

2.3.4 Generalized Density Functional Theory Models

For more general models, $K(\rho)$ is considered as a functional of $\rho, \nabla\rho, \ldots, \nabla^k\rho$. Thus

$$K(\rho, \nabla\rho, \nabla^2\rho, \ldots, \nabla^k\rho) = \int_\Omega \Bbbk(\rho, \nabla\rho, \nabla^2\rho, \ldots, \nabla^k\rho)d\vec{x}. \tag{2.3.55}$$

This is natural to introduce higher order derivatives of ρ into $K(\rho)$.

2.4 Self-Consistent Method

Both Hartree-Fock method and density functional theory are solved by Self-consistent iterative methods. Section 2.4.1 and Section 2.4.2 will give an introduction to Self-consistent iterative methods, and how they are applied to Hartree-Fock method and density functional theory.

2.4.1 Hartree-Fock

The general procedure for solving (2.2.39), (2.2.44) and (2.2.45) is Self-consistent method [129]. One begins with initial guesses on ϕ_i's, and solves corresponding effective one-electron eigenvalue equation. Then we get the n lowest eigenvalues for (2.2.39), n_α lowest eigenvalues for (2.2.44) and n_β lowest eigenvalues for (2.2.45). Calculate the difference between resulting ϕ_i's and ϕ_i's calculated in last step. This process is then repeated until the difference is smaller than the error threshold.

2.4.2 Density Functional Theory

The general procedure for solving (2.3.36), (2.3.37), (2.3.49) and (2.3.50) are very similar with Hartree-Fock [81]. One begins with an assumed ρ, and constructs corresponding $\widehat{J}$ and $\widehat{K}$. Then solve corresponding effective one-electron eigenvalue equation, and get the n_α lowest eigenvalues for (2.3.36) and (2.3.49) and the n_β lowest eigenvalues for (2.3.37) and (2.3.50). A new electronic density ρ can be constructed, and calculate the difference between resulting ρ and ρ calculated in last step. This process is then repeated until the difference is smaller than the error threshold.

Chapter 3

Analysis of Continuous and Discrete Electronic Structure Models

This chapter gives a complete mathematical analysis on Hartree-Fock and density functional theory models [146, 67]. Section 3.1 shows the motivation why we feel interested in the mathematical properties of these models. The physical system and some formulations are defined in Section 3.2. The definitions, theorems and notations, which are used in our proofs, are given in Section 3.3. In Section 3.4, we prove the existence of a minimizer for these models. In Section 3.5 and Section 3.6, we prove the convergence of finite element approximation and the convergence of finite element approximation with numerical quadratures. The convergence of pseudopotential approximation is shown in Section 3.7. Section 3.8 is our conclusion.

The paper of Suryanarayana et al. [146] presented the existence of a minimizer, the convergence of finite element approximation, the convergence of finite element approximation with numerical quadratures and the convergence of pseudopotential approximation for local spin density approximation model. We generalize the proofs of Suryanarayana et al. [146] to restricted Hartree-Fock, unrestricted Hartree-Fock, generalized gradient approximation model, meta generalized gradient approximation model and generalized density functional theory models [67].

3.1 Motivation

Though countless papers have been published in solving Hartree-Fock and density functional theory problems, only few papers [3, 46, 84, 146] discuss the mathematical properties of these models, such as the existence of a minimizer, the convergence of finite element approximation, and the convergence of numerical method.

We generalize the work of Suryanarayana et al. [146] on local spin density approximation model, and develop a mathematical framework for Hartree-Fock models and density functional theory models. Under certain conditions, we show the existence of the solution, and the convergence of finite element approximations including numerical quadratures and pseudopotential approximations [67]. We view this work as a fundamental and important step to analyze these models from a mathematical point of view and it will also have great value for future physics models and numerical methods design.

3.2 Physical System and Formulation

We will introduce the external field V_{ext} in Section 3.2.1 [146, 67]. We assume the nuclear charge is uniformly distributed in a small ball in Section 3.2.2 [146, 67]. We also rewrite some equations in Chapter 2 in Section 3.2.3 [146, 67].

3.2.1 External Field

Suppose electrons are moving in the external field generated by M nuclei, and the external field is given by:

$$V_{ext}(\vec{x}) = \sum_{i=1}^{M} \frac{z_i}{|\vec{x} - \vec{x}_{z_i}|}, \tag{3.2.1}$$

where z_i represents the nuclear charge magnitude of the ith nucleus and $\vec{x}_{z_i}$ represents the nuclear postion of the ith nucleus.

The repulsive energy among the nuclei is:

$$B = \frac{1}{2} \sum_{i,j=1}^{M} \frac{z_i z_j}{|\vec{x}_{z_i} - \vec{x}_{z_j}|}, \tag{3.2.2}$$

which is an inconsequential constant depending only on the nuclear charges.

3.2.2 Ball Approximation for Singularity

We follow the work of Suryanarayana et al. [146] and use a small ball approximation $b(x)$ to smooth the singularities of nucleus. The external potential could be rewritten as:

$$V_{ext} = \int \frac{b(\vec{x}')}{|\vec{x} - \vec{x}'|} d\vec{x}', \tag{3.2.3}$$

where

$$b(\vec{x}) = -\sum_{i=1}^{M} z_i \delta_{\vec{x}_{z_i}}(\vec{x}). \tag{3.2.4}$$

Here z_i is the nuclear charge magnitude at $\vec{x}_{z_i}$, $-z_i\delta_{\vec{x}_{z_i}}(\vec{x})$ has a compact support in a neighborhood around $\vec{x}_{z_i}$ and total charge $-z_i$.

The repulsive energy among the nuclei (3.2.2) is:

$$B = \frac{1}{2} \int \int \frac{b(\vec{x})b(\vec{x}')}{|\vec{x} - \vec{x}'|} d\vec{x} d\vec{x}', \tag{3.2.5}$$

3.2.3 Formulation

In order to simplify the notations and proofs, we follow Section 2 of Suryanarayana et al. paper [146] and rewrite some equations in Chapter 2 [67].

Restricted Hartree-Fock

The electron density can be defined similarly as density functional theory:

$$\rho(\vec{x}) = \sum_{i=1}^{2n} \phi_i^*(\vec{x})\phi_i(\vec{x}) = \sum_{i=1}^{2n} |\phi_i(\vec{x})|^2, \tag{3.2.6}$$

where

$$\phi_{n+i} = \phi_i, i = 1, 2, \ldots, n \tag{3.2.7}$$

Because of the orthonormality of spin factors, (2.2.18), (2.2.19), (2.2.20), (2.2.21) can be written as:

$$T(\Phi) = 2\sum_{i=1}^{n} T_i = \sum_{i=1}^{n} \int \phi_i^*(\vec{x}) \nabla^2 \phi_i(\vec{x}) d\vec{x} = \frac{1}{2} \sum_{i=1}^{2n} \int \phi_i^*(\vec{x}) \nabla^2 \phi_i(\vec{x}) d\vec{x}. \tag{3.2.8}$$

$$
\begin{aligned}
V(\Phi) &= 2\sum_{i=1}^{n} V_i = 2\sum_{i=1}^{n}\int \phi_i^*(\vec{x}) V_{ext}(\vec{x})\phi_i(\vec{x})d\vec{x}, \\
&= \sum_{i=1}^{2n}\int \phi_i^*(\vec{x}) V_{ext}(\vec{x})\phi_i(\vec{x})d\vec{x} = \int\int \frac{\rho(\vec{x})b(\vec{x}')}{|\vec{x}-\vec{x}'|}d\vec{x}d\vec{x}'. \qquad (3.2.9)
\end{aligned}
$$

$$
\begin{aligned}
J(\Phi) &= 2\sum_{i,j=1}^{n} J_{ij} = 2\sum_{i,j=1}^{n}\int\int \frac{\phi_i^*(\vec{x})\phi_j^*(\vec{x}')\phi_i(\vec{x})\phi_j(\vec{x}')}{|\vec{x}-\vec{x}'|}d\vec{x}d\vec{x}', \\
&= \frac{1}{2}\sum_{i,j=1}^{2n}\int\int \frac{\phi_i^*(\vec{x})\phi_j^*(\vec{x}')\phi_i(\vec{x})\phi_j(\vec{x}')}{|\vec{x}-\vec{x}'|}d\vec{x}d\vec{x}', \\
&= \frac{1}{2}\int\int \frac{\rho(\vec{x})\rho(\vec{x}')}{|\vec{x}-\vec{x}'|}d\vec{x}d\vec{x}'. \qquad (3.2.10)
\end{aligned}
$$

$$
\begin{aligned}
K(\Phi) &= -\sum_{i,j=1}^{n} K_{ij} = -\sum_{i,j=1}^{n}\int\int \frac{\phi_i^*(\vec{x})\phi_j^*(\vec{x}')\phi_j(\vec{x})\phi_i(\vec{x}')}{|\vec{x}-\vec{x}'|}d\vec{x}d\vec{x}', \\
&= -\frac{1}{4}\sum_{i,j=1}^{2n}\int\int \frac{\phi_i^*(\vec{x})\phi_j^*(\vec{x}')\phi_j(\vec{x})\phi_i(\vec{x}')}{|\vec{x}-\vec{x}'|}d\vec{x}d\vec{x}'. \qquad (3.2.11)
\end{aligned}
$$

Following the Equation (17) of Suryanarayana et al. paper [146], we have

$$
\begin{aligned}
& V(\Phi) + J(\Phi) + B, \\
= \quad & \int\int \frac{\rho(\vec{x})b(\vec{x}')}{|\vec{x}-\vec{x}'|}d\vec{x}d\vec{x}' + \frac{1}{2}\int\int \frac{\rho(\vec{x})\rho(\vec{x}')}{|\vec{x}-\vec{x}'|}d\vec{x}d\vec{x}' + \frac{1}{2}\int\int \frac{b(\vec{x})b(\vec{x}')}{|\vec{x}-\vec{x}'|}d\vec{x}d\vec{x}', \\
= \quad & -\inf_{\varphi}\left\{\frac{1}{8\pi}\int |\nabla\varphi(\vec{x})|^2 d\vec{x} - \int(\rho(\vec{x}) + b(\vec{x}))\varphi(\vec{x})d\vec{x}\right\}, \qquad (3.2.12)
\end{aligned}
$$

where

$$
\varphi(x) = \int \frac{\rho(\vec{x}')}{|\vec{x}-\vec{x}'|}dx' + \int \frac{b(\vec{x}')}{|\vec{x}-\vec{x}'|}d\vec{x}' = V_\rho(\vec{x}) + V_{ext}(\vec{x}). \qquad (3.2.13)
$$

Also the exact exchange energy term (3.2.11) can be written as:

$$
\begin{aligned}
K(\Phi) &= -\frac{1}{4}\sum_{i,j=1}^{2n}\int\int \frac{\phi_i^*(\vec{x})\phi_j^*(\vec{x}')\phi_j(\vec{x})\phi_i(\vec{x}')}{|\vec{x}-\vec{x}'|}d\vec{x}d\vec{x}', \\
&= -\frac{1}{4}\sum_{i,j=1}^{2n}\int \phi_i^*(\vec{x})\phi_j(\vec{x})\int \frac{\phi_i(\vec{x}')\phi_j^*(\vec{x}')}{|\vec{x}-\vec{x}'|}d\vec{x}'d\vec{x}. \qquad (3.2.14)
\end{aligned}
$$

If we define $\varphi_{K_{ij}}(\vec{x})$ as:

$$\varphi_{K_{ij}}(\vec{x}) = \int \frac{\phi_i(\vec{x}')\phi_j^*(\vec{x}')}{|\vec{x} - \vec{x}'|} d\vec{x}', \tag{3.2.15}$$

then

$$-\frac{1}{4\pi}\nabla^2\varphi_{K_{ij}}(\vec{x}) = \phi_i(\vec{x})\phi_j^*(\vec{x}), \tag{3.2.16}$$

and we have:

$$K(\Phi) = \frac{1}{2}\sum_{i,j=1}^{2n} \min_{\varphi_{K_{ij}}} \left\{ \int [\frac{1}{8\pi}(\nabla\varphi_{K_{ij}}(\vec{x}))^2 - \phi_i^*(\vec{x})^*\phi_j(\vec{x})\varphi_{K_{ij}}(\vec{x})]dx \right\}. \tag{3.2.17}$$

The total energy functional $\mathcal{E}$ can be written in the following form:

$$\mathcal{E}(\Phi) = T(\Phi) + \mathcal{J}(\Phi) + K(\Phi), \tag{3.2.18}$$

where

$$\mathcal{J}(\Phi) = -\min_{\varphi} \left\{ \frac{1}{8\pi} \int |\nabla\varphi(\vec{x})|^2 d\vec{x} - \int (\rho(\vec{x}) + b(\vec{x}))\varphi(\vec{x}) d\vec{x} \right\}. \tag{3.2.19}$$

Unrestricted Hartree-Fock

The electron density can be defined similarly as previously in density functional theory:

$$\rho(\vec{x}) = \sum_{i=1}^{n_\alpha+n_\beta} \phi_i^*(\vec{x})\phi_i(\vec{x}) = \sum_{i=1}^{n_\alpha+n_\beta} |\phi_i(\vec{x})|^2 \tag{3.2.20}$$

Total energy (2.2.43) can be rewritten similarly as done in restricted Hartree-Fock:

$$\mathcal{E} = T(\Phi) + \mathcal{J}(\Phi) + K(\Phi), \tag{3.2.21}$$

where

$$T(\Phi) = \frac{1}{2}\sum_{i=1}^{n_\alpha+n_\beta} \int \phi_i^*(\vec{x})\nabla^2\phi_i(\vec{x}) d\vec{x}, \tag{3.2.22}$$

$$K(\Phi) = -\frac{1}{2}\left[\sum_{i,j=1}^{n_\alpha} + \sum_{i,j=n_\alpha+1}^{n_\alpha+n_\beta}\right] \int\int \frac{\phi_i^*(\vec{x})\phi_j^*(\vec{x}')\phi_j(\vec{x})\phi_i(\vec{x}')}{|\vec{x} - \vec{x}'|} d\vec{x}d\vec{x}', \tag{3.2.23}$$

and $\mathcal{J}(\Phi)$ (3.2.19).

We can rewrite the exchange energy term (3.2.23) similarly as done in restricted Hartree-Fock:

$$K(\Phi) = \frac{1}{2}\left[\sum_{i,j=1}^{n_\alpha} + \sum_{i,j=n_\alpha+1}^{n_\alpha+n_\beta}\right] \min_{\varphi_{K_{ij}}} \left\{ \int [\frac{1}{8\pi}(\nabla\varphi_{K_{ij}}(\vec{x}))^2 - \phi_i^*(\vec{x})\phi_j(\vec{x})\varphi_{K_{ij}}(\vec{x})]dx \right\}, \tag{3.2.24}$$

where

$$\varphi_{K_{ij}}(\vec{x}) = \int \frac{\phi_i(\vec{x}')\phi_j^*(\vec{x}')}{|\vec{x} - \vec{x}'|} d\vec{x}'. \tag{3.2.25}$$

Density Functional Theory

Total energy functional $\mathcal{E}$ (2.3.28) can be written as:

$$\mathcal{E}(\rho) = T(\rho) + \mathcal{J}(\rho) + K(\rho), \tag{3.2.26}$$

where

$$\mathcal{J}(\rho) = -\min_{\varphi} \left\{ \frac{1}{8\pi} \int |\nabla\varphi(\vec{x})|^2 d\vec{x} - \int (\rho(\vec{x}) + b(\vec{x}))\varphi(\vec{x}) d\vec{x} \right\}, \tag{3.2.27}$$

φ (3.2.13), $T(\rho)$ (2.3.29) and $K(\rho)$ (2.3.32).

3.3 Definition and Theorem

Definitions and theorems are presented in Section 3.3.1 and Section 3.3.2, and Section 3.3.3 gives us function space notations [1, 35, 144, 31, 16, 49, 76, 28]. All the definitions, theorems and notations in this section are highly correlated with the proofs in this chapter.

3.3.1 Definition

Euclidean n-space is called Cartesian space, which is the space of all n-turples of real numbers. A Banach space is a complete normed vector space. A functional is a real-valued function on a normed vector space. Euclidean space, Banach space and nonlinear functional notations are given in Definition 3.3.1.

Definition 3.3.1 ([1, 35])**.**

$\mathbb{R}^n$ *denotes* n*-dimensional real Euclidean space*.

$\bar{\mathbb{R}} = \mathbb{R} \cup \{+\infty, -\infty\}$.

$\mathbb{R}^+ = \{x \in \mathbb{R} | x > 0\}$.

U, V *denote subsets of* $\mathbb{R}^n$.

Z *denotes a Banach space, and* Y *is a subset of* Z.

$F : Y \to \bar{\mathbb{R}}$ *is a generally nonlinear functional*.

A function space is a topological vector space, consisting of a set of functions satisfied certain properties. The notations of function spaces are defined in Definition 3.3.2.

Definition 3.3.2 ([1, 35])**.** *Function spaces :*

$C(U) = \{u : U \to \mathbb{R} | u$ *is continuous*$\}$.

$C^k(U) = \{u : U \to \mathbb{R} | u$ *is* k*-times continuously differentiable*$\}$

$C^\infty(U) = \{u : U \to \mathbb{R} | u$ *is infinitely differentiable*$\}$.

$C_c(U), C_c^k(U),$ *etc. denote these functions in* $C(U), C^k(U)$*, etc. with compact support*.

$L^p(U) = \{u : U \to \mathbb{R} | u$ *is Lebesgue measurable,* $\|u\|_{L^p(U)} < +\infty\}$.

where $\|u\|_{L^p(U)} = \begin{cases} \left(\int_U |u|^p d\vec{x}\right)^{\frac{1}{p}}, \ (1 \le p < +\infty), \\ ess\sup_U |u|, \ (p = +\infty), ess\sup \text{ means essential supremum}. \end{cases}$

$L^p_{loc}(U) = \{u : U \to \mathbb{R} | u \in L^p(V)$ *for each* $V \subset\subset U\}$,

where $V \subset\subset U$ *means* V *is compactly contained in* U.

A weak derivative is a generalization of the derivative concept of functions, which are not differentiable, but only integrable. Definition 3.3.3 gives the definition of weak derivation.

Definition 3.3.3 ([1, 35])**.** *(Weak Derivatives). Suppose* $u, v \in L^1_{loc}(U)$ *and* α *is a multiindex. We say that* v *is* α*th-weak partial derivative of* u*, written*

$$D^\alpha u = v, \tag{3.3.1}$$

provided

$$\int_U uD^\alpha \vartheta d\vec{x} = (-1)^{|\alpha|} \int_U v\vartheta d\vec{x} \tag{3.3.2}$$

for all test functions $\vartheta \in C_c^\infty(U)$.

In partial differential equation theory, as it is hard to prove the solutions we construct belong to some well-behavioral space, we need to define some other kinds of spaces for less smooth functions. We design Sobolev spaces, consisting of functions with some, but not great smooth properties. A Sobolev space, defined in Definition 3.3.4, is a vector space of functions with a norm that is a combination of L^p norms of the function itself as well as its derivatives.

Definition 3.3.4 ([1, 35]). *(Sobolev Spaces). Let* $1 \leq p \leq +\infty$ *and* k *be a nonnegative integer. The Sobolev space*

$$W^{k,p}(U) \tag{3.3.3}$$

consists of all locally summable functions $u : U \to \mathbb{R}$ *such that for each multiindex* α *with* $|\alpha| \leq k$, $D^\alpha u$ *exists in the weak sense and belongs to* $L^p(U)$.
If $p = 2$, *we usually write*

$$H^k(U) = W^{k,2}(U). \tag{3.3.4}$$

Norms with derivatives are nature extensions of L^p norms, which is used by Sobolev space, and it needs more strict conditions than L^p norms. The norms of Sobolev spaces are defined in Definition 3.3.5 and Definition 3.3.6,

Definition 3.3.5 ([1, 35]). *Let* $\Phi = (\phi_i)_{i=1}^N \in (W^{k,p}(U))^N$. *We define*

$$\|\Phi\|_{W^{k,p}(U)} = \left(\sum_{i=1}^N \|\phi_i\|_{W^{k,p}(U)}^p \right)^{\frac{1}{p}}. \tag{3.3.5}$$

Definition 3.3.6 ([1, 35]). *If* $u \in W^{k,p}(U)$, *we define its norm to be*

$$\|u\|_{W^{k,p}(U)} = \begin{cases} \left(\sum_{|\alpha| \leq k} \int_U |D^\alpha u|^p d\vec{x} \right)^{\frac{1}{p}}, (1 \leq p < +\infty), \\ \sum_{|\alpha| \leq k} ess\sup_U |D^\alpha u|, (p = +\infty). \end{cases} \tag{3.3.6}$$

For optimization problems, we always design the convex objective functionals, or we need to transform a non-convex functional to a convex functional by adding some constraints. The definition of a convex functional is in Definition 3.3.7,

Definition 3.3.7 ([144, 35]). *$\forall u, v \in Y$ and $\alpha \in [0, 1]$, if*

$$F(\alpha u + (1 - \alpha)v) \leq \alpha F(u) + (1 - \alpha)F(v) \tag{3.3.7}$$

satisfied, then F is convex.

A sequence is said to be convergent if it approaches some limit. Definition 3.3.8 shows how the convergence of a function sequence is defined in a function space.

Definition 3.3.8 ([144, 35]). *We say a sequence $\{u_j\} \in Y$ converges to $u \in Y$, written*

$$u_j \to u, \tag{3.3.8}$$

when $\forall \varepsilon > 0$, if there exists N, then $\|u - u_j\|_Y < \varepsilon$ satisfied for all $j > N$.

Semi-continuity is a generalized property of continuity, which allows a more general class of functionals. Semi-continuity is weaker than continuity, and it is defined in Definition 3.3.9.

Definition 3.3.9 ([144, 35]). *$\forall$ sequence $\{u_j\} \in Y$ and $u_j \to u$, if*

$$F(u) \leq \lim_{j->\infty} \inf F(u_j) \tag{3.3.9}$$

satisfied, then F is lower semi-continuous.

Coercivity is a class of functionals, which are bounded by below. Definition 3.3.10 defines the coercivity of a functional.

Definition 3.3.10 ([144, 35]). *$\forall u \in Y$, if*

$$F(u) \geq C_0 \|u\|_Y^2 - C_1, \tag{3.3.10}$$

where C_0 and C_1 are constants, then F is coercive in Y.

The dual space of a function space is a collection of functionals, and the weak topology is defined using the dual space. The definitions of a bounded linear operator and dual space are given in Definition 3.3.11.

Definition 3.3.11 ([144, 35]). *A bounded linear operator $u^* : Y \to \mathbb{R}$ is called a bounded linear functionals on Y. We write Y^* to denote the collection of all bounded linear functionals on Y; Y^* is the dual space of Y.*

Weak convergence is the convergence of a sequence in the weak topology of a space. Every strongly convergent sequence is also weakly convergent. Definition 3.3.12 shows the definition of weak convergence in a function space.

Definition 3.3.12 ([144, 35]). *We say a sequence $\{u_j\} \in Y$ converges weakly to $u \in Y$, written*

$$u_j \rightharpoonup u, \tag{3.3.11}$$

when $\forall f \in Y^*$, $f(u_j) \to f(u)$.

Weakly lower semi-continuity is lower semi-continuity in the weak topology of a space. Contrary to weak convergence, weakly lower semi-continuity is a stronger condition than lower semi-continuity, which means every weakly lower semi-continuous sequence is also lower semi-continuous. Definition 3.3.13 gives us the definition of weakly lower semi-continuity.

Definition 3.3.13 ([144, 35]). *$\forall$ sequence $\{u_j\} \in Y$ and $u_j \rightharpoonup u$, if*

$$F(u) \leq \lim \inf_{j->\infty} F(u_j) \tag{3.3.12}$$

satisfied, then F is weakly lower semi-continuous.

Weak coercivity is the coercivity in the weak topology of a space, and it is a weaker condition than strong coercivity, which is defined in Definition 3.3.14.

Definition 3.3.14 ([31]). *If Y is reflective. $\forall u \in Y, F(u)$ tends to $+\infty$ as $\|u\|_Y$ tends to $+\infty$, then F is coercive in the weak topology of Y.*

Among variational convergences, Γ-convergence plays an important role for its compactness properties and the large number of results concerning Γ-limits of integral functionals. Also we can express all other variational convergence easily in the language of Γ-convergence. Γ-convergence of a sequence of a functions is defined in Definition 3.3.15.

Definition 3.3.15 ([31, 16]). *The sequence (F_h) Γ-converges to a functional F in Y, if for any sequence $(u_h) \to u$ in Y, we have $\liminf_{h\to 0} F_h(u_h) \geq F(u)$ and a recovery sequence can be constructed.*

Equi-coercive sequence is an important assumption in Γ-convergence analysis. Definition 3.3.16 defines weak equi-coercive sequence.

Definition 3.3.16 ([31, 16]). *The sequence (F_h) is equi-coercive in the weak topology of Y, if there exits a weakly coercive function $\Lambda : Y \to \bar{\mathbb{R}}$ independent of h such that $F_h \geq \Lambda$ on Y for every h.*

3.3.2 Theorem

Weak lower semi-continuity is a stronger condition than lower semi-continuity, and Theorem 3.3.17 shows how we could get weakly lower semi-continuity of a functional from lower semi-continuity.

Theorem 3.3.17 ([144]). *If F is convex and lower semi-continous, then F is weakly lower semi-continuous.*

Theorem 3.3.18 is an inequality, which connects the first order derivative of a function and the first order derivative of the absolute value of this function. It will be used to bound the kinetic energy term.

Theorem 3.3.18 ([49]). $|\nabla|\phi_i|| \leq |\nabla\phi_i|$, *if* $\phi_i \in H_0^1(\Omega)$.

Theorem 3.3.19 gives inclusions among L^p spaces, and inequalities among norms of L^p spaces.

Theorem 3.3.19 ([76]). *Let U be a finite bounded domain and $1 \leq p < q \leq +\infty$. For any function f, if $f \in L^p(U), L^q(U)$, then*

$$\|f\|_{L^p(U)} \leq C\|f\|_{L^q(U)}, \tag{3.3.13}$$

$$L^q(U) \subset L^p(U), \tag{3.3.14}$$

where C is a positive constant.

Theorem 3.3.20 shows the properties between function spaces, where one function space is compactly embedded in the other function space.

Theorem 3.3.20 ([1]). *If $Y \subset\subset Z$, then $\|u\|_Z \leq C\|u\|_Y$ for a constant C, and each bounded sequence in Z is precompact in Y.*

Theorem 3.3.21 shows the error associated with n-th numerical quadrature, which is our starting point to prove the convergence of finite element convergence with numerical quadratures.

Theorem 3.3.21 ([28]). *Suppose $F = \int f(\vec{x})d\vec{x}$. If the quadrature rule is n-th order, then the error associated with numerical quadrature $\widetilde{F}$ is given by*

$$|\widetilde{F} - F| \leq C^{n+1} \int |D^{n+1} f(\vec{x})| d\vec{x}. \tag{3.3.15}$$

Young's Inequality, Hölder inequality and Poincar$\acute{e}$ inequality are standard tools in our proofs, and they are described carefully in Theorem 3.3.22, Theorem 3.3.23 and Theorem 3.3.24.

Theorem 3.3.22 ([1, 35]). *(Young's Inequality). Let $1 < p, q < +\infty$, $\frac{1}{p} + \frac{1}{q} = 1$. Then*

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}, \ (a, b > 0). \tag{3.3.16}$$

Theorem 3.3.23 ([1, 35]). *(Hölder Inequality). Assume $1 \leq p, q \leq +\infty$, $\frac{1}{p} + \frac{1}{q} = 1$. Then if $u \in L^p(U)$, $v \in L^q(U)$, we have*

$$\int_U |uv| d\vec{x} \leq \|u\|_{L^p(U)} \|v\|_{L^q(U)}. \tag{3.3.17}$$

Theorem 3.3.24 ([1, 35]). *(Poincar$\acute{e}$ Inequality). Assume that $1 \leq p \leq +\infty$ and U is bounded with a Lipchitz boundary. Then there exists a constant C depending only on U and p, such that for every function u in the Sobolev Space $W^{1,p}(U)$,*

$$\|u\|_{L^p(U)} \leq C \|\nabla u\|_{L^p(U)}. \tag{3.3.18}$$

Inverse inequality relates the norms of functions in finite element approximate spaces, which is useful in finite element analysis. Theorem 3.3.25 gives a clear description of inverse inequality.

Theorem 3.3.25 ([28]). *(Inverse Inequality). Let T_h be a family of triangulations of mesh size $0 < h \leq 1$. Assume that Y_h is a finite dimensional subspace of Y consisting of functions whose restriction to every cell in T_h is a polynomial, and Y_h is a subspace of $W^{l,p}$ and $W^{m,q}$, where $1 \leq p \leq +\infty$, $1 \leq q \leq +\infty$ and $0 \leq m \leq l$. Then there exists a constant $C = C(l, p, q, n, Y)$ such that for all $u \in Y_h$.*

$$\|u\|_{W^{l,p}} \leq C h^{m-l+n/p-n/q} \|\nabla u\|_{W^{m,q}}. \tag{3.3.19}$$

Sobolev embedding theorem gives inclusions between certain Sobolev spaces, and shows that Sobolev spaces are compactly embedded in others under stronger conditions.

Theorem 3.3.26 ([1, 35]). *(Sobolev Embedding Theorem). Let $j \geq 0$, $m \geq 1$, $1 \leq p < +\infty$ and j, m, p, q be integers. Let n be the dimension of U.*
If $n > mp$, then

$$\begin{aligned} &W^{j+m,p}(U) \subset W^{j,q}(U), (1 \leq q \leq +\infty) \\ &W^{j+m,p}(U) \subset\subset W^{j,q}(U), (1 \leq q < +\infty) \end{aligned} \tag{3.3.20}$$

If $n = mp$, then

$$\begin{aligned} &W^{j+m,p}(U) \subset W^{j,q}(U), (1 \leq q < +\infty) \\ &W^{j+m,p}(U) \subset\subset W^{j,q}(U), (1 \leq q < +\infty) \end{aligned} \tag{3.3.21}$$

If $0 < n - mp < n$, then

$$\begin{aligned} &W^{j+m,p}(U) \subset W^{j,q}(U), (1 \leq q \leq p^* = \frac{np}{n-mp}) \\ &W^{j+m,p}(U) \subset\subset W^{j,q}(U), (1 \leq q < p^* = \frac{np}{n-mp}) \end{aligned} \tag{3.3.22}$$

Theorem 3.3.27 shows is a basic property of Sobolev spaces. The Banach space is reflexive if the map from it to its second dual space is surjective.

Theorem 3.3.27 ([144, 35]). *Sobolev spaces are reflexive Banach spaces.*

Theorem 3.3.28 concerns the continuity of a functional, where the integrand is bounded by polynomials. Theorem 3.3.30 concerns the existence of a minimizer from coercivity and lower semi-continuity.

Theorem 3.3.28 ([16]). *If f is continuous and satisfies a growth condition of the form $|f(\vec{x})| \leq C(1 + |\vec{x}|^p)$, where C is a constant. Then $\mathcal{F} = \int_U f d\vec{x}$ is continuous on $L^p(U)$.*

Theorem 3.3.29 ([31, 16]). *Assume that the function $F : Y \to \bar{\mathbb{R}}$ is coercive and lower semi-continuous. Then F has a minimizer in Y.*

Theorem 3.3.30 concerns the convergence of the minimizers of an equi-coercive sequence of functions. Theorem 3.3.31 shows a sum property of Γ-limit. Both of them are used to prove the convergence of finite element approximation and finite element approximation with numerical quadratures.

Theorem 3.3.30 ([31, 16]). *Suppose that (F_h) is a sequence of functionals, which are equi-coercive in Y. And if (F_h) Γ-converges to a functional F in Y, then F is coercive and*

$$\min_{u\in Y} F(u) = \lim_{h\to 0} \inf_{u\in Y} F_h(u). \tag{3.3.23}$$

Theorem 3.3.31 ([31, 16]). *Suppose that (F_h) and (G_h) are sequences of functionals. If (F_h) Γ-converges to F and (G_h) is continuously convergent to G, then $(F_h + G_h)$ Γ-converges to $F + G$.*

3.3.3 Function Space Notation

Let $\Omega \subset \mathbb{R}^3$ be a finite bounded Lipchitz domain and the electronic system be restricted in this domain. In our analysis, we use standard notations for Sobolev spaces and their norms defined in (3.3.3) and (3.3.4) [1, 35]. We denote the space of m-times weakly differentiable functions in $L^2(\Omega)$ and $L^p(\Omega)$ by $H^k(\Omega)$ and $W^{k,p}(\Omega)$, and the subspace of $H^k(\Omega)$ and $W^{k,p}(\Omega)$ with zero boundary conditions by $H_0^k(\Omega)$ and $W_0^{k,p}(\Omega)$. We denote by X the solution space.

3.4 Existence of a Minimizer

The Section 3.2 of Suryanarayana et al. paper [146] presented the existence of a minimizer for the local spin density approximation model. We generalize the proofs of Suryanarayana et al. [146] to restricted Hartree-Fock, unrestricted Hartree-Fock, generalized gradient approximation models, meta generalized gradient approximation models and generalized density functional theory models [67].

The existence of a minimizer is the prerequisite for the ground state energy calculation. We use Theorem 3.3.29 to show the existence of a minimizer. The coercivity

is easier to be satisfied if we have more converging sequences, while the lower semi-continuity is easier to verify if we have less converging sequences. And the number of converging sequences is determined by the choice of topology on X. In the following proof, we choose weak topology of X for the proof.

Section 3.4.1 shows the proofs of Hartree-Fock methods. Section 3.4.2 shows the proof of density functional theory models.

3.4.1 Hartree-Fock

The proofs of restricted Hartree-Fock and unrestricted Hartree-Fock will be discussed in this section [67].

Restricted Hartree-Fock

The total energy functional $\mathcal{E}$ has been defined in (3.2.18). Suppose the solution space is [67]:

$$X_{RHF} = \left\{ \begin{array}{l} \Phi = (\phi_i)_{i=1}^{2n} \in (H_0^1(\Omega))^{2n} | \langle \phi_i, \phi_j \rangle_{(L^2(\Omega), L^2(\Omega))} = \delta_{ij}; \\ \varphi_{K_{ij}} \in H_0^1(\Omega), \varphi \in H_0^1(\Omega), i, j = 1, 2, \ldots, 2n. \end{array} \right\}. \tag{3.4.1}$$

The well-defined property shows the total energy functional $\mathcal{E}$ will not be infinite in a given function space, and Lemma 3.4.1 shows $\mathcal{E}$ is well-defined in X_{RHF}.

Lemma 3.4.1 ([67]). *$\mathcal{E}$ is well-defined in X_{RHF}.*

Proof. By the definitions (3.2.8), (3.2.17), (3.2.19) and Hölder Inequality, we can have

$$\begin{aligned} \mathcal{E} \leq \quad & C \Big\{ \|\Phi\|_{H^1(\Omega)}^2 + \|\varphi\|_{H^1(\Omega)}^2 + \|\rho\|_{L^2(\Omega)} \|\varphi\|_{L^2(\Omega)} + \|\varphi\|_{L^1(\Omega)} \\ & + \textstyle\sum_{i,j=1}^{2n} \left(\|\varphi_{K_{ij}}\|_{H^1(\Omega)}^2 + \|\varphi_{K_{ij}}\|_{L^2(\Omega)} \|\phi_i\|_{L^4(\Omega)} \|\phi_j\|_{L^4(\Omega)} \right) \Big\}, \end{aligned} \tag{3.4.2}$$

where C is a positive constant and ρ is defined in (3.2.6). It is also straightforward to show:

$$\|\rho\|_{L^2(\Omega)} = \|\Phi\|_{L^4(\Omega)}^2 \tag{3.4.3}$$

By Sobolev Embedding Theorem, we have $H_0^1(\Omega)$ compactly embedded in $L^1(\Omega), L^2(\Omega)$ and $L^4(\Omega)$. Then (3.4.1) and Theorem 3.3.20 show $\mathcal{E}$ is well-defined in X_{RHF}. Lemma 3.4.1 □

X_{RHF} is a subspace of $(H_0^1(\Omega))^{2n}$. X_{RHF} is closed if every sequence in X_{RHF} that converge in $(H_0^1(\Omega))^{2n}$, will converge to a function in X_{RHF}. Lemma 3.4.2 proves the X_{RHF} is closed in the weak topology of $(H_0^1(\Omega))^{2n}$, which will be used to show the lower semi-continuity and coercivity of $\mathcal{E}$ in the weak topology of X_{RHF}.

Lemma 3.4.2 ([146, 67]). *X_{RHF} is closed in the weak topology of $(H_0^1(\Omega))^{2n}$.*

Proof. The proof is exactly same as Lemma 1 of Suryanarayana et al. paper [146]. Consider an arbitrary sequence $\{\Phi_l\} \subset X_{RHF}$ with $\Phi_l \rightharpoonup \Phi$ in $H_0^1(\Omega)$. By Sobolev Embedding Theorem and Theorem 3.3.20, $H_0^1(\Omega)$ has a compact embedding into $L^2(\Omega)$, which means $\{\Phi_l\}$ has a subsequence $\Phi_l' \rightharpoonup \Phi'$ from which it follows that $\delta_{ij} = \langle \phi_{i,l}', \phi_{j,l}' \rangle_{(L^2(\Omega),L^2(\Omega))} = \langle \phi_i, \phi_j \rangle_{(L^2(\Omega),L^2(\Omega))}$ for $i, j = 1, 2, \ldots, 2n$ as $l \to +\infty$. Therefore $\Phi \in X_{RHF}$, which implied that X_{RHF} is closed in the weak topology of $(H_0^1(\Omega))^{2n}$. □

Lemma 3.4.3 shows the lower semi-continuity of the kinetic energy functional $T(\Phi)$. Since $T(\Phi)$ is a part of $\mathcal{E}\Phi)$, Lemma 3.4.3 will be used to show the lower semi-continuity of $\mathcal{E}\Phi)$.

Lemma 3.4.3 ([67]). *$T(\Phi)$ is lower semi-continuous if $\Phi \in X_{RHF}$.*

Proof. The proof is exactly same as Node (d) in Section 3.2 of The paper of Suryanarayana et al. [146]. As $\sum_{i=1}^{2n} \int \phi_i^*(\vec{x}) \nabla^2 \phi_i(\vec{x}) d\vec{x}$ is continuous in the strong topology of $H_0^1(\Omega)$ and convex, it follows that $\sum_{i=1}^{2n} \int \phi_i^*(\vec{x}) \nabla^2 \phi_i(\vec{x}) d\vec{x}$ is lower semi-continuous in the weak topology of $H_0^1(\Omega)$ by Theorem 3.3.17. □

Lemma 3.4.4 and Lemma 3.4.5 prove the continuous property of $\mathcal{J}(\Phi)$ and $K(\Phi)$, which are the preconditions of weakly lower semi-continuity.

Lemma 3.4.4 ([67]). *$\mathcal{J}(\Phi)$ is continuous if $\Phi \in (L^4(\Omega))^{2n}$.*

Proof. [45, Lemma 2]. If φ_u denote the minimizer of $\mathcal{J}(\Phi_u, \rho = \rho_u)$ (3.2.19), then

$$\frac{1}{4\pi} \int_\Omega \nabla \varphi_u \nabla \chi d\vec{x} = \int_\Omega (\rho_u + b) \chi d\vec{x}, \forall \chi \in H_0^1(\Omega). \tag{3.4.4}$$

For every ρ_u, ρ_v, we have

$$\frac{1}{4\pi} \int_\Omega \nabla (\varphi_u - \varphi_v) \nabla \chi d\vec{x} = \int_\Omega (\rho_u - \rho_v) \chi d\vec{x}, \forall \chi \in H_0^1(\Omega). \tag{3.4.5}$$

By Hölder Inequality and Poincaré Inequality, it is immediate that

$$\|\varphi_u - \varphi_v\|_{H_0^1(\Omega)} \leq C\|\rho_u - \rho_v\|_{L^2(\Omega)}, \tag{3.4.6}$$

where C is a positive constant. By (3.4.3), it shows ρ is continuous in $L^2(\Omega)$ if $\Phi \in (L^4(\Omega))^{2n}$. Therefore,

$$\begin{aligned} |\mathcal{J}(\Phi_u) - \mathcal{J}(\Phi_v)| \leq \quad & \|\nabla(\varphi_u - \varphi_v)\|_{L^2(\Omega)}\|\nabla(\varphi_u + \varphi_v)\|_{L^2(\Omega)} \\ & + \|\rho_u - \rho_v\|_{L^2(\Omega)}\|\varphi_u\|_{L^2(\Omega)} + \|\rho_v\|_{L^2(\Omega)}\|\varphi_u - \varphi_v\|_{L^2(\Omega)} \\ & + \|\varphi_u - \varphi_v\|_{L^2(\Omega)}\|b\|_{L^2(\Omega)} \end{aligned} \tag{3.4.7}$$

By (3.2.19) and (3.4.6), $\mathcal{J}(\Phi)$ is continuous if $\Phi \in (L^4(\Omega))^{2n}$. □

Lemma 3.4.5 ([67]). *$K(\Phi)$ is continuous if $\Phi \in (L^4(\Omega))^{2n}$.*

Proof. From (3.2.17), we have

$$K(\Phi) = -\sum_{i,j=1}^{2n} K_{ij}(\Phi) = \frac{1}{2}\sum_{i,j=1}^{2n} \min_{\varphi_{K_{ij}}} \left\{ \int [\frac{1}{8\pi}(\nabla\varphi_{K_{ij}}(\vec{x}))^2 - \phi_i^*(\vec{x})\phi_j(\vec{x})\varphi_{K_{ij}}(\vec{x})]d\vec{x} \right\}. \tag{3.4.8}$$

If φ_{K_{ij},Φ_u} denote the minimizer of $-K_{ij}(\Phi_u)$, then

$$\frac{1}{4\pi}\int_\Omega \nabla\varphi_{K_{ij}}\nabla\chi d\vec{x} = \int_\Omega \phi_i\phi_j^*\chi d\vec{x}, \forall\chi \in H_0^1(\Omega). \tag{3.4.9}$$

For every Φ_u, Φ_v, we have

$$\frac{1}{4\pi}\int_\Omega \nabla(\varphi_{K_{ij},\Phi_u} - \varphi_{K_{ij},\Phi_v})\nabla\chi d\vec{x} = \int_\Omega (\phi_{i,u}\phi_{j,u}^* - \phi_{i,v}\phi_{j,v}^*)\chi d\vec{x}, \forall\chi \in H_0^1(\Omega). \tag{3.4.10}$$

By Hölder Inequality and Poincaré Inequality, it is immediate that,

$$\frac{1}{4\pi}\|\varphi_{K_{ij},\Phi_u} - \varphi_{K_{ij},\Phi_v}\|_{H_0^1(\Omega)} \leq C\|\phi_{i,u}\phi_{j,u}^* - \phi_{i,v}\phi_{j,v}^*\|_{L^2(\Omega)}, \tag{3.4.11}$$

where C is a positive constant. Thus,

$$\begin{aligned} & |K(\Phi_u) - K(\Phi_v)| \\ \leq \quad & C\sum_{i,j=1}^{2n} \Big\{ \|\nabla(\varphi_{K_{ij},\Phi_u} - \varphi_{K_{ij},\Phi_v})\|_{L^2(\Omega)}\|\nabla(\varphi_{K_{ij},\Phi_u} + \varphi_{K_{ij},\Phi_v})\|_{L^2(\Omega)} \\ & + \|(\phi_{i,u}\phi_{j,u}^* - \phi_{i,v}\phi_{j,v}^*)\|_{L^2(\Omega)}\|\varphi_{K_{ij},\Phi_u}\|_{L^2(\Omega)} \\ & + \|\phi_{i,v}\phi_{j,v}^*\|_{L^2(\Omega)}\|\varphi_{K_{ij},\Phi_u} - \varphi_{K_{ij},\Phi_v}\|_{L^2(\Omega)} \Big\}. \end{aligned} \tag{3.4.12}$$

Then by (3.4.1) and (3.4.11), $K(\Phi)$ is continuous if $\Phi \in (L^4(\Omega))^{2n}$.

□

Lemma 3.4.6 is used to show the weakly lower semi-continuity of $\mathcal{J}(\Phi)$ and $K(\Phi)$ from their continuous property shown in Lemma 3.4.4 and Lemma 3.4.5.

Lemma 3.4.6 ([67]). *Y and Z are two topological spaces. If a functional F is continuous in the strong topology of Z, and Y is compactly embedded in Z ($Y \subset\subset Z$), then F is lower semi-continuous in the weak topology of Y.*

Proof. Consider an arbitrary sequence $(\phi_n) \subset Y$ with $\phi_n \rightharpoonup \phi$ in Y. If $F(\phi_n) \leq a$ for all n and a is a real number, then we can find a subsequence $(\phi_{nj}) \subset Y$ with $\phi_{nj} \to \phi$ in Z by Theorem 3.3.20. Because F is continuous in the strong topology of Z, we have $F(\phi) = \lim_{nj\to\infty} F(\phi_{nj}) \leq a$, which means $\{\phi \in Y | F(\phi) \leq a\}$ is weakly closed in X, and this is another definition of weakly lower semi-continuity. Thus, F is lower semi-continuous in the weak topology of Y. □

Lemma 3.4.7 shows the weakly lower semi-continuity of $\mathcal{E}$, and it is a precondition of the existence of a minimizer.

Lemma 3.4.7 ([67]). *$\mathcal{E}$ is lower semi-continuous in the weak topology of X_{RHF}.*

Proof. From Lemma 3.4.4 and Lemma 3.4.5, $K(\Phi)$ and $\mathcal{J}(\Phi)$ are continuous in $(L^4(\Omega))^{2n}$. Because $H_0^1(\Omega) \subset\subset L^4(\Omega)$ and Lemma 3.4.6, $K(\Phi)$ and $\mathcal{J}(\Phi)$ are lower semi-continuous in the weak topology of X_{RHF}.$T(\Phi)$ is also lower semi-continuous in the weak topology of X_{RHF} by Lemma 3.4.3.Therefore, $\mathcal{E}$ is lower semi-continuous in the weak topology of X_{RHF}. □

Lemma 3.4.8 shows the weak coercivity of $\mathcal{E}$, and it is another precondition of the existence of a minimizer.

Lemma 3.4.8 ([67]). *$\mathcal{E}$ is coercive in the weak topology of X_{RHF}.*

Proof. Suppose $\eta(\vec{x}, \vec{x}') = \phi_i(\vec{x})\phi_j(\vec{x}') - \phi_j(\vec{x})\phi_i(\vec{x}')$, then

$$\begin{aligned}
& \int\int \frac{\eta^*(\vec{x},\vec{x}')\eta(\vec{x},\vec{x}')}{|\vec{x}-\vec{x}'|} d\vec{x}d\vec{x}' \\
= \quad & \int\int \frac{\phi_i^*(\vec{x})\phi_j^*(\vec{x}')\phi_i(\vec{x})\phi_j(\vec{x}')}{|\vec{x}-\vec{x}'|} d\vec{x}d\vec{x}' - \int\int \frac{\phi_i^*(\vec{x})\phi_j^*(\vec{x}')\phi_j(\vec{x})\phi_i(\vec{x}')}{|\vec{x}-\vec{x}'|} d\vec{x}d\vec{x}' \\
& - \int\int \frac{\phi_j^*(\vec{x})\phi_i^*(\vec{x}')\phi_i(\vec{x})\phi_j(\vec{x}')}{|\vec{x}-\vec{x}'|} d\vec{x}d\vec{x}' + \int\int \frac{\phi_j^*(\vec{x})\phi_i^*(\vec{x}')\phi_j(\vec{x})\phi_i(\vec{x}')}{|\vec{x}-\vec{x}'|} d\vec{x}d\vec{x}' \\
= \quad & 2(J_{ij} - K_{ij}) \geq 0. \qquad (3.4.13)
\end{aligned}$$

From (3.2.10), (3.2.11) and (3.4.13), we have

$$J(\Phi) + K(\Phi) \geq 0. \tag{3.4.14}$$

We point out that $V_{ext}(\vec{x})$ is poinwise bounded and it holdes $|V_{ext}(\vec{x})| \leq C$ almost everywhere in Ω and some constant C. Therefore, $|V(\Phi)| \leq C \int_\Omega \rho dx = 2nC$. Similarly, you can have $|B| \leq C$. Also $T(\Phi) = \frac{1}{2}\|\nabla\Phi\|^2_{L^2(\Omega)}$. Thus,

$$\mathcal{E} \geq C_1\|\Phi\|^2_{H^1_0(\Omega)} + C_2, \tag{3.4.15}$$

where C_1 and C_2 are positive constants. This ensures $\mathcal{E} \to +\infty$ as $\|\Phi\|_{H^1_0(\Omega)} \to +\infty$. This is the coercivity of $\mathcal{E}$ in the weak topology of X_{RHF}. □

Theorem 3.4.9 shows the existence of a minimizer for the restricted Hartree-Fock model.

Theorem 3.4.9 ([67]). *$\mathcal{E}$ has a minimizer in X_{RHF}.*

Proof. This is an immediate result of Lemma 3.4.7, Lemma 3.4.8. and Theorem 3.3.29. □

Unrestricted Hartree-Fock

The total energy functional $\mathcal{E}$ has been defined in (3.2.21). Suppose the solution space is [67]:

$$X_{UHF} = \left\{ \begin{array}{l} \Phi = (\phi_i)_{i=1}^{n_\alpha+n_\beta} \in (H^1_0(\Omega))^{n_\alpha+n_\beta} | \langle \phi_i, \phi_j \rangle_{(L^2(\Omega),L^2(\Omega))} = \delta_{ij}, \\ \varphi_{K_{ij}} \in H^1_0(\Omega), \varphi \in H^1_0(\Omega), i,j = 1,2,\ldots,n_\alpha + n_\beta \end{array} \right\}. \tag{3.4.16}$$

Theorem 3.4.10 shows the existence of a minimizer for the unrestricted Hartree-Fock model.

Theorem 3.4.10 ([67]). *$\mathcal{E}$ has a minimizer in X_{UHF}.*

Proof. The only difference between restricted Hartree-Fock and unrestricted Hartree-Fock is the number of spin-up electrons and spin-down electrons. The proof is exactly the same as Section restricted Hartree-Fock. □

3.4.2 Density Functional Theory

We discuss the existence of a minimizer of density functional theory models in this section, including local spin density approximation, generalized gradient approximation, meta generalized gradient approximation, and generalized density functional theory models [67].

Local Spin Density Approximation

The existence of a minimizer of local spin density approximation has been carefully studied in the Section 3.2 of Suryanarayana et al. paper [146]. More to the local spin density approximation, we will discuss other density functional theory models below [67].

Generalized Gradient Approximation

Suppose the exchange and correlation energy $K(\rho, \bigtriangledown\rho)$ can be written as:

$$K(\rho, \bigtriangledown\rho) = \int_{\Omega} \Bbbk(\rho, \bigtriangledown\rho) d\vec{x}. \tag{3.4.17}$$

Suppose the solution space is [67]:

$$X_{GGA} = \left\{ \begin{array}{l} \Phi = (\phi_i)_{i=1}^{n_\alpha+n_\beta} \in (H_0^1(\Omega))^{n_\alpha+n_\beta} | \langle \phi_i, \phi_j \rangle_{(L^2(\Omega), L^2(\Omega))} = \delta_{ij}, \\ \varphi \in H_0^1(\Omega), i, j = 1, 2, \ldots, n_\alpha + n_\beta \end{array} \right\}. \tag{3.4.18}$$

We make the following hypothesis on $\Bbbk(\rho, \bigtriangledown\rho)$ [67]:

(A1) The density $\Bbbk(\rho, \bigtriangledown\rho)$ is convex and continuous in $\mathbb{R}^+$.

(A2) The growth condition

$$\Bbbk(\rho, \bigtriangledown\rho) = \sum_{\alpha,\beta} \Bbbk_{\alpha,\beta}, |\Bbbk_{\alpha,\beta}| \leq c_{1(\alpha,\beta)} |\rho|^\alpha |\bigtriangledown\rho|^\beta + c_{2(\alpha,\beta)}, \tag{3.4.19}$$

holds for positive constants $c_{1(\alpha,\beta)}, c_{2(\alpha,\beta)}$ and exponents satisfy:

$$\left\{ \begin{array}{l} 1 < k_{1(\alpha,\beta)}, k_{2(\alpha,\beta)}, k_{3(\alpha,\beta)} < \infty, 1/k_{1(\alpha,\beta)} + 1/k_{2(\alpha,\beta)} + 1/k_{3(\alpha,\beta)} = 1; \\ 1 \leq 2k_{1(\alpha,\beta)}\alpha, k_{2(\alpha,\beta)}\beta < 3; \\ 1 \leq k_{3(\alpha,\beta)}\beta < 2. \end{array} \right. \tag{3.4.20}$$

The number of (α, β) pairs is finite.

Lemma 3.4.11 is an inequality, and it will be used to analyze the complex exchange and correlation functional.

Lemma 3.4.11 ([67]). *If $a_i \geq 0, i = 1, 2, \ldots, N$ and $q > 0$, there exists constant C such that*

$$\left(\sum_{i=1}^{N} a_i\right)^q \leq C \left(\sum_{i=1}^{N} a_i^q\right)$$

.

Proof. Suppose $a_{i0} = \max(a_i), i = 1, \ldots, N$, then

$$\left(\sum_{i=1}^{N} a_i\right)^q \leq (N a_{i0})^q \leq N^q a_{i0}^q \leq N^q \sum_{i=1}^{N} a_i^q. \tag{3.4.21}$$

Therefore, we can choose $C = N^q$. □

Lemma 3.4.12 shows the exchange and correlation functional $K(\rho, \nabla\rho)$ is well-defined. $K(\rho, \nabla\rho)$ is a part of $\mathcal{E}$, and Lemma 3.4.12 will be used to show $\mathcal{E}$ is well-defined.

Lemma 3.4.12 ([67]). *Let (A1), (A2) hold, $K(\rho, \nabla\rho)$ is well-defined in X_{GGA}.*

Proof. By Lemma 3.4.11, Theorem 3.3.18 and Young's Inequality, we have:

$$\begin{aligned} \mathbb{k}_{\alpha,\beta}(\rho, \nabla\rho) &\leq c_{1(\alpha,\beta)} |\rho|^\alpha |\nabla\rho|^\beta + c_{2(\alpha,\beta)} \\ &\leq c_1 \sum_{i=1}^{n_\alpha + n_\beta} \left(|\phi_i|^{2k_{1(\alpha,\beta)}\alpha} + |\phi_i|^{k_{2(\alpha,\beta)}\beta} + |\nabla\phi_i|^{k_{3(\alpha,\beta)}\beta} \right) + c_2, \end{aligned} \tag{3.4.22}$$

which yields

$$K_{\alpha,\beta}(\rho, \nabla\rho) \leq c_1 \left[\|\Phi\|^{2k_{1(\alpha,\beta)}\alpha}_{L^{2k_{1(\alpha,\beta)}\alpha}(\Omega)} + \|\Phi\|^{k_{2(\alpha,\beta)}\beta}_{L^{k_{2(\alpha,\beta)}\beta}(\Omega)} + \|\nabla\Phi\|^{k_3\beta}_{L^{k_{3(\alpha,\beta)}\beta}(\Omega)} \right] + c_2, \tag{3.4.23}$$

where c_1 and c_2 are two positive constants, and they don't need to be same in the proof. Because $1 \leq k_{3(\alpha,\beta)}\beta < 2$, Ω is a finite bounded domain and Theorem 3.3.19, we have

$$K_{\alpha,\beta}(\rho, \nabla\rho) \leq c_1 \left[\|\Phi\|^{2k_{1(\alpha,\beta)}\alpha}_{L^{2k_{1(\alpha,\beta)}\alpha}(\Omega)} + \|\Phi\|^{k_{2(\alpha,\beta)}\beta}_{L^{k_{2(\alpha,\beta)}\beta}(\Omega)} + \|\nabla\Phi\|^2_{L^2(\Omega)} \right] + c_2, \tag{3.4.24}$$

By (A2) and Sobolev Embedding Theorem, $H_0^1(\Omega)$ is compactly embedded in $L^{2k_{1(\alpha,\beta)}\alpha}(\Omega)$, $L^{k_{2(\alpha,\beta)}\beta}(\Omega)$, which means $K_{\alpha,\beta}(\rho, \triangledown\rho)$ is well-defined. Because the number of (α, β) pairs is finite, $K(\rho, \triangledown\rho)$ is well-defined. □

Lemma 3.4.13 shows the total energy functional $\mathcal{E}$ of generalized gradient approximation models are well-defined in X_{GGA}.

Lemma 3.4.13 ([67]). *Let (A1), (A2) hold, $\mathcal{E}$ is well-defined in X_{GGA}.*

Proof. By (A2), Lemma 3.4.1 and Lemma 3.4.12, it is immediate that the total energy $\mathcal{E}$ is well-defined. □

Lemma 3.4.14 shows the X_{GGA} is weakly closed of $(H_0^1(\Omega))^{n_\alpha+n_\beta}$, which will be used to show the weakly lower semi-continuity and weak coercivity of $\mathcal{E}$.

Lemma 3.4.14 ([67, 146]). *Let (A1), (A2) hold, X_{GGA} is weakly closed of $(H_0^1(\Omega))^{n_\alpha+n_\beta}$*

Proof. The proof is exact same as Lemma 3.4.2, which follows Lemma 1 of Suryanarayana et al. paper [146]. □

Lemma 3.4.15 shows the weakly lower semi-continuity of $K(\rho, \triangledown\rho)$, which is the major different part between generalized gradient approximation model and Hartree-Fock models.

Lemma 3.4.15 ([67]). *Let (A1), (A2) hold, $K(\rho, \triangledown\rho)$ is lower semi-continuous in the weak topology of X_{GGA}.*

Proof. From (3.4.22), we have

$$\mathbb{k}_{\alpha,\beta}(\rho, \triangledown\rho) \leq c_1 \sum_{i=1}^{n_\alpha+n_\beta} \left(|\phi_i|^{2k_{1(\alpha,\beta)}\alpha} + |\phi_i|^{k_{2(\alpha,\beta)}\beta} + |\triangledown\phi_i|^{k_{3(\alpha,\beta)}\beta} \right) + c_2,$$

By Theorem 3.3.28, The three terms of $K_{\alpha,\beta}(\rho, \triangledown\rho)$ is strongly continuous in $L^{2k_1\alpha}(\Omega)$, $L^{k_2\beta}(\Omega)$ and $W_0^{1,k_3\beta}(\Omega)$ correspondingly. By Theorem 3.3.19 and Sobolev Embedding Theorem, $H_0^1(\Omega)$ is a subset of $L^{2k_1\alpha}(\Omega)$, $L^{k_2\beta}(\Omega)$ and $W_0^{1,k_3\beta}(\Omega)$. Therefore, $K(\rho, \triangledown\rho)$ is continuous and lower semi-continuous in the strong topology of $(H_0^1(\Omega))^{n_\alpha+n_\beta}$. By (A1), $K(\rho, \triangledown\rho)$ is convex. Thus, $K(\rho, \triangledown\rho)$ is weakly lower semi-continuous in

$(H_0^1(\Omega))^{n_\alpha+n_\beta}$, and as $X_{GGA} \subset (H_0^1(\Omega))^{n_\alpha+n_\beta}$ and weakly closed, the claimed lower semi-continuity in the weak topology of X_{GGA} follows.

□

Lemma 3.4.16 shows $\mathcal{E}$ is lower semi-continuous in the weak topology of X_{GGA}, and it is a precondition to show the existence of a minimizer.

Lemma 3.4.16 ([67])**.** *Let (A1), (A2) hold, $\mathcal{E}$ is lower semi-continuous in the weak topology of X_{GGA}.*

Proof. Lemma 3.4.7 shows the weakly lower semi-continuity of $T(\rho)$ and $\mathcal{J}(\rho)$. Lemma 3.4.15 shows the weakly lower semi-continuity of $K(\rho, \nabla\rho)$. Thus, $\mathcal{E}$ is lower semi-continuous in the weak topology of X_{GGA}. □

Lemma 3.4.17 shows an inequality that $(T(\rho) + \mathcal{J}(\rho))$ satisfied, and it will be used to show the weak coercivity of $T(\rho) + \mathcal{J}(\rho)$.

Lemma 3.4.17 ([67, 146])**.**

$$T(\rho) + \mathcal{J}(\rho) \geq c_0\|\nabla\Phi\|^2_{L^2(\Omega)} + c_1\|\Phi\|^3_{L^3(\Omega)} - c_2\|\Phi\|^2_{L^2(\Omega)} - c_3, \tag{3.4.25}$$

where c_0, c_1, c_2 and c_3 are positive constants.

Proof. The proof is exactly same as Lemma 3 of Suryanarayana et al. paper [146]. $\mathcal{J}(\rho)$ can be rewritten as

$$\mathcal{J}(\rho) = J(\rho) + V(\rho) + B, \tag{3.4.26}$$

where J_ρ is the electrostatic interaction energy among electrons, V_ρ is the electronic energy between electrons and the external field and B is the repulsive energy among the nuclei.

$$J(\rho) = -\min_{\varphi_\rho \in H_0^1(\Omega)} \left\{ \frac{1}{8\pi} \int |\nabla\varphi_\rho(\vec{x})|^2 d\vec{x} - \int \rho(\vec{x})\varphi_\rho(\vec{x}) d\vec{x} \right\}. \tag{3.4.27}$$

$J(\rho)$ is super-linear, we can have

$$\begin{aligned} J(\rho) &= J\Big(\sum_{i=1}^{n_\alpha+n_\beta} |\phi_i|^2\Big) \geq \sum_{i=1}^{n_\alpha+n_\beta} J(|\phi_i|^2) \\ &\geq \sum_{i=1}^{n_\alpha+n_\beta} \left[\max_{\varphi_\rho \in H_0^1(\Omega)} \left\{ \int |\phi_i(\vec{x})|^2 \varphi_{i,\rho}(\vec{x}) d\vec{x} - \frac{1}{8\pi} \int |\nabla\varphi_{i,\rho}(\vec{x})|^2 d\vec{x} \right\} \right]. \end{aligned} \tag{3.4.28}$$

We use $\varphi_{i,\rho} = c_0|\phi_i|$ as test functions and constant c_0 can be determined later. Therefore,

$$J(\rho) \geq c_0\|\Phi\|^3_{L^3(\Omega)} - \frac{c_0}{8\pi}\|\nabla\Phi\|^2_{L^2(\Omega)}. \tag{3.4.29}$$

Becuase external potential V_{ext} is pointwise bounded, we conclude

$$\mathcal{J}(\rho) \geq c_0\|\Phi\|^3_{L^3(\Omega)} - \frac{c_0}{8\pi}\|\nabla\Phi\|^2_{L^2(\Omega)} - c_1\|\Phi\|^2_{L^2(\Omega)} - c_2. \tag{3.4.30}$$

If we choose $c_0 < 4\pi$, then $\frac{1}{2} - \frac{c_0}{8\pi} = c_0' > 0$. We end up with

$$T(\rho) + \mathcal{J}(\rho) \geq c_0\|\nabla\Phi\|^2_{L^2(\Omega)} + c_1\|\Phi\|^3_{L^3(\Omega)} - c_2\|\Phi\|^2_{L^2(\Omega)} - c_3, \tag{3.4.31}$$

where c_0, c_1, c_2 and c_3 are positive constants. □

Lemma 3.4.18 shows $\mathcal{E}$ is coercive in the weak topology of X_{GGA}, and it is a precondition to show the existence of a minimizer.

Lemma 3.4.18 ([67]). *Let (A1), (A2) hold, $\mathcal{E}$ is coercive in the weak topology of X_{GGA}.*

Proof. By Lemma 3.9., (3.4.22) and (3.4.23), we have

$$\begin{aligned}
\mathcal{E} &= T(\Phi) + J(\rho) + K(\rho, \nabla\rho) \\
&\geq c_0\|\nabla\Phi\|^2_{L^2(\Omega)} + c_1\|\Phi\|^3_{L^3(\Omega)} - c_2\|\Phi\|^2_{L^2(\Omega)} \\
&\quad -c_3\sum_{\alpha,\beta}\left[\|\Phi\|^{2k_{1(\alpha,\beta)}\alpha}_{L^{2k_{1(\alpha,\beta)}\alpha}(\Omega)} + \|\Phi\|^{k_{2(\alpha,\beta)}\beta}_{L^{k_{2(\alpha,\beta)}\beta}(\Omega)} + \|\nabla\Phi\|^{k_3\beta}_{L^{k_{3(\alpha,\beta)}\beta}(\Omega)}\right] - c_4,
\end{aligned} \tag{3.4.32}$$

where c_0, c_1, c_2, c_3 and c_4 are positive constants. From (A2), $2k_{1(\alpha,\beta)}\alpha, k_{2(\alpha,\beta)}\beta \leq 3$ and $k_{3(\alpha,\beta)}\beta \leq 2$, we have $\mathcal{E} \to +\infty$ as $\|\Phi\|_{H^1_0(\Omega)} \to +\infty$, thus $\mathcal{E}$ is coercive in the weak topology of X_{GGA}.

□

Theorem 3.4.19 shows total energy functional $\mathcal{E}$ of generalized gradient approximation models have minimizers in X_{GGA}.

Theorem 3.4.19 ([67]). *Let (A1), (A2) hold, $\mathcal{E}$ possesses a minimizer in X_{GGA}.*

Proof. This is an immediate result of Lemma 3.4.16, Lemma 3.4.18. and Theorem 3.3.29. □

Meta Generalized Gradient Approximation

The exchange-correlation energy $K(\rho, \nabla\rho, \nabla^2\rho, \tau)$ of meta generalized gradient approximation models could be written as:

$$K(\rho, \nabla\rho, \nabla^2\rho, \tau) = \int_\Omega \Bbbk(\rho, \nabla\rho, \nabla^2\rho, \tau)d\vec{x}, \tag{3.4.33}$$

where $\tau = \frac{1}{2}\sum_i |\nabla\phi_i|^2$.

Suppose the solution space is [67]:

$$X_{Meta-GGA} = \left\{ \begin{array}{l} \Phi = (\phi_i)_{i=1}^N = (\phi_{1i})_{i=1}^{n_\alpha+n_\beta} + i(\phi_{2i})_{i=1}^{n_\alpha+n_\beta}; \\ Z = (\zeta_i)_{i=1}^{2n_\alpha+2n_\beta} = ((\phi_{1i})_{i=1}^{n_\alpha+n_\beta}, (\phi_{2i})_{i=1}^{n_\alpha+n_\beta}); \\ Z \in (W_0^{m,p}(\Omega))^{2n_\alpha+n_\beta}; \\ \rho = \sum_{i=1}^{n_\alpha+n_\beta} |\phi_i|^2 = \sum_{i=1}^{2n_\alpha+2n_\beta} \zeta_i^2; \\ \langle\phi_i, \phi_j\rangle_{(L^2(\Omega),L^2(\Omega))} = \delta_{ij}, \phi \in H_0^1(\Omega), i, j = 1, 2, \ldots, n_\alpha + n_\beta. \end{array} \right\}. \tag{3.4.34}$$

We make the following hypothesis on $K(\rho, \nabla\rho, \nabla^2\rho, \tau)$ [67]:

(B1) The density $K(\rho, \nabla\rho, \nabla^2\rho, \tau)$ is continuous in $\mathbb{R}^+$.

(B2) The growth condition

$$\Bbbk(\rho, \nabla\rho, \nabla^2\rho, \tau) = \sum_{\alpha,\beta} \Bbbk_{\alpha,\beta}, |\Bbbk_{\alpha,\beta}| \leq c_{1(\alpha,\beta)}|\rho|^{\alpha_0}|\nabla\rho|^{\alpha_1}|\nabla^2\rho|^{\alpha_2}|\tau|^\beta + c_{2(\alpha,\beta)}, \tag{3.4.35}$$

holds for positive $c_{1(\alpha,\beta)}, c_{2(\alpha,\beta)}$ and exponents satisfy

$$\left\{ \begin{array}{l} 1 < k_{1(\alpha,\beta)}, k_{2(\alpha,\beta)}, \ldots, k_{7(\alpha,\beta)} < \infty, \sum_{i=1}^7 1/k_{i(\alpha,\beta)} = 1; \\ W_0^{m,p}(\Omega) \subset\subset L^{2k_1\alpha_0}(\Omega), W_0^{m,p}(\Omega) \subset\subset L^{k_2\alpha_1}(\Omega), W_0^{m,p}(\Omega) \subset\subset L^{k_4k_6\alpha_2}(\Omega); \\ W_0^{m,p}(\Omega) \subset\subset L^{2k_4\alpha_2}(\Omega), W_0^{m,p}(\Omega) \subset\subset L^4(\Omega), W_0^{m,p}(\Omega) \subset\subset W^{1,k_3\alpha_1}(\Omega); \\ W_0^{m,p}(\Omega) \subset\subset W^{1,2k_5\beta}(\Omega), W_0^{m,p}(\Omega) \subset\subset W^{2,k_4k_7\alpha_2}(\Omega), W_0^{m,p}(\Omega) \subset H_0^1(\Omega). \end{array} \right. \tag{3.4.36}$$

The number of (α, β) pairs is finite.

(B3) $\mathcal{E}$ is coercive in the weak topology of $X_{Meta-GGA}$.

Lemma 3.4.20 shows the total energy functional $\mathcal{E}$ of meta generalized gradient approximation models are well-defined in $X_{Meta-GGA}$.

Lemma 3.4.20 ([67]). *Let (B1), (B2) hold, $\mathcal{E}$ is well-defined in $X_{Meta-GGA}$.*

Proof. By Lemma 3.4.11, Theorem 3.3.18 and Young's Inequality, we have:

$$\begin{aligned}\mathbb{k}_{\alpha,\beta} \leq \quad & C \sum_{i=1}^{2n_\alpha+2n_\beta} (|\zeta_i|^{2k_{1(\alpha,\beta)}\alpha_0} + |\zeta_i|^{k_{2(\alpha,\beta)}\alpha 1} + |\zeta_i|^{k_{4(\alpha,\beta)}k_6\alpha_2} + |\zeta_i|^{2k_{4(\alpha,\beta)}\alpha_2} \\ & + |\nabla\zeta_i|^{k_{3(\alpha,\beta)}\alpha_1} + |\nabla\zeta_i|^{2k_{5(\alpha,\beta)}\beta} + |\nabla^2\zeta_i|^{k_{4(\alpha,\beta)}k_{7(\alpha,\beta)}\alpha_2}). \end{aligned} \tag{3.4.37}$$

From (B2), we can get $K(\rho, \nabla\rho, \nabla^2\rho, \tau)$ well-defined. Also $W_0^{m,p}(\Omega) \subset H_0^1(\Omega)$, which means other terms in $\mathcal{E}$ is well-defined. Thus, $\mathcal{E}$ is well-defined in $X_{Meta-GGA}$. □

Lemma 3.4.21 proves $X_{Meta-GGA}$ is weakly closed of $(W_0^{m,p}(\Omega))^{2n_\alpha+2n_\beta}$, and it will be used to show the weakly lower semi-continuity of $\mathcal{E}$.

Lemma 3.4.21 ([67]). *Let (B1), (B2) hold, $X_{Meta-GGA}$ is weakly closed of $(W_0^{m,p}(\Omega))^{2n_\alpha+2n_\beta}$.*

Proof. Because $L^2(\Omega) \subset\subset W_0^{m,p}(\Omega)$, the proof is exact same as Lemma 3.4.2, which follows Lemma 1 of The paper of Suryanarayana et al. [146]. □

Lemma 3.4.22 proves $\mathcal{E}$ is lower semi-continuous in the weak topology of $X_{Meta-GGA}$, and it is a precondition of existence of a minimizer.

Lemma 3.4.22 ([67]). *Let (B1), (B2) hold, $\mathcal{E}$ is lower semi-continuous in the weak topology of $X_{Meta-GGA}$.*

Proof. As a result of (3.4.37), $K(\rho, \nabla\rho, \nabla^2\rho, \tau)$ is lower semi-continuous in the weak topology of $X_{Meta-GGA}$ by Theorem 3.3.28 and Lemma 3.4.6. Because $W_0^{m,p}(\Omega) \subset\subset L^4(\Omega)$, $W_0^{m,p}(\Omega) \subset H_0^1(\Omega)$, Lemma 3.4.3 and Lemma 3.4.6, $T(\rho)$ and $\mathcal{J}(\rho)$ are lower semi-continuous in the weak topology of $X_{Meta-GGA}$. Therefore, $\mathcal{E}$ is lower semi-continuous in the weak topology of $X_{Meta-GGA}$. □

Theorem 3.4.23 shows the total energy functional $\mathcal{E}$ of meta generalized gradient approximation models have minimizers.

Theorem 3.4.23 ([67]). *Let (B1)- (B3) hold, $\mathcal{E}$ possesses a minimizer in $X_{Meta-GGA}$.*

Proof. Therefore, it is an immediate result of Lemma 3.4.2, (B3). and Theorem 3.3.29. □

Generalized Density Functional Theory Models

The exchange-correlation energy $K(\rho, \nabla\rho, \nabla^2\rho, \ldots, \nabla^j\rho)$ could be written as:

$$K(\rho, \nabla\rho, \nabla^2\rho, \ldots, \nabla^j\rho) = \int_\Omega \Bbbk(\rho, \nabla\rho, \nabla^2\rho, \ldots, \nabla^j\rho)d\vec{x}. \tag{3.4.38}$$

Suppose the solution space is [67]:

$$X_{G-DFT} = \left\{ \begin{array}{l} \Phi = (\phi_i)_{i=1}^N = (\phi_{1i})_{i=1}^{n_\alpha+n_\beta} + i(\phi_{2i})_{i=1}^{n_\alpha+n_\beta}; \\ Z = (\zeta_i)_{i=1}^{2n_\alpha+2n_\beta} = ((\phi_{1i})_{i=1}^{n_\alpha+n_\beta}, (\phi_{2i})_{i=1}^{n_\alpha+n_\beta}) \in (W_0^{m,p}(\Omega))^{2n_\alpha+n_\beta}; \\ \rho = \sum_{i=1}^{n_\alpha+n_\beta} |\phi_i|^2 = \sum_{i=1}^{2n_\alpha+2n_\beta} \zeta_i^2; \\ \langle \phi_i, \phi_j \rangle_{(L^2(\Omega),L^2(\Omega))} = \delta_{ij}, \phi \in H_0^1(\Omega), i, j = 1, 2, \ldots, n_\alpha + n_\beta. \end{array} \right\}. \tag{3.4.39}$$

We make the following hypothesis on $k(\rho, \nabla\rho, \nabla^2\rho, \ldots, \nabla^j\rho)$ [67]:

(C1) The density $\Bbbk(\rho, \nabla\rho, \nabla^2\rho, \ldots, \nabla^j\rho)$ is continuous in $\mathbb{R}^+$.

(C2) The growth condition

$$\begin{aligned} &\Bbbk(\rho, \nabla\rho, \nabla^2\rho, \ldots, \nabla^j\rho) = \sum_\alpha \Bbbk_\alpha, \\ &|\Bbbk_\alpha| \leq c_{1(\alpha)}|\rho|^{\alpha_0}|\nabla\rho|^{\alpha_1}|\nabla^2\rho|^{\alpha_2} \cdots |\nabla^j\rho|^{\alpha_j} + c_{2(\alpha)} \end{aligned}$$

holds for positive $c_{1(\alpha)}, c_{2(\alpha)}$ and exponents satisfy

$$\left\{ \begin{array}{l} 1 < k_{0(\alpha)}, k_{1(\alpha)}, \ldots, k_{j(\alpha)} < \infty, \sum_{i=1}^j 1/k_{i(\alpha)} = 1; \\ 1 < k_{s,i-s(\alpha)} < \infty, s = 0, \ldots, i \text{ and } i = 0, \ldots, j; \\ 1/k_{s,i-s(\alpha)} + 1/k_{i-s,s(\alpha)} = 1, \text{if } s \neq i - s; \\ 1/k_{s,s(\alpha)} + 1/k'_{s,s(\alpha)} = 1, \text{if } s = i - s; \\ W_0^{m,p}(\Omega) \subset\subset \cap_{i=0}^j \cap_{s=0}^i W^{s,k_{s,i-s(\alpha)}\alpha_i k_i}(\Omega); \\ W_0^{m,p}(\Omega) \subset H_0^1(\Omega). \end{array} \right.$$

The number of $k(\alpha)$ terms is finite.

(C3) $\mathcal{E}$ is coercive in the weak topology of X_{G-DFT}.

Theorem 3.4.24 shows the total energy functional $\mathcal{E}$ of generalized density functional theory models have minimizers.

Theorem 3.4.24 ([67]). *Let (C1)-(C3) hold, $\mathcal{E}$ possesses a minimizer in X_{G-DFT}.*

Proof. The proof is very similar with the section meta generalized gradient approximation. By Lemma 3.4.11, Theorem 3.3.18 and Young's Inequality, we can get

$$\mathbb{k}_\alpha \leq \sum_{i=0}^{j}\sum_{s=0}^{i}\left[\sum_{t=1}^{2n_\alpha+2n_\beta}\left|\nabla^s\zeta_t\right|^{k_{s,i-s(\alpha)}\alpha_i k_{i(\alpha)}}\right]. \tag{3.4.40}$$

Combined with (C2) Theorem 3.3.28 and Lemma 3.4.6, we can show $\mathcal{E}$ is lower semicontinuous in the weak topology of X_{G-DFT}. Finally, $\mathcal{E}$ possesses a minimizer in X_{G-DFT} by (C3) and Theorem 3.3.29.

□

3.5 Convergence of Finite Element Approximation

The Section 3.3 of Suryanarayana et al. paper [146] presented the convergence of the finite element approximation for local spin density approximation model. We generalize the proofs of Suryanarayana et al. [146] to restricted Hartree-Fock, unrestricted Hartree-Fock, generalized gradient approximation model, meta generalized gradient approximation model, and generalized density functional theory models [67].

Finite element approximation is an important method to discretize the continuous partial differential equation into matrix. In this section, we will show the convergence of the ground state energy of a system computed with a finite element approximation for Hartree-Fock and density functional theory models.

We use Theorem 3.3.30 to prove the convergnce of finite element approximation. Weak topology is choosen to satisfy both the equi-coercivity and Γ-convergence property.

Let h be the size of triangulation mesh T_h of Ω. Let X_h be the corresponding sequence of subspaces of X consisting of functions whose restriction to every cell in T_h

is a polynomial. By P_k we denote the ring of polynomials of non-negative degree less than or equal to k for some fixed $k \geq 1$.

Section 3.5.1 shows the proofs of Hartree-Fock methods. Section 3.5.2 shows the proofs of density functional theory models.

3.5.1 Hartree-Fock

The convergence of finite element approximation of restricted Hartree-Fock and unrestricted Hartree-Fock models will be discussed in this section [67].

Restricted Hartree-Fock

Suppose finite elment approximate space X_h is [67]:

$$X_{h,RHF} = \left\{ \begin{array}{l} \Phi_h = (\phi_{h,i})_{i=1}^{2n} \in (H_0^1(\Omega))^{2n} | \langle \phi_{h,i}, \phi_{h,j} \rangle_{(L^2(\Omega),L^2(\Omega))} = \delta_{ij}; \\ \varphi_{h,K_{ij}} \in H_0^1(\Omega), \varphi_h \in H_0^1(\Omega), i,j = 1,2,\ldots,2n; \\ (\phi_{h,i})_{i=1}^{2n}, \varphi_{h,K_{ij}}, \varphi_h \in P_k \text{ for every cell in } T_h. \end{array} \right\}. \quad (3.5.1)$$

We define the discrete energy functional (3.2.18) with a finite element approximation as [146, 67]

$$\mathcal{E}_h(\Phi) = \left\{ \begin{array}{l} T(\Phi) + \mathcal{J}_h(\Phi) + K_h(\Phi), \text{ if } \Phi \in X_{h,RHF}, \\ +\infty, \text{ if } \Phi \in X \backslash X_{h,RHF}, \end{array} \right. \quad (3.5.2)$$

where

$$\mathcal{J}_h(\Phi) = -\min_{\varphi \in H_0^1(\Omega)} I_h(\Phi, \varphi), \quad (3.5.3)$$

$$K_h(\Phi) = \min_{\varphi_{K_{ij}} \in H_0^1(\Omega)} L_h(\Phi, \varphi_{K_{ij}}), \quad (3.5.4)$$

where

$$I_h(\Phi, \varphi) = \left\{ \begin{array}{l} I(\Phi, \varphi), \text{ if } \Phi, \varphi \in X_{h,RHF}, \\ +\infty, \text{ otherwise}, \end{array} \right. \quad (3.5.5)$$

$$L_h(\Phi, \varphi_{K_{ij}}) = \left\{ \begin{array}{l} L(\Phi, \varphi_{K_{ij}}), \text{ if } \Phi, \varphi_{K_{ij}} \in X_{h-RHF}, \\ +\infty, \text{ otherwise}, \end{array} \right. \quad (3.5.6)$$

$$I(\Phi, \varphi) = \frac{1}{8\pi} \int_\Omega |\nabla \varphi(\vec{x})|^2 d\vec{x} - \int_\Omega (\rho(\vec{x}) + b(\vec{x})) \varphi(\vec{x}) d\vec{x}, \quad (3.5.7)$$

$$L(\Phi, \varphi_{K_{ij}}) = \frac{1}{2}\sum_{i,j=1}^{2n}\left\{\int_\Omega [\frac{1}{8\pi}(\nabla\varphi_{K_{ij}}(\vec{x}))^2 - \phi_i^*(\vec{x})\phi_j(\vec{x})\varphi_{K_{ij}}(\vec{x})]d\vec{x}\right\}. \tag{3.5.8}$$

Lemma 3.5.1 and Lemma 3.5.2 shows $\lim_{h\to 0}\mathcal{J}_h(\Phi_h) = \mathcal{J}(\Phi)$ and $\lim_{h\to 0} K_h(\Phi_h) = K(\Phi)$, and they will be used to show the weakly Γ-convergence and weak equi-coercivity.

Lemma 3.5.1 ([45, 67]). *If $(\Phi_h) \in (X_{h,RHF})$ is a sequecnce such that $\Phi_h \to \Phi$ in $(L^4(\Omega))^{2n}$, then*

$$\lim_{h\to 0}\mathcal{J}_h(\Phi_h) = \mathcal{J}(\Phi). \tag{3.5.9}$$

Proof. The proof is exactly same as Theorem 9 of Gavinni's paper [45]. We will show Γ-convergence property of I_h in part (1) of the proof and equi-coercivity of I_h in part (2) of the proof.

(1) Consider a sequence $\varphi_h \rightharpoonup \varphi$ in X_{RHF}. If there is no subsequence $\varphi_{h_k} \in X_{h_k,RHF}$, then

$$+\infty = \lim\inf_{k\to\infty} I_{h_k}(\varphi_{h_k}) \geq I(\varphi). \tag{3.5.10}$$

Otherwise,

$$\begin{aligned}\lim_{k\to\infty}\int_\Omega(\rho_{h_k}(\vec{x}) + b(\vec{x}))\varphi_{h_k}(\vec{x})d\vec{x} &= \lim_{k\to\infty}\int_\Omega(\rho_{h_k}(\vec{x}) + b(\vec{x}))\varphi(\vec{x})d\vec{x} \\ &= \int_\Omega(\rho(\vec{x}) + b(\vec{x}))\varphi(\vec{x})d\vec{x}. \end{aligned}\tag{3.5.11}$$

Because $\Phi_h \to \Phi$ in $(L^4(\Omega))^{2n}$, we have $\rho_h \to \rho$ in $L^2(\Omega)$, the first equality in (3.5.11) is from the definition of weak convergence, and we can also get the second equality in (3.5.11) by Hölder inequality easily. Combined with the weakly lower semi-continuity of $\frac{1}{8\pi}\int_\Omega |\nabla\varphi(\vec{x})|^2 d\vec{x}$, we have:

$$\lim\inf_{k\to\infty} I_{h_k}(\varphi_{h_k}) \geq I(\varphi). \tag{3.5.12}$$

Also a recovery sequence (φ_h) can be constructed from interpolated function, it follows that $\lim_{h\to 0} I_h(\varphi_h) = I(\varphi)$. Thus, we have $I_h(\varphi_h) \to I(\varphi)$ (in the Γ-sense) in the weak topology of X_{RHF}.

(2) If $\varphi \notin X_{RHF}$, then $I_h(\varphi) = +\infty$.

Otherwise, by Hölder inequality we have:

$$I_h(\varphi) \geq c_0\|\varphi\|^2_{H_0^1(\Omega)} - c_1\|\rho_h + b\|_{L_2(\Omega)}\|\varphi\|_{L_2(\Omega)}, \tag{3.5.13}$$

where c_0 and c_1 are positive constants. Because $\rho_h \in L_2(\Omega)$ and b is pointwise bounded,

$$I_h(\varphi) \geq c_0\|\varphi\|^2_{H_0^1(\Omega)} - c_1\|\varphi\|_{L_2(\Omega)}. \tag{3.5.14}$$

The right side in (3.5.14) is coercive in the weak topology of X_{RHF}.
Based on (1), (2) and Theorem 3.3.30, we have (3.5.9). □

Lemma 3.5.2 ([67]). *If $(\Phi_h) \in (X_{h,RHF})$ is a sequecnce such that $\Phi_h \to \Phi$ in $(L^4(\Omega))^{2n}$, then*

$$\lim_{h\to 0} K_h(\Phi_h) = K(\Phi). \tag{3.5.15}$$

Proof. We will show Γ-convergence property of K_h in part (1) of the proof and equi-coercivity of K_h in part (2) of the proof.
(1) Consider a sequence $\varphi_{h,K_{ij}} \rightharpoonup \varphi_{K_{ij}}$ in X_{RHF}. If there is no subsequence $\varphi_{h_k,K_{ij}} \in X_{h_k,RHF}$, then

$$+\infty = \lim\inf_{k\to\infty} L_{h_k}(\varphi_{h_k,K_{ij}}) \geq L(\varphi). \tag{3.5.16}$$

Otherwise,

$$\begin{aligned}\lim_{k\to\infty}\int_\Omega \phi^*_{h_k,i}(\vec{x})\phi_{h_k,j}(\vec{x})\varphi_{h_k,K_{ij}}(\vec{x})d\vec{x} = &\quad \lim_{k\to\infty}\int_\Omega \phi^*_{h_k,i}(\vec{x})\phi_{h_k,j}(\vec{x})\varphi_{K_{ij}}(\vec{x})d\vec{x}\\ = &\quad \lim_{k\to\infty}\int_\Omega \phi^*_{i}(\vec{x})\phi_{j}(\vec{x})\varphi_{K_{ij}}(\vec{x})d\vec{x}.\end{aligned} \tag{3.5.17}$$

Because $\Phi_h \to \Phi$ in $(L^4(\Omega))^{2n}$, the first equality in (3.5.17) is from the definition of weak convergence, we can also get the second equality in (3.5.17) by Hölder inequality easily. Combined with the weakly lower semi-continuity of $\frac{1}{8\pi}\int_\Omega |\nabla\varphi_{h_k,K_{ij}}(\vec{x})|^2 d\vec{x}$, we have:

$$\lim\inf_{k\to\infty} L_{h_k}(\varphi_{h_k,K_{ij}}) \geq L(\varphi). \tag{3.5.18}$$

Also a recovery sequence $(\varphi_{h,K_{ij}})$ can be constructed from interpolated function, it follows that $\lim_{h\to 0} L_h(\varphi_{h,K_{ij}}) = L(\varphi_{K_{ij}})$. Thus, we have $L_h(\varphi_{h,K_{ij}}) \to L(\varphi_{K_{ij}})$ (in the Γ-sense) in the weak topology of X_{RHF}.
(2) If $\varphi_{K_{ij}} \notin X_{RHF}$, then $L_h(\varphi_{K_{ij}}) = +\infty$.
Otherwise, by Hölder inequality we have:

$$L_h(\varphi_{K_{ij}}) \geq c_0\sum_{i,j=1}^{2n}\|\varphi_{K_{ij}}\|^2_{H_0^1(\Omega)} - c_1\sum_{i,j=1}^{2n}\|\phi_{h,i}\|_{L_4(\Omega)}\|\phi_{h,j}\|_{L_4(\Omega)}\|\varphi_{K_{ij}}\|_{L_2(\Omega)}, \tag{3.5.19}$$

where c_0 and c_1 are positive constants. Because $\Phi_h \in L_4(\Omega)$, we have

$$L_h(\varphi_{K_{ij}}) \geq c_0 \sum_{i,j=1}^{2n} \|\varphi_{K_{ij}}\|^2_{H_0^1(\Omega)} - c_1 \sum_{i,j=1}^{2n} \|\varphi_{K_{ij}}\|_{L_2(\Omega)}. \tag{3.5.20}$$

The right side in (3.5.20) is coercive in the weak topology of X_{RHF}. L_h is equi-coercive in the weak topology of X_{RHF} .

Based on (1), (2) and Theorem 3.3.30, we have (3.5.15).

□

Lemma 3.5.3 proves $\mathcal{E}_h \to \mathcal{E}$ (in the Γ-sense) in the weak topology of X_{RHF}, and it is a precondition to show the convergence of finite element approximation.

Lemma 3.5.3 ([67]). *$\mathcal{E}_h \to \mathcal{E}$ (in the Γ-sense) in the weak topology of X_{RHF}.*

Proof. Γ-convergence needs to satisfy two conditions: (1). the $\liminf$ inequality, (2) construct a recovery sequence.

(1). Consider a sequence $\Phi_h \rightharpoonup \Phi$ in X_{RHF}.

(1a). There is no subsequence $\Phi_{h_k} \in X_{h_k,RHF}$, then

$$+\infty = \liminf_{k\to\infty} \mathcal{E}_{h_k}(\Phi_{h_k}) \geq \mathcal{E}(\Phi). \tag{3.5.21}$$

(1b). There is a subsequence $\Phi_{h_k} \in X_{h_k,RHF}$, then

$$\liminf_{k\to\infty} \mathcal{E}_{h_k,RHartree-Fock}(\Phi_{h_k}) \geq \liminf_{k\to\infty} \left\{ \frac{1}{2}\|\nabla\Phi_{h_k}\|^2_{L^2(\Omega)} + \mathcal{J}_{h_k}(\Phi_{h_k}) + K(\Phi_{h_k}) \right\}. \tag{3.5.22}$$

$\frac{1}{2}\|\nabla\Phi_{h_k}\|^2_{L^2(\Omega)}$ is lower semi-continuous in the weak topology of X_{RHF}. And because of $X_{RHF} \subset\subset L^4$, Lemma 3.5.1 and Lemma 3.5.2, we have $\lim_{k\to\infty} \mathcal{J}_{h_k}(\Phi_{h_k}) = \mathcal{J}(\Phi)$ and $\lim_{k\to\infty} K_{h_k}(\Phi_{h_k}) = K(\Phi)$. This shows that

$$\liminf_{k\to\infty} \mathcal{E}_{h_k}(\Phi_{h_k}) \geq \mathcal{E}(\Phi). \tag{3.5.23}$$

This is the $\liminf$ inequality.

(2) construct a recovery sequence.

We can always construct the interpolation functions of successive triangulations such that $\Phi_h \to \Phi$, it follows that $\lim_{h\to 0} \mathcal{E}_h(\Phi_h) = \mathcal{E}(\Phi)$ by Lemma 3.5.1, Lemma 3.5.2 and continuity of $T(\Phi)$.

By (1) and (2), we can have $\mathcal{E}_h \to \mathcal{E}$ (in the Γ-sense) in the weak topology of X_{RHF}. □

Lemma 3.5.4 proves $(\mathcal{E}_h)$ is equi-coercive in the weak topology of X_{RHF}, and it is a precondition to show the convergence of finite element approximation.

Lemma 3.5.4 ([67]). *$(\mathcal{E}_h)$ is equi-coercive in the weak topology of X_{RHF}.*

Proof. (1) If $\Phi \notin X_{RHF}$, then

$$\mathcal{E}_h(\Phi) = +\infty. \tag{3.5.24}$$

(2) If $\Phi \in X_{RHF}$, then by the coercive property of $\mathcal{E}$ in Lemma 3.4.8, we can show

$$\begin{aligned} \mathcal{E}_h(\Phi) &\geq \frac{1}{2}\|\nabla\Phi\|^2_{L^2(\Omega)} + \mathcal{J}_h(\Phi) + K_h(\Phi) \\ &\geq C_1\|\Phi\|^2_{H^1_0(\Omega)} + C_2. \end{aligned} \tag{3.5.25}$$

Because the right side of (3.5.25) is a lower semi-continuous coercive functional in the weak topology of X_{RHF}, it follows that $(\mathcal{E}_h)$ is equi-coercive in the weak topology of X_{RHF}. □

Theorem 3.5.5 shows the convergence of finite element approximation for the restricted Hartree-Fock model.

Theorem 3.5.5 ([67]).

$$\lim_{h\to 0} \inf_{X_{RHF}} \mathcal{E}_h = \min_{X_{RHF}} \mathcal{E}. \tag{3.5.26}$$

Proof. This is an immediate result of Lemma 3.5.3, Lemma 3.5.4 and Theorem 3.3.30. □

Unrestricted Hartree-Fock

Theorem 3.5.6 shows the convergence of finite element approximation for the unrestricted Hartree-Fock model.

Theorem 3.5.6 ([67]).

$$\lim_{h\to 0} \inf_{X_{UHF}} \mathcal{E}_h = \min_{X_{UHF}} \mathcal{E}. \tag{3.5.27}$$

Proof. The only difference between restricted Hartree-Fock and unrestricted Hartree-Fock is the number of spin-up electrons and spin-down electrons. The proof is exactly the same as Section restricted Hartree-Fock. □

3.5.2 Density Functional Theory

In this section, we discuss the convergence of finite element approximation of density functional theory models, including local spin density approximation, generalized gradient approximation, meta generalized gradient approximation and generalized density functional theory models.

Local Spin Density Approximation

The convergence of the finite element approximation of local spin density approximation has been carefully studied in the Section 3.3 of Suryanarayana et al. paper [146]. More to the local spin density approximation, we will discuss other density functional theory models below [67].

Generalized Gradient Approximation

Suppose finite element approximate space X_h is [67]:

$$X_{h,GGA} = \left\{ \begin{array}{l} \Phi = (\phi_i)_{i=1}^{n_\alpha+n_\beta} \in (H_0^1(\Omega))^{n_\alpha+n_\beta} | \langle \phi_i, \phi_j \rangle_{(L^2(\Omega),L^2(\Omega))} = \delta_{ij}; \\ \varphi \in H_0^1(\Omega), i,j = 1,2,\ldots,n_\alpha+n_\beta; \\ (\phi_{h,i})_{i=1}^{n_\alpha+n_\beta}, \varphi_h \in P_k \text{ for every cell in } T_h. \end{array} \right\}. \tag{3.5.28}$$

We define the discrete energy functional (3.2.26) with a finite element approximation as [67]

$$\mathcal{E}_h(\rho) = \left\{ \begin{array}{l} T(\rho) + \mathcal{J}_h(\rho) + K(\rho), \text{ if } \Phi \in X_{h,GGA}, \\ +\infty, \text{ if } \Phi \in X \backslash X_{h,GGA}, \end{array} \right. \tag{3.5.29}$$

where $\mathcal{J}_h(\rho)$ is defined in (3.5.3).

Theorem 3.5.7 shows the convergence of finite element approximation for generalized gradient approximation models.

Theorem 3.5.7 ([67])**.** *Let (A1), (A2) hold, then*

$$\lim_{h\to 0} \inf_{X_{GGA}} \mathcal{E}_h = \min_{X_{GGA}} \mathcal{E}. \tag{3.5.30}$$

Proof. We will show Γ-convergence property of $\mathcal{E}_h$ in part (1) of the proof and equi-coercivity of $\mathcal{E}_h$ in part (2) of the proof.

(1) Consider a sequence $\Phi_h \rightharpoonup \Phi$ in X_{GGA}. If there is no subsequence $\Phi_{h_k} \in X_{h_k,GGA}$, then

$$+\infty = \lim \inf_{k\to\infty} \mathcal{E}_{h_k}(\rho_{h_k}) \geq \mathcal{E}(\rho). \tag{3.5.31}$$

If there is a subsequence $\Phi_{h_k} \in X_{h_k,GGA}$, by (3.5.3) we have:

$$\lim \inf_{k\to\infty} \mathcal{E}_{h_k}(\rho_{h_k}) \geq \lim \inf_{k\to\infty} \left\{ \frac{1}{2} \| \nabla \Phi_{h_k} \|^2_{L^2(\Omega)} + \mathcal{J}_{h_k}(\rho_{h_k}) + K(\rho_{h_k}) \right\}. \tag{3.5.32}$$

Lemma 3.4.7 and Lemma 3.4.15 show the weakly lower semi-continuity of $\frac{1}{2}\| \nabla \Phi_{h_k} \|^2_{L^2(\Omega)}$ and $K(\rho_{h_k})$. And because $X_{GGA}(\Omega) \subset\subset L^4(\Omega)$, we have $\lim\inf_{k\to\infty} \mathcal{J}_{h_k}(\rho_{h_k}) = \mathcal{J}(\rho)$ by Lemma 3.5.1. This shows

$$\lim \inf_{k\to\infty} \mathcal{E}_{h_k}(\rho_{h_k}) \geq \mathcal{E}(\rho). \tag{3.5.33}$$

The interpolation functions of successive triangulations can also be constructed to be a recovery sequence. By continuity of each individual term in $\mathcal{E}_h(\rho)$, we have $\lim_{h\to 0} \mathcal{E}_h = \mathcal{E}$. Thus, we get the Γ-convergence property.

(2) If $\Phi \notin X_{h,GGA}$, then $\mathcal{E}_h = +\infty$.

Otherwise, by (3.4.32) we have:

$$\begin{aligned} \mathcal{E}_h(\rho) \quad \geq \quad & c_0 \| \nabla \Phi \|^2_{L^2(\Omega)} + c_1 \| \Phi \|^3_{L^3(\Omega)} - c_2 \| \Phi \|^2_{L^2(\Omega)} \\ & - c_3 \left[\| \Phi \|^{2k_1\alpha}_{L^{2k_1\alpha}(\Omega)} + \| \Phi \|^{k_2\beta}_{L^{k_2\beta}(\Omega)} + \| \nabla \Phi \|^{k_3\beta}_{L^{k_3\beta}(\Omega)} \right] - c_4. \end{aligned} \tag{3.5.34}$$

where c_0, c_1, c_2, c_3 and c_4 are positive constants. Because the right side of (3.5.34) is a coercive function in the weak topology of X_{GGA} independent of h, it follows that $\mathcal{E}_h$ is equi-coercive in the weak topology of X_{GGA}.

Based on (1), (2) and Theorem 3.3.30, we have (3.5.30). □

Meta Generalized Gradient Approximation

Suppose X_h is [67]:

$$X_{h,Meta-GGA} = \left\{ \begin{array}{l} \Phi_h = (\phi_{h,i})_{i=1}^{N} = (\phi_{h,1i})_{i=1}^{n_\alpha+n_\beta} + i(\phi_{h,2i})_{i=1}^{n_\alpha+n_\beta}; \\ Z_h = (\zeta_{h,i})_{i=1}^{2n_\alpha+2n_\beta} = ((\phi_{h,1i})_{i=1}^{n_\alpha+n_\beta}, (\phi_{h,2i})_{i=1}^{n_\alpha+n_\beta}); \\ Z_h \in (W_0^{m,p}(\Omega))^{2n_\alpha+n_\beta}; \\ \rho_h = \sum_{i=1}^{n_\alpha+n_\beta} |\phi_{h,i}|^2 = \sum_{i=1}^{2n_\alpha+2n_\beta} \zeta_{h,i}^2; \\ \langle \phi_{h,i}, \phi_{h,j} \rangle_{(L^2(\Omega),L^2(\Omega))} = \delta_{ij}, i, j = 1, \ldots, n_\alpha + n_\beta; \\ \varphi_h \in H_0^1(\Omega), \end{array} \right\}. \quad (3.5.35)$$

We define the discrete energy functional (3.2.26) with a finite element approximation as [67]

$$\mathcal{E}_h(\rho) = \left\{ \begin{array}{l} T(\rho) + \mathcal{J}_h(\rho) + K(\rho), \text{ if } \Phi \in X_{h,Meta-GGA}, \\ +\infty, \text{ if } \Phi \in X \backslash X_{h,Meta-GGA}, \end{array} \right. \quad (3.5.36)$$

where $\mathcal{J}_h(\rho)$ is defined in (3.5.3).

Theorem 3.5.8 shows the convergence of finite element approximation for meta generalized gradient approximation models.

Theorem 3.5.8 ([67]). *Let (B1)-(B3) hold, then*

$$\lim_{h \to 0} \inf_{X_{Meta-GGA}} \mathcal{E}_h = \min_{X_{Meta-GGA}} \mathcal{E}. \quad (3.5.37)$$

Proof. We will show Γ-convergence property of $\mathcal{E}_h$ in part (1) of the proof and equi-coercivity of $\mathcal{E}_h$ in part (2) of the proof.

1) Consider a sequence $\Phi_h \rightharpoonup \Phi$ in $X_{Meta-GGA}$. If there is no subsequence $\Phi_{h_k} \in X_{h_k,Meta-GGA}$, then

$$+\infty = \lim \inf_{k \to \infty} \mathcal{E}_{h_k}(\rho_{h_k}) \geq \mathcal{E}(\rho). \quad (3.5.38)$$

If there is a subsequence $\Phi_{h_k} \in X_{h_k,Meta-GGA}$, by (3.5.3) we have:

$$\lim \inf_{k \to \infty} \mathcal{E}_{h_k}(\rho_{h_k}) \geq \lim \inf_{k \to \infty} \left\{ \frac{1}{2} \| \nabla \Phi_{h_k} \|_{L^2(\Omega)}^2 + \mathcal{J}_{h_k}(\rho_{h_k}) + K(\rho_{h_k}) \right\}. \quad (3.5.39)$$

Lemma 3.4.22 shows the weakly lower semi-continuity of $\frac{1}{2} \| \nabla \Phi_{h_k} \|_{L^2(\Omega)}^2$ and $K(\rho_{h_k})$. And because $X_{Meta-GGA}(\Omega) \subset\subset L^4(\Omega)$, we have $\liminf_{k \to \infty} \mathcal{J}_{h_k}(\rho_{h_k}) = \mathcal{J}(\rho)$ by Lemma 3.5.1. This shows

$$\lim \inf_{k \to \infty} \mathcal{E}_{h_k}(\rho_{h_k}) \geq \mathcal{E}(\rho). \quad (3.5.40)$$

The interpolation functions of successive triangulations can also be constructed to be a recovery sequence. By continuity of each individual term in $\mathcal{E}_h(\rho)$, we have $\lim_{h\to 0}\mathcal{E}_h = \mathcal{E}$. Thus, we get the Γ-convergence property.
(2) If $\Phi \notin X_{h,GGA}$, then $\mathcal{E}_h = +\infty$.
Otherwise,

$$\mathcal{E}_h(\rho) \geq \mathcal{E}. \tag{3.5.41}$$

By (B3), $\mathcal{E}$ is a coercive function in the weak topology of $X_{Meta-GGA}$ independent of h, it follows that $\mathcal{E}_h$ is equi-coercive in the weak topology of $X_{Meta-GGA}$.
Based on (1), (2) and Theorem 3.3.30, we have (3.5.37). □

Generalized Density Functional Theory Models

Suppose finite element approximate spaceX_h is [67]:

$$X_{h,G-DFT} = \left\{ \begin{array}{l} \Phi_h = (\phi_{h,i})_{i=1}^{N} = (\phi_{h,1i})_{i=1}^{n_\alpha+n_\beta} + i(\phi_{h,2i})_{i=1}^{n_\alpha+n_\beta}; \\ Z_h = (\zeta_{h,i})_{i=1}^{2n_\alpha+2n_\beta} = ((\phi_{h,1i})_{i=1}^{n_\alpha+n_\beta}, (\phi_{h,2i})_{i=1}^{n_\alpha+n_\beta}) \\ Z_h \in (W_0^{m,p}(\Omega))^{2n_\alpha+n_\beta}; \\ \rho_h = \sum_{i=1}^{n_\alpha+n_\beta} |\phi_{h,i}|^2 = \sum_{i=1}^{2n_\alpha+2n_\beta} \zeta_{h,i}^2; \\ \langle \phi_{h,i}, \phi_{h,j} \rangle_{(L^2(\Omega),L^2(\Omega))} = \delta_{ij}, i, j = 1, \ldots, n_\alpha + n_\beta; \\ \varphi_h \in H_0^1(\Omega), \end{array} \right\}. \tag{3.5.42}$$

We define the discrete energy functional (3.2.26) with a finite element approximation as [67]

$$\mathcal{E}_h(\rho) = \left\{ \begin{array}{l} T(\rho) + \mathcal{J}_h(\rho) + K(\rho), \text{ if } \Phi \in X_{h,G-DFT}, \\ +\infty, \text{ if } \Phi \in X \backslash X_{h,G-DFT}, \end{array} \right. \tag{3.5.43}$$

where $\mathcal{J}_h(\rho)$ is defined in (3.5.3).

Theorem 3.5.9 shows the convergence of finite element approximation for meta generalized gradient approximation models.

Theorem 3.5.9 ([67]). *Let (C1)-(C3) hold, then*

$$\lim_{h\to 0} \inf_{X_{G-DFT}} \mathcal{E}_h = \min_{X_{G-DFT}} \mathcal{E}. \tag{3.5.44}$$

Proof. The proof is almost same as the proof of Theorem 3.5.8. By Theorem 3.3.30, (3.4.40) and (C3), we get (3.5.44). □

3.6 Convergence of Finite Element Approximation with Numerical Quadratures

The Section 3.4 of Suryanarayana et al. paper [146] presented the convergence of the finite element approximation with numerical quadratures for local spin density approximation model. We generalize the proofs of Suryanarayana et al. [146] to restricted Hartree-Fock, unrestricted Hartree-Fock, generalized gradient approximation model, meta generalized gradient approximation model and generalized density functional theory models [67].

Coeffient matrix of finite element approximation needs to be calculated by numerical quadratures. In this section, we will discuss the conditions which need to be satisfied in order to have the convergence of the finite element approximation with numerical quadrature.

Suppose we use a m-th order quadrature method. Let $\widetilde{\mathcal{E}}$ and $Q\{\mathcal{E}\}$ denote a m-th order quadrature of $\mathcal{E}$. Let k denote the degree of the polynomial used in the finite element approximation.

Theorem 3.3.30 is used to prove the convergence of the finite element approximation with numerical quadratures, which needs to satisfy both Γ-convergence and equi-coercivity.

Section 3.6.1 shows the proof of Hartree-Fock methods [67]. Section 3.6.2 shows the proof of density functional theory models [67].

3.6.1 Hartree-Fock

The convergence of finite element approximation with numerical quadratures of restricted Hartree-Fock and unrestricted Hartree-Fock models will be discussed in this section [67].

Restricted Hartree-Fock

By the finite element approximation with numerical quadratures, we define the total energy functional $\widetilde{\mathcal{E}}_h$, coulomb energy functional $\widetilde{\mathcal{J}}_h$ and electrons exact exchange

energy $\widetilde{K}_h$ as [146, 67]

$$\widetilde{\mathcal{E}}_h(\Phi) = \begin{cases} \widetilde{T}(\Phi) + \widetilde{\mathcal{J}}_h(\Phi) + \widetilde{K}_h(\Phi), \text{ if } \Phi \in X_{h,RHF}; \\ +\infty, \text{ if } \Phi \in X \backslash X_{h,RHF}, \end{cases} \tag{3.6.1}$$

where

$$\widetilde{\mathcal{J}}_h(\Phi) = -\min_{\phi \in H_0^1(\Omega)} \widetilde{I}_h(\Phi, \varphi), \tag{3.6.2}$$

$$\widetilde{K}_h(\Phi) = \min_{\varphi_{K_{ij}} \in H_0^1(\Omega)} \widetilde{L}_h(\Phi, \varphi_{K_{ij}}), \tag{3.6.3}$$

where

$$\widetilde{I}_h(\Phi, \varphi) = \begin{cases} \widetilde{I}(\Phi, \varphi), \text{ if } \Phi, \varphi \in X_{h,RHF}, \\ +\infty, \text{ otherwise}, \end{cases} \tag{3.6.4}$$

$$\widetilde{L}_h(\Phi, \varphi_{K_{ij}}) = \begin{cases} \widetilde{L}(\Phi, \varphi_{K_{ij}}), \text{ if } \Phi, \varphi_{K_{ij}} \in X_{h,RHF}, \\ +\infty, \text{ otherwise}, \end{cases} \tag{3.6.5}$$

where $I(\Phi, \varphi)$ and $L(\Phi, \varphi_{K_{ij}})$ have been defined in (3.5.7) and (3.5.8).

Lemma 3.6.1 and Lemma 3.6.2 prove $\lim_{h\to 0} \widetilde{\mathcal{J}}_h(\Phi_h) = \mathcal{J}(\Phi)$ and $\lim_{h\to 0} \widetilde{K}_h(\Phi_h) = K(\Phi)$, and they will be used to show the weakly Γ-convergence and weak equi-coercivity of $\widetilde{\mathcal{E}}_h$.

Lemma 3.6.1 ([67])**.** *If $(\Phi_h) \in (X_{h,RHF})$ is a sequence such that $\Phi_h \rightharpoonup \Phi$ in X_{RHF} and $m - 2k + 3 > 0$, then*

$$\lim_{h\to 0} \widetilde{\mathcal{J}}_h(\Phi_h) = \mathcal{J}(\Phi). \tag{3.6.6}$$

Proof. [45, Theorem 16]. We define $\triangle I_h$ as

$$\begin{aligned} \triangle I_h(\Phi, \varphi) &= \widetilde{I}_h(\Phi, \varphi) - I_h(\Phi, \varphi) \\ &= \begin{cases} \widetilde{I}(\Phi, \varphi) - I(\Phi, \varphi) \text{ if } \Phi, \varphi \in X_{h,RHF}, \\ 0, \text{ otherwise.} \end{cases} \end{aligned} \tag{3.6.7}$$

As a result of Theorem 3.3.21, we have

$$|\triangle I_h(\Phi, \varphi)| \leq C h^{m+1} \int_\Omega \left| D^{m+1} \left[(\nabla \varphi(\vec{x}))^2 - (\rho(\vec{x}) + b(\vec{x}))\varphi(\vec{x}) \right] \right| d\vec{x}. \tag{3.6.8}$$

If $m-2k+3>0$, $D^{m+1}|(\nabla\varphi)^2|=0$. Therefore,

$$\begin{aligned}
|\triangle I_h(\Phi,\varphi)| \leq \quad & C_0 h^{m+1}\int_\Omega \left|D^{m+1}\left[(\rho(\vec{x})+b(\vec{x}))\varphi(\vec{x})\right]\right| d\vec{x},\\
\leq \quad & C_0 h\int_\Omega \left|D\left[(\rho(\vec{x})+b(\vec{x}))\varphi(\vec{x})\right]\right| d\vec{x},\\
\leq \quad & C_0 h\sum_{i=1}^{2n}\left\{\|D\phi_i\|_{L^2(\Omega)}\|\phi_i\|_{L^4(\Omega)}\|\varphi\|_{L^4(\Omega)}\right.\\
& \left.+(\|\phi_i\|_{L^4(\Omega)}\|\phi_i\|_{L^4(\Omega)}+C_1)\|D\varphi\|_{L^2(\Omega)}\right\},
\end{aligned}\tag{3.6.9}$$

where C_0 and C_1 are positive constants, which don't need to be same everywhere in the proof. The second inequality in (3.6.9) is by Inverse Inequality. The third inequality in (3.6.9) is by Hölder inequality and pointwise bounded b.

We will show Γ-convergence property of $\widetilde{I}_h$ in part (1) of the proof and equi-coercivity of $\widetilde{I}_h$ in part (2) of the proof.

(1) If $(\Phi_h)\in(X_{h,RHF})$ is a sequence such that $\Phi_h \rightharpoonup \Phi$ in X_{RHF}. If $\varphi\notin X_{h,RHF}$, $\triangle I_h(\Phi_h,\varphi)=0$. Otherwise,

$$\begin{aligned}
|\triangle I_h(\Phi_h,\varphi)| \leq \quad & C_0 h\sum_{i=1}^{2n}\left\{\|D\phi_{h,i}\|_{L^2(\Omega)}\|\phi_{h,i}\|_{L^4(\Omega)}\|\varphi\|_{L^4(\Omega)}\right.\\
& \left.+(\|\phi_{h,i}\|_{L^4(\Omega)}\|\phi_{h,i}\|_{L^4(\Omega)}+C_1)\|D\varphi\|_{L^2(\Omega)}\right\}.
\end{aligned}\tag{3.6.10}$$

By the definition of X_{RHF}, all the norms on the right side of (3.6.10) are uniformly bounded. Hence, it follows $\triangle I_h(\Phi_h,\varphi)$ is continuously convergent to the zero function. Because

$$\widetilde{I}_h(\Phi_h,\varphi)=I_h(\Phi_h,\varphi)+\triangle I_h(\Phi_h,\varphi),\tag{3.6.11}$$

Lemma 3.5.1 and Theorem 3.3.31, it follows that $\widetilde{I}_h$ Γ-converges to I in the weak topology of X_{RHF}.

(2) If $\varphi\in X_{RHF}$, from (3.6.9),

$$\begin{aligned}
\widetilde{I}_h(\Phi,\varphi) \geq \quad & I_h(\Phi_h,\varphi)-C_0 h\sum_{i=1}^{2n}\left\{\|D\phi_i\|_{L^2(\Omega)}\|\phi_i\|_{L^4(\Omega)}\|\varphi\|_{L^4(\Omega)}\right.\\
& \left.+(\|\phi_i\|_{L^4(\Omega)}\|\phi_i\|_{L^4(\Omega)}+C_1)\|D\varphi\|_{L^2(\Omega)}\right\},\\
\geq \quad & C_1\|\varphi\|^2_{H^1_0(\Omega)}-C_2 h\|\nabla\varphi\|_{L^2(\Omega)}-C_3 h\|\varphi\|_{L^4(\Omega)},\\
\geq \quad & C_1\|\varphi\|^2_{H^1_0(\Omega)}-C_2 h\|\nabla\varphi\|_{L^2(\Omega)}-C_3 h^{\frac{1}{4}}\|\varphi\|_{L^2(\Omega)},
\end{aligned}\tag{3.6.12}$$

where C_1, C_2, C_3 are postive constants. The third inequality in (3.6.12) is by Inverse Inequality. If $\varphi \notin X_{RHF}$, $\widetilde{I}_h = +\infty$. Thus, we have (3.6.12) satisfied. Because the right side in (3.6.12) is coercive in the weak toplogy of X_{RHF}, $\widetilde{I}_h$ is equi-coercive in the weak toplogy of X_{RHF}.

Based on (1), (2) and Theorem 3.3.30, we have (3.6.6). □

Lemma 3.6.2 ([67]). *If $(\Phi_h) \in (X_{h,RHF})$ is a sequence such that $\Phi_h \rightharpoonup \Phi$ in X_{RHF} and $m - 2k + 3 > 0$, then*

$$\lim_{h\to 0} \widetilde{K}_h(\Phi_h) = K(\Phi). \tag{3.6.13}$$

Proof. We define $\triangle L_h$ as

$$\begin{aligned} \triangle L_h(\Phi, \varphi_{K_{ij}}) &= \widetilde{L}_h(\Phi, \varphi_{K_{ij}}) - L_h(\Phi, \varphi_{K_{ij}}) \\ &= \begin{cases} \widetilde{L}(\Phi, \varphi_{K_{ij}}) - L(\Phi, \varphi_{K_{ij}}), \text{ if } \Phi_h, \varphi_{K_{ij}} \in X_{h,RHF}, \\ 0, \text{ otherwise.} \end{cases} \end{aligned} \tag{3.6.14}$$

As a result of Theorem 3.3.21, we have

$$|\triangle L_h(\Phi, \varphi_{K_{ij}})| \le Ch^{m+1} \sum_{i,j=1}^{2n} \int_\Omega \left| D^{m+1} \left[(\nabla \varphi_{K_{ij}})^2 - \phi_i^* \phi_j \varphi_{K_{ij}} \right] \right| dx. \tag{3.6.15}$$

If $m - 2k + 3 > 0$, $D^{m+1}|(\nabla \varphi_{K_{ij}})^2| = 0$. Therefore,

$$\begin{aligned} |\triangle L_h(\Phi, \varphi_{K_{ij}})| &\le Ch^{m+1} \sum_{i,j=1}^{2n} \int_\Omega \left| D^{n+1}(\phi_i^* \phi_j \varphi_{K_{ij}}) \right| d\vec{x}, \\ &\le Ch \int_\Omega \sum_{i,j=1}^{2n} \left| D(\phi_i^* \phi_j \varphi_{K_{ij}}) \right| dx, \\ &\le Ch \sum_{i,j=1}^{2n} \big\{ \|\nabla \phi_i\|_{L^2(\Omega)} \|\phi_j\|_{L^4(\Omega)} \|\varphi_{K_{ij}}\|_{L^4(\Omega)} \\ &\quad + \|\phi_i\|_{L^4(\Omega)} \|\nabla \phi_j\|_{L^2(\Omega)} \|\varphi_{K_{ij}}\|_{L^4(\Omega)} \\ &\quad + \|\phi_i\|_{L^4(\Omega)} \|\phi_j\|_{L^4(\Omega)} \|\nabla \varphi_{K_{ij}}\|_{L^2(\Omega)} \big\}, \end{aligned} \tag{3.6.16}$$

where C is a positive constant, which don't need to be same everywhere in the proof. The second inequality in (3.6.16) is by Inverse Inequality. The third inequality in (3.6.16) is by Hölder inequality.

We will show Γ-convergence property of $\widetilde{L}_h$ in part (1) of the proof and equi-coercivity

of $\widetilde{L}_h$ in part (2) of the proof.

(1)If $(\Phi_h) \in (X_{h,RHF})$ is a sequence such that $\Phi_h \rightharpoonup \Phi$ in X_{RHF}. If $\varphi_{K_{ij}} \notin X_{h,RHF}$, then $\triangle L_h(\Phi_h, \varphi_{K_{ij}}) = 0$. Otherwise,

$$\begin{aligned} |\triangle L_h(\Phi_h, \varphi_{K_{ij}})| \quad \leq \quad & Ch \sum_{i,j=1}^{2n} \big\{ \|\nabla \phi_{h,i}\|_{L^2(\Omega)} \|\phi_{h,j}\|_{L^4(\Omega)} \|\varphi_{K_{ij}}\|_{L^4(\Omega)} \\ & + \|\phi_{h,i}\|_{L^4(\Omega)} \|\nabla \phi_{h,j}\|_{L^2(\Omega)} \|\varphi_{K_{ij}}\|_{L^4(\Omega)} \\ & + \|\phi_{h,i}\|_{L^4(\Omega)} \|\phi_{h,j}\|_{L^4(\Omega)} \|\nabla \varphi_{K_{ij}}\|_{L^2(\Omega)} \big\} \end{aligned} \tag{3.6.17}$$

By the definition of X_{RHF}, all the norms on the right side of (3.6.17) are uniformly bounded. Hence, it follows $\triangle L_h(\Phi_h, \varphi_{K_{ij}})$ is continuously convergent to the zero function. Because

$$\widetilde{L}_h(\Phi_h, \varphi_{K_{ij}}) = L_h(\Phi_h, \varphi_{K_{ij}}) + \triangle L_h(\Phi_h, \varphi_{K_{ij}}), \tag{3.6.18}$$

Lemma 3.5.1 and Theorem 3.3.31, it follows that $\widetilde{L}_h$ Γ-converges to L in the weak topology of X_{RHF}.

(2) If $\varphi_{K_{ij}} \in X_{RHF}$, from (3.6.17),

$$\begin{aligned} \widetilde{L}_h(\Phi, \varphi_{K_{ij}}) \geq \quad & L_h(\Phi_h, \varphi_{K_{ij}}) - Ch \sum_{i,j=1}^{2n} \big\{ \|\nabla \phi_{h,i}\|_{L^2(\Omega)} \|\phi_{h,j}\|_{L^4(\Omega)} \|\varphi_{K_{ij}}\|_{L^4(\Omega)} \\ & + \|\phi_{h,i}\|_{L^4(\Omega)} \|\nabla \phi_{h,j}\|_{L^2(\Omega)} \|\varphi_{K_{ij}}\|_{L^4(\Omega)} \\ & + \|\phi_{h,i}\|_{L^4(\Omega)} \|\phi_{h,j}\|_{L^4(\Omega)} \|\nabla \varphi_{K_{ij}}\|_{L^2(\Omega)} \big\} \\ \geq \quad & \sum_{i,j=1}^{2n} \left\{ C_1 \|\varphi_{K_{ij}}\|^2_{H^1_0(\Omega)} - C_2 h \|\nabla \varphi_{K_{ij}}\|_{L^2(\Omega)} - C_3 h \|\varphi_{K_{ij}}\|_{L^4(\Omega)} \right\}, \\ \geq \quad & \sum_{i,j=1}^{2n} \left\{ C_1 \|\varphi_{K_{ij}}\|^2_{H^1_0(\Omega)} - C_2 h \|\nabla \varphi_{K_{ij}}\|_{L^2(\Omega)} - C_3 h^{\frac{1}{4}} \|\varphi_{K_{ij}}\|_{L^2(\Omega)} \right\}, \end{aligned} \tag{3.6.19}$$

where C_1, C_2, C_3 are postive constants. The third inequality in (3.6.19) is by Inverse Inequality. If $\varphi_{K_{ij}} \notin X_{RHF}$, then $\widetilde{L}_h = +\infty$. Thus, we have (3.6.19) satisfied. Because the right side in (3.6.19) is coercive in the weak toplogy of X_{RHF}, $\widetilde{L}_h$ is equi-coercive in the weak toplogy of X_{RHF}.

Based on (1), (2) and Theorem 3.3.30, we have (3.6.13). □

Lemma 3.6.3 will be used to show the convergence of finite element approximation with numerical quadratures.

Lemma 3.6.3 ([67]). *If $(\Phi_h) \in (X_{h,RHF})$ is a sequence such that $\Phi_h \rightharpoonup \Phi$ in X_{RHF} and $m - 2k + 3 > 0$, then*

$$\lim_{h\to 0}\{\widetilde{\mathcal{E}}_{h,RHF}(\Phi_h) - \mathcal{E}_{h,RHF}(\Phi_h)\} = 0. \tag{3.6.20}$$

Proof. We know

$$\begin{aligned} |\widetilde{\mathcal{E}}_h(\Phi_h) - \mathcal{E}_h(\Phi_h)| \;\leq\; & Ch^{m+1}\int_\Omega \left| D^{m+1}(\sum_{i=1}^{2n} |\nabla\phi_{h,i}|^2)\right| dx + |\widetilde{\mathcal{J}}_h(\Phi_h) - \mathcal{J}_h(\Phi_h)| \\ & + |\widetilde{K}_h(\Phi_h) - K_h(\Phi_h)|. \end{aligned} \tag{3.6.21}$$

Due to $m-2k+3 > 0$, the first term in (3.6.21) is zero. And as a result of Lemma 3.5.1, Lemma 3.5.2, Lemma 3.6.1 and Lemma 3.6.2, the last two terms in (3.6.21) is zero as h goes to zero. Hence, we can get (3.6.20). □

Theorem 3.6.4 shows the convergence of finite element approximation with numerical quadratures for the restricted Hartree-Fock model.

Theorem 3.6.4 ([67]). *If $m - 2k + 3 > 0$, then*

$$\lim_{h\to 0}\inf_{X_{RHF}} \widetilde{\mathcal{E}}_h = \min_{X_{RHF}} \mathcal{E}. \tag{3.6.22}$$

Proof. We will show Γ-convergence property of $\widetilde{\mathcal{E}}_h$ in part (1) of the proof and equicoercivity of $\widetilde{\mathcal{E}}_h$ in part (2) of the proof.
(1) Let (Φ_h) be a sequence and $\Phi_h \rightharpoonup \Phi$ in X_{RHF}. If there is no subsequence $\Phi_{h_k} \in X_{h_k,RHF}$, then

$$+\infty = \liminf_{k\to\infty} \mathcal{E}_{h_k}(\Phi_{h_k}) \geq \mathcal{E}(\Phi). \tag{3.6.23}$$

If there is a subsequence $\Phi_{h_k} \in X_{h_k,RHF}$, then

$$\begin{aligned} \liminf_{k\to\infty} \widetilde{\mathcal{E}}_{h_k}(\Phi_{h_k}) \;&\geq\; \liminf_{k\to\infty} (\widetilde{\mathcal{E}}_{h_k}(\Phi_{h_k}) - \mathcal{E}_{h_k}(\Phi_{h_k})) + \liminf_{k\to\infty} \mathcal{E}_{h_k}(\Phi_{h_k}) \\ &\geq\; \liminf_{k\to\infty} \mathcal{E}_{h_k}(\Phi_{h_k}) \\ &\geq\; \mathcal{E}(\Phi) \end{aligned} \tag{3.6.24}$$

The second inequality in (3.6.24) is by Lemma 3.6.1 and the third inequality in (3.6.24) is by Lemma 3.5.3. Thus, we have the liminf property.

The recovery sequence is also from the interpolated functions of successive triangulations. Thus, we can get $\widetilde{\mathcal{E}}_h$ Γ-converges to $\mathcal{E}$ in the weak topology of X_{RHF}.
(2) Due to Lemma 3.5.4,

$$\mathcal{E}_h(\Phi) \geq C_1 \|\nabla\Phi\|^2_{L^2(\Omega)} + C_2, \tag{3.6.25}$$

where C_1 and C_2 are positive constants. Then by (3.6.21),

$$\widetilde{\mathcal{E}}_h(\Phi) \geq C_1 \|\nabla\Phi\|^2_{L^2(\Omega)} + C_2 - |\widetilde{\mathcal{J}}_h(\Phi) - \mathcal{J}_h(\Phi)| - |\widetilde{K}_h(\Phi) - K_h(\Phi)|. \tag{3.6.26}$$

By Lemma 3.5.1, Lemma 3.5.2, Lemma 3.6.1 and Lemma 3.6.2, there exists a bound h_0 such that for all $h < h_0$, we have

$$\widetilde{\mathcal{E}}_h(\Phi) \geq C_1 \|\nabla\Phi\|^2_{L^2(\Omega)} + C_2, \tag{3.6.27}$$

where C_1 and C_2 are positive constants independent of h. Because the right side of (3.6.27) is a weakly coercive lower semi-continuous functional, we have $\widetilde{\mathcal{E}}_h(\Phi)$ is equi-coercive in the weak topology of X_{RHF}.
Based on (1), (2) and Theorem 4.3., we have (3.6.22). □

Unrestricted Hartree-Fock

Theorem 3.6.5 shows the convergence of finite element approximation with numerical quadratures for the unrestricted Hartree-Fock model.

Theorem 3.6.5 ([67])**.**

$$\lim_{h\to 0} \inf_{X_{UHF}} \widetilde{\mathcal{E}}_h = \min_{X_{UHF}} \mathcal{E}. \tag{3.6.28}$$

Proof. The only difference between restricted Hartree-Fock and unrestricted Hartree-Fock is the number of spin-up electrons and spin-down electrons. The proof is exactly the same as Section restricted Hartree-Fock. □

3.6.2 Density Functional Theory

In this section, we discuss the convergence of finite element approximation with numerial quadratures of density functional theory models, including local spin density approximation, generalized gradient approximation, meta generalized gradient approximation and generalized density functional theory models [67].

Local Spin Density Approximation

The convergence of the finite element approximation of local spin density approximation has been carefully studied in Section 3.4 of Suryanarayana et al. paper [146]. More to the local spin density approximation, we will discuss other density functional theory models below [67].

Generalized Gradient Approximation

We define the total energy functional with a finite element approximation where all integrations are performed with numerical quadratures as [67]

$$\widetilde{\mathcal{E}}_h(\rho) = \begin{cases} \widetilde{T}(\rho) + \widetilde{J}_h(\rho) + \widetilde{K}(\rho), \text{ if } \Phi \in X_{h,GGA}; \\ +\infty, \text{ if } \Phi \notin X_{h,GGA}, \end{cases} \tag{3.6.29}$$

where $\mathcal{J}_h(\rho)$ is defined in (3.5.3).

Theorem 3.6.6 shows the convergence of finite element approximation with numerical quadratures for generalized gradient approximation models.

Theorem 3.6.6 ([67]). *Let (A1), (A2) hold,* $m - 2k + 3 > 0$ *and* $\|\nabla \mathbb{k}\|_{L^1(\Omega)}$ *is bounded independently of* h*, then*

$$\lim_{h\to 0} \inf_{X_{GGA}} \widetilde{\mathcal{E}}_h = \min_{X_{GGA}} \mathcal{E}. \tag{3.6.30}$$

Proof. We will show Γ-convergence property of $\widetilde{\mathcal{E}}_h$ in part (1) of the proof and equicoercivity of $\widetilde{\mathcal{E}}_h$ in part (2) of the proof.
(1) If $(\Phi_h) \in (X_{h,GGA})$ is a sequence such that $\Phi_h \rightharpoonup \Phi$ in X_{GGA}, then

$$\begin{aligned} & |\widetilde{\mathcal{E}}_h(\rho_h) - \mathcal{E}_h(\rho_h)| \\ \leq \quad & C_0 h^{m+1} \int_\Omega \left| D^{m+1}\Big(\sum_{i=1}^{n_\alpha+n_\beta} |\nabla \phi_{h,i}|^2 + \mathbb{k}(\rho)\Big)\right| d\vec{x} + |\widetilde{\mathcal{J}}_h(\rho_h) - \mathcal{J}_h(\rho_h)| \\ \leq \quad & C_0 h^{m+1} \int_\Omega \left| D^{m+1}\Big(\sum_{i=1}^{n_\alpha+n_\beta} |\nabla \phi_{h,i}|^2\Big)\right| d\vec{x} + |\widetilde{J}_h(\rho_h) - J_h(\rho_h)| + C_1 h \|\nabla \mathbb{k}\|_{L^1(\Omega)}. \end{aligned} \tag{3.6.31}$$

Due to $m - 2k + 3 > 0$, the first term in (3.6.31) is zero. Because $\|\nabla \mathbb{k}\|_{L^1(\Omega)}$ is bounded independently of h and Lemma 3.6.1,

$$\lim_{h\to 0} \{\widetilde{\mathcal{E}}_h(\rho_h) - \mathcal{E}_h(\rho_h)\} = 0. \tag{3.6.32}$$

Let (Φ_h) be a sequence and $\Phi_h \rightharpoonup \Phi$ in X_{GGA}. If there is no subsequence $\Phi_{h_k} \in X_{h_k,GGA}$, then

$$+\infty = \liminf_{k\to\infty} \mathcal{E}_{h_k}(\rho_{h_k}) \geq \mathcal{E}(\rho). \tag{3.6.33}$$

If there is a subsequence $\rho_{h_k} \in X_{h_k,GGA}$, then

$$\begin{aligned} \liminf_{k\to\infty} \widetilde{\mathcal{E}}_{h_k}(\rho_{h_k}) &\geq \liminf_{k\to\infty} (\widetilde{\mathcal{E}}_{h_k}(\rho_{h_k}) - \mathcal{E}_{h_k}(\rho_{h_k})) + \liminf_{k\to\infty} \mathcal{E}_{h_k}(\rho_{h_k}) \\ &\geq \liminf_{k\to\infty} \mathcal{E}_{h_k}(\rho_{h_k}) \\ &\geq \mathcal{E}(\rho) \end{aligned} \tag{3.6.34}$$

The second inequality in (3.6.34) is by (3.6.33) and the third inequality in (3.6.34) is by Theorem 3.5.7. Thus, we have the liminf property. The interpolation functions of successive triangulations can also be constructed to be a recovery sequence. We have

$$\lim_{h\to 0} \widetilde{\mathcal{E}}_h = \lim_{h\to 0} \mathcal{E}_h = \mathcal{E}. \tag{3.6.35}$$

Thus, $\widetilde{\mathcal{E}}_h$ Γ-converges to $\mathcal{E}$ in the weak topology of X_{GGA}.

(2) Due to Theorem 3.6.31, we have

$$\begin{aligned} \mathcal{E}_h(\rho) &\geq c_0\|\nabla\Phi\|^2_{L^2(\Omega)} + c_1\|\Phi\|^3_{L^3(\Omega)} - c_2\|\Phi\|^2_{L^2(\Omega)} \\ &\quad -c_3\left[\|\Phi\|^{2k_1\alpha}_{L^{2k_1\alpha}(\Omega)} + \|\Phi\|^{k_2\beta}_{L^{k_2\beta}(\Omega)} + \|\nabla\Phi\|^{k_3\beta}_{L^{k_3\beta}(\Omega)}\right] - c_4. \end{aligned} \tag{3.6.36}$$

Then by (3.6.31),

$$\begin{aligned} \widetilde{\mathcal{E}}_h(\rho) \geq\ & c_0\|\nabla\Phi\|^2_{L^2(\Omega)} + c_1\|\Phi\|^3_{L^3(\Omega)} - c_2\|\Phi\|^2_{L^2(\Omega)} \\ & -c_3\left[\|\Phi\|^{2k_1\alpha}_{L^{2k_1\alpha}(\Omega)} + \|\Phi\|^{k_2\beta}_{L^{k_2\beta}(\Omega)} + \|\nabla\Phi\|^{k_3\beta}_{L^{k_3\beta}(\Omega)}\right] - c_4 \\ & -|\widetilde{J}_h(\rho_h) - J_h(\rho_h)| - c_5 h\|\nabla\mathbb{k}\|_{L^1(\Omega)}, \end{aligned} \tag{3.6.37}$$

where c_1, c_2, c_3, c_4, c_5 are positive constants. There exists a bound h_0 such that for all $h < h_0$, we have

$$\begin{aligned} \widetilde{\mathcal{E}}_h(\rho) &\geq c_0\|\nabla\Phi\|^2_{L^2(\Omega)} + c_1\|\Phi\|^3_{L^3(\Omega)} - c_2\|\Phi\|^2_{L^2(\Omega)} \\ &\quad -c_3\left[\|\Phi\|^{2k_1\alpha}_{L^{2k_1\alpha}(\Omega)} + \|\Phi\|^{k_2\beta}_{L^{k_2\beta}(\Omega)} + \|\nabla\Phi\|^{k_3\beta}_{L^{k_3\beta}(\Omega)}\right] - c_4. \end{aligned} \tag{3.6.38}$$

where c_1, c_2, c_3, c_4, c_5 are positive constants independent of h. Because the right side of (3.6.38) is coercive in the weak topology of X_{GGA}, $\widetilde{\mathcal{E}}_h(\rho)$ is equi-coercive in the weak topology of X_{GGA}.

Based on (1), (2) and Theorem 4.3., we have (3.6.30). □

Meta Generalized Gradient Approximation

We define the total energy functional with a finite element approximation where all integrations are performed with numerical quadratures as [67]

$$\widetilde{\mathcal{E}}_h(\rho) = \begin{cases} \widetilde{T}(\rho) + \widetilde{J}_h(\rho) + \widetilde{K}(\rho), \text{ if } \Phi \in X_{h,Meta-GGA}, \\ +\infty, \text{ if } \Phi \notin X_{h,Meta-GGA}, \end{cases} \tag{3.6.39}$$

where $\mathcal{J}_h(\rho)$ is defined in (3.5.3).

Theorem 3.6.7 shows the convergence of finite element approximation with numerical quadratures for meta generalized gradient approximation models.

Theorem 3.6.7 ([67]). *Let (B1)-(B3) hold,* $m - 2k + 3 > 0$ *and* $\|\nabla \mathbb{k}\|_{L^1(\Omega)}$ *is bounded independently of* h*, then*

$$\lim_{h\to 0} \inf_{X_{Meta-GGA}} \widetilde{\mathcal{E}}_h = \min_{X_{Meta-GGA}} \mathcal{E}. \tag{3.6.40}$$

Proof. We will show Γ-convergence property of $\widetilde{\mathcal{E}}_h$ in part (1) of the proof and equi-coercivity of $\widetilde{\mathcal{E}}_h$ in part (2) of the proof.
(1) Theorem 3.5.8 shows $\mathcal{E}_h$ Γ-converges to $\mathcal{E}$ in the weak topology of $X_{Meta-GGA}$. The rest of the proof is almost same with part (1) of Theorem 3.6.6.
(2) There exists a bound h_0 such that for all $h < h_0$, we have

$$\widetilde{\mathcal{E}}_h(\rho) \geq \mathcal{E}_h(\rho) - C, \tag{3.6.41}$$

where C is a positive constant. By equi-coercivity of $\mathcal{E}_h(\rho)$ in Theorem 3.5.8, $\widetilde{\mathcal{E}}_h$ is equi-coercive in the weak topology of $X_{Meta-GGA}$.
Based on (1), (2) and Theorem 4.3., we have (3.6.40). □

Generalized Density Functional Theory Models

We define the total energy functional with a finite element approximation where all integrations are performed with numerical quadratures $\widetilde{\mathcal{E}}_h$ as [67]

$$\widetilde{\mathcal{E}}_h(\rho) = \begin{cases} \widetilde{T}(\rho) + \widetilde{J}_h(\rho) + \widetilde{K}(\rho), \text{ if } \Phi \in X_{h,G-DFT}, \\ +\infty, \text{ if } \Phi \notin X_{h,G-DFT}, \end{cases} \tag{3.6.42}$$

where $\mathcal{J}_h(\rho)$ is defined in (3.5.3).

Theorem 3.6.8 shows the convergence of finite element approximation with numerical quadratures for generalized density functional theory models.

Theorem 3.6.8 ([67]). *Let (C1)-(C3) hold,* $m - 2k + 3 > 0$ *and* $\|\nabla \mathbb{k}\|_{L^1(\Omega)}$ *is bounded independently of* h*, then*

$$\lim_{h \to 0} \inf_{X_{G-DFT}} \widetilde{\mathcal{E}}_h = \min_{X_{G-DFT}} \mathcal{E}. \tag{3.6.43}$$

Proof. The proof is the same as the proof of Theorem 3.6.7. □

3.7 Convergence of Pseudopotential Approximation

The Section 3.5 of Suryanarayana et al. paper [146] presented the convergence of pseudopotential approximation for local spin density approximation model. We generalize the proofs of Suryanarayana et al. [146] to restricted Hartree-Fock, unrestricted Hartree-Fock, generalized gradient approximation model, meta generalized gradient approximation model and generalized density functional theory models [67].

Pseudopotential approximation is commonly used to smooth the nuclear electrons coulomb interaction, and it can be broadly classified as local and non-local.

A local pseudopotential is an explicit function V_{ext}^{PS} with $\|V_{ext}^{PS}\|_{L^\infty(\Omega)} < C$, where C is a positive constant. The local pseudopotential approximation can be incoporated into our problem by replacing b by $b^{PS} = -C\nabla^2 V_{ext}^{PS}$. All the results presented in the previous sections are applicable and hence existence of a minimizer, convergence of the finite element appoximation and convergence of fthe inite-element appoximation with numerical quadratures follow.

For non-local pseudopotential, we can write total energy functional as [146, 67]:

$$\mathcal{E}^{PS} = \mathcal{E} + P, \tag{3.7.1}$$

$$P = \sum_{i=1}^{N} \sum_{J=1}^{M} \sum_{l,t} \frac{1}{G_{l,t}^J} \left| \int_\Omega f(\vec{x}, \vec{x}_J) \phi_i(\vec{x}) d\vec{x} \right|^2. \tag{3.7.2}$$

Where l is the azimuthal quantum number, t is the magnetic quantum number, J is the atom number, N is the total number of electrons after the pseudopotential approximation and $G_{l,t}^J$ is a nonzero constant. In our proof below, we only consider non-local pseudopotential.

Section 3.7.1 shows the proof of Hartree-Fock methods [67]. Section 3.7.2 shows the proof of density functional theory models [67].

3.7.1 Hartree-Fock

The convergence of pseudopotential approximation of restricted Hartree-Fock and unrestricted Hartree-Fock models will be discussed in this section [67].

Restricted Hartree-Fock

Theorem 3.7.1 shows the convergence of pseudopotential approximation for the restricted Hartree-Fock model.

Theorem 3.7.1 ([67]). *Let $m - 2k + 3 > 0$, and $f \in H_0^1(\Omega)$. Then $\mathcal{E}^{PS}$ possesses a minimizer in X_{RHF}, and*

$$\lim_{h\to 0} \inf_{X_{RHF}} \mathcal{E}_h^{PS} = \min_{X_{RHF}} \mathcal{E}. \tag{3.7.3}$$

$$\lim_{h\to 0} \inf_{X_{RHF}} \widetilde{\mathcal{E}}_h^{PS} = \min_{X_{RHF}} \mathcal{E}. \tag{3.7.4}$$

Proof. By $f \in H_0^1(\Omega)$ and Hölder inequality, $P(\Phi)$ is continuous in $(L^2(\Omega))^{2n}$, and

$$P(\Phi) \geq -C\|\Phi\|^2_{L^2(\Omega)}, \tag{3.7.5}$$

where C is a positive constant dependent on $f, G^J_{l,t}$. By (3.4.13), we have

$$\frac{1}{2}J(\Phi) + K(\Phi) \geq 0. \tag{3.7.6}$$

By (3.4.30), we have

$$\mathcal{J}(\Phi) \geq c_0\|\Phi\|^3_{L^3(\Omega)} - \frac{c_0}{8\pi}\|\nabla\Phi\|^2_{L^2(\Omega)} - c1\|\Phi\|^2_{L^2(\Omega)} - c_2. \tag{3.7.7}$$

If we choose $c_0 < 4\pi$, then $\frac{1}{2} - \frac{c_0}{8\pi} = c_0' > 0$. We end up with

$$\begin{aligned} \mathcal{E}(\Phi) = \quad & T(\Phi) + \mathcal{J}(\Phi) + K(\Phi), \\ \geq \quad & c_0\|\nabla\Phi\|^2_{L^2(\Omega)} + c_1\|\Phi\|^3_{L^3(\Omega)} - c_2\|\Phi\|^2_{L^2(\Omega)} - c_3, \end{aligned} \tag{3.7.8}$$

where c_0, c_1, c_2 and c_3 are positive constants, which don't need to be same in the proof. Thus,

$$\begin{aligned} \mathcal{E}^{PS}(\Phi) = & \quad \mathcal{E}(\Phi) + P(\Phi), \\ \geq & \quad c_0\|\nabla\Phi\|^2_{L^2(\Omega)} + c_1\|\Phi\|^3_{L^3(\Omega)} - c_2\|\Phi\|^2_{L^2(\Omega)} - c_3, \end{aligned} \tag{3.7.9}$$

(1) Because $P(\Phi)$ is continuous in $(L^2(\Omega))^{2n}$ and Lemma 3.4.7, $\mathcal{E}^{PS}(\Phi)$ is lower semi-continuous in the weak topology of X_{RHF}. (3.7.8) shows $\mathcal{E}^{PS}(\Phi)$ is coercive in the weak topology of X_{RHF}. Thus, $\mathcal{E}^{PS}$ possesses a minimizer in X_{RHF} by Theorem 3.3.29.
(2) The Γ-convergence property of $\mathcal{E}^{PS}_h$ in the weak topology of X_{RHF} follows from continuous property of $P(\Phi)$ and the process in Lemma 3.5.3. The equi-coercivity of $\mathcal{E}^{PS}_h$ in the weak topology of X_{RHF} follows from (3.7.8) and the process in Lemma 3.5.4. Hence, we can get (3.7.3) by Theorem 3.5.5.
(3) The error of $P(\Phi)$ due to numerical quadrature is

$$\begin{aligned} |P(\Phi) - \widetilde{P}_h(\Phi)| \leq & \quad Ch^{m+1}\sum_{i=1}^{2n}\sum_{J=1}^{M}\sum_{l,t}\left|\int_\Omega D^{m+1}(f(\vec{x},\vec{x}_J)\phi_i(\vec{x}))d\vec{x}\right|^2, \\ \leq & \quad Ch\sum_{i=1}^{2n}\sum_{J=1}^{M}\sum_{l,t}\left[\|\nabla f\|_{L^2(\Omega)}\|\phi_i\|_{L^2(\Omega)} + \|f\|_{L^2(\Omega)}\|\nabla\phi_i\|_{L^2(\Omega)}\right], \end{aligned} \tag{3.7.10}$$

where C is a positive constant. The second inequality in (3.7.10) is by Inverse Inequality. Because $f \in H^1_0(\Omega)$ and (3.7.10), we have $\lim_{h\to 0}|P(\Phi) - \widetilde{P}_h(\Phi)| = 0$. By the same process in Theorem 3.6.4, we can get (3.7.4). □

Unrestricted Hartree-Fock

Theorem 3.7.2 shows the convergence of pseudopotential approximation for the unrestricted Hartree-Fock model.

Theorem 3.7.2 ([67]). *Let $m - 2k + 3 > 0$, and $f \in H^1_0(\Omega)$. If $P(\Phi)$ is lower bounded, $\mathcal{E}^{PS}$ possesses a minimizer in X_{UHF}, and*

$$\lim_{h\to 0}\inf_{X_{UHF}}\mathcal{E}^{PS}_h = \min_{X_{UHF}}\mathcal{E}. \tag{3.7.11}$$

$$\lim_{h\to 0}\inf_{X_{UHF}}\widetilde{\mathcal{E}}^{PS}_h = \min_{X_{UHF}}\mathcal{E}. \tag{3.7.12}$$

Proof. Because $P(\Phi)$ is lower bounded and (3.4.15), we have

$$\mathcal{E} \geq C_1 \|\Phi\|^2_{H_0^1(\Omega)} - C_2, \tag{3.7.13}$$

where C_1 and C_2 are positive constants.
All the proof left is the same as Part (1), Part (2) and Part (3) of Theorem 3.7.1.

□

3.7.2 Density Functional Theory

In this section, we discuss the convergence of pseudopotential approximation of density functional theory models, including local spin density approximation, generalized gradient approximation, meta generalized gradient approximation and generalized density functional theory models.

Local Spin Density Approximation

The convergence of the finite element approximation of local spin density approximation model has been carefully studied in Section 3.5 of Suryanarayana et al. paper [146]. More to the local spin density approximation, we will discuss other density functional theory models below [67].

Generalized Gradient Approximation

Theorem 3.7.3 shows the convergence of pseudopotential approximation for generalized gradient approximation models.

Theorem 3.7.3 ([67])**.** *Let (A1), (A2) hold,* $m - 2k + 3 > 0$, $\|\nabla \mathbb{k}\|_{L^1(\Omega)}$ *is bounded independently of* h*, and* $f \in H_0^1(\Omega)$*. Then* $\mathcal{E}^{PS}$ *possesses a minimizer in* X_{GGA}*, and*

$$\lim_{h\to 0} \inf_{X_{GGA}} \mathcal{E}_h^{PS} = \min_{X_{GGA}} \mathcal{E}. \tag{3.7.14}$$

$$\lim_{h\to 0} \inf_{X_{GGA}} \widetilde{\mathcal{E}}_h^{PS} = \min_{X_{GGA}} \mathcal{E}. \tag{3.7.15}$$

Proof. By (3.4.32) and (3.7.5)

$$\begin{aligned}\mathcal{E}^{PS} \geq\ & c_0\|\nabla\Phi\|^2_{L^2(\Omega)} + c_1\|\Phi\|^3_{L^3(\Omega)} - c_2\|\Phi\|^2_{L^2(\Omega)} \\ & -c_3\sum_{\alpha,\beta}\left[\|\Phi\|^{2k_{1(\alpha,\beta)}\alpha}_{L^{2k_{1(\alpha,\beta)}\alpha}(\Omega)} + \|\Phi\|^{k_{2(\alpha,\beta)}\beta}_{L^{k_{2(\alpha,\beta)}\beta}(\Omega)} + \|\nabla\Phi\|^{k_3\beta}_{L^{k_{3(\alpha,\beta)}\beta}(\Omega)}\right] - c_4,\end{aligned} \tag{3.7.16}$$

(1) Because $P(\Phi)$ is continuous in $(L^2(\Omega))^{2n}$ and Lemma 3.4.18, $\mathcal{E}^{PS}$ is lower semi-continuous in the weak topology of X_{GGA}. (3.7.16) shows $\mathcal{E}^{PS}$ is coercive in the weak topology of X_{GGA}. Thus, $\mathcal{E}^{PS}$ possesses a minimizer in X_{GGA} by Theorem 3.3.29.
(2) We can get (3.7.14) by the process in Theorem 3.5.7.
(3) We can get (3.7.15) by (3.7.10) and the process in Theorem 3.6.6. □

Meta Generalized Gradient Approximation

Theorem 3.7.4 shows the convergence of pseudopotential approximation for meta generalized gradient approximation models.

Theorem 3.7.4 ([67]). *Let (B1)-(B3) hold, $m-2k+3>0$, $\|\nabla\mathbb{k}\|_{L^1(\Omega)}$ is bounded independently of h, and $f\in H^1_0(\Omega)$. If $\mathcal{E}^{PS}$ is coercive in the weak topology of $X_{Meta-GGA}$, $\mathcal{E}^{PS}$ possesses a minimizer in $X_{Meta-GGA}$, and*

$$\lim_{h\to 0}\inf_{X_{Meta-GGA}}\mathcal{E}^{PS}_h = \min_{X_{Meta-GGA}}\mathcal{E}. \tag{3.7.17}$$

$$\lim_{h\to 0}\inf_{X_{Meta-GGA}}\widetilde{\mathcal{E}}^{PS}_h = \min_{X_{Meta-GGA}}\mathcal{E}. \tag{3.7.18}$$

Proof. (1) Because $P(\Phi)$ is continuous in $(L^2(\Omega))^{2n}$ and Lemma 3.4.22, $\mathcal{E}^{PS}(\Phi)$ is lower semi-continuous in the weak topology of $X_{Meta-GGA}$. Also, $\mathcal{E}^{PS}$ is coercive in the weak topology of $X_{Meta-GGA}$. Thus, $\mathcal{E}^{PS}$ possesses a minimizer in $X_{Meta-GGA}$ by Theorem 3.3.29.
(2) We can get (3.7.14) by the process in Theorem 3.5.8.
(3) We can get (3.7.15) by (3.7.10) and the process in Theorem 3.6.7. □

Generalized Density Functional Theory Models

Theorem 3.7.5 shows the convergence of pseudopotential approximation for generalized density functional theory models.

Theorem 3.7.5 ([67]). *Let (C1)-(C3) hold, $m - 2k + 3 > 0$, $\|\nabla \mathbb{k}\|_{L^1(\Omega)}$ is bounded independently of h, and $f \in H_0^1(\Omega)$. Then $\mathcal{E}^{PS}$ possesses a minimizer in X_{G-DFT}, and*

$$\lim_{h\to 0} \inf_{X_{G-DFT}} \mathcal{E}_h^{PS} = \min_{X_{G-DFT}} \mathcal{E}. \tag{3.7.19}$$

$$\lim_{h\to 0} \inf_{X_{G-DFT}} \widetilde{\mathcal{E}}_h^{PS} = \min_{X_{G-DFT}} \mathcal{E}. \tag{3.7.20}$$

Proof. 1) Because $P(\Phi)$ is continuous in $(L^2(\Omega))^{2n}$ and Lemma 3.4.24, $\mathcal{E}^{PS}(\Phi)$ is lower semi-continuous in the weak topology of X_{G-DFT}. Also, $\mathcal{E}^{PS}$ is coercive in the weak topology of X_{G-DFT}. Thus, $\mathcal{E}^{PS}$ possesses a minimizer in X_{G-DFT} by Theorem 3.3.29.
(2) We can get (3.7.19) by the process in Theorem 3.5.9.
(3) We can get (3.7.20) by (3.7.10) and the process in Theorem 3.6.8. □

3.8 Conclusion

We construct a mathematical framework for Hartree-Fock and density functional theory models, including restricted Hartree-Fock, unrestricted Hartree-Fock, local spin density approximation, generalized gradient approximation, meta generalized gradient approximation and more generalized density functional theory models [67]. The well-posedness of these models, such as the existence of the minimizers, the convergence of finite element approximation, the convergence of finite element approximation with numerical quadratures, and the well-posedness of the pseudopotential approximation have been proved. It will be a great tool and extremely helpful for the development of Hartree-Fock and density functional theory models.

Chapter 3, is currently begin prepared for submission for publication of the material. Holst, M., Hu, H., and Zhu, Y. The dissertation author was the primary investigator and author of this material.

Chapter 4

New Linear Scaling Methods for Exact Exchange

We describe an efficient numerical algorithm to solve the Hartree-Fock equation [71]. The Hartree-Fock equation is discretized by the finite element method. Priori adaptive mesh is adopted to optimize the efficiency of the algorithm. The conjugate gradient method with an algebraic multigrid preconditioner is employed to solve the Poisson and bound-state Helmholtz equations. The computational cost scales linearly with the number of bases. A variety of numerical experiments for atoms and molecules demonstrate reliable precision and speed.

Section 4.1 shows the motivation to develop the linear scaling method. Section 4.2 discusses the algorithm of our method. Section 4.3 shows several numerical examples. Section 4.4 is our conclusion.

4.1 Motivation

Various approaches have been developed to investigate materials in the last century. Quantum methods, solving the Schrödinger equation with different types of approximations, have best accuracy comparing with other methods. The most commonly used quantum methods are wavefunction method [129] and density functional method [66, 81]. Wavefunction method is based on obtaining the wavefunction of the system. Density functional method studies the properties of the system through its electronic

density. A great deal of effort has been devoted to solve quantum models numerically. The plane-wave basis [82, 137, 51] is one of most-widely used basis for studying materials systems. The plane-wave basis functions form a complete and orthonormal set, and compute electrostatic interactions efficiently through Fourier transformations. However, there are still some notable disadvantages for the plane-wave basis. The plane-wave basis calculations are restricted to periodic boundary conditions and cannot handle most realistic systems with complicated boundary. Further, the uniform spatial resolution provided by the plane-wave basis is very inefficient to solve the non-periodic systems, where we need to concentrate the bases for some regions and provides low resolution elsewhere. Moreover, the scalability of computations on parallel computation is affected by the huge data flow during the orthogonalization process. Atomic-orbital type basis such as Gaussian basis [57, 161, 75] is another popular basis used. Atomic-orbital type basis includes important physics properties in the basis sets, and this reduces the amount of basis sets needed in the calculation dramatically. However, the self-consistent solver of this approach is difficult to get convergent due to the ill-conditioned matrices as the number of nonorthogonal basis functions increased. Furthermore, it is not flexible for complex geometries and boundary conditions, and it is not scalable on parallel computing platfomrs.

In order to avoid these drawbacks, much effort has been devoted to develop scalable real-space methods [8, 143, 26, 23, 48, 15, 141, 102] over the past decade, such as finite difference method, wavelet method, and finite element method. These real-space calculations are performed on meshes, and systematic convergence can be achieved. Among real-space methods, finite element method [160, 151, 152, 153, 113, 114, 115, 166, 22, 87, 146, 91, 36, 6, 105, 104, 96, 132, 27] is flexible for unstructured mesh, which allows for consideration of complex geometries and boundary conditions. Also finite element method has very good scalability on parallel computing platforms. Furthermore, a lot of numerical algorithms and existing software implemented by the finite element method such as FEniCS [34] could be used. The flexibility of mesh generation is helpful in Hartree-Fock and density functional theory calculations. The electronic wave functions have rapid variations near singular regions of the atomic center and decay quickly away from the atomic center. Thus, we need to provide high resolution for

regions with high electron density and a coarser resolution elsewhere. The mesh adaption technique is proposed to overcome this problem. Priori mesh adaption technique have been studied in [105, 104, 22, 87, 146, 36, 89, 128]. Most of them use existing knowledge of atomic wavefunction as a criterior to construct adaptive mesh, while some work [105, 104] proposed to use the error function of finite element approximation to generate adaptive function. Posteriori mesh adaption technique is also applied to density functional theory recently in papers [166, 6, 27].

Difficulties arise when finite element method is used to solve a model with exact exchange energy, such as Hartree-Fock model. Because exact exchange energy operator is non-local and finite element method needs more bases comparing with atomic-orbital type basis, the evaluation of exact exchange energy becomes a major computational bottleneck. However, exact exchange energy is included in most modern hybrid functionals [12, 120, 78, 145, 10, 86, 157, 11, 65]. Thus an efficient way to solve the exact exchange energy problem with finite element method needs to be developed. A few examples by assuming an exponential asymptotic behavior in the density matrix are proposed for solving Hartree-Fock exchange in the Gaussian bases [133, 135, 134]. Several algorithms have been developed to only solve the one dimensional Hartree-Fock equation in radial direction by the symmetry [42, 43, 59, 61, 60], but these algorithms are limited to small symmetrical systems. A recent Hartree-Fock method [2] implementation using finite element basis introduces an acceleration technique for the exact exchange energy which uses an approximate X-α formulation. Another interesting approach to evaluate the Hartree-Fock exact exchange energy used multiresolution self-consistent method [55, 164], which employs a fast integral convolution method with some approximations.

In this paper, we described a general linear scaling method with finite element basis to solve Hartree-Fock model. The linear scaling method will guarantee both the linear time cost and linear memory cost. This method could be easily generalized to density functional theory, and hybrid models [71], but we will focus only on the Hartree-Fock model in this paper.

4.2 Formulation

The formulations presented below is only for restricted Hartree-Fock, and they can be generalized to unrestricted Hartree-Fock by a straightforward way. This section follows Section 2.2 very closely. Section 4.2.1 introduces the Hartree-Fock equations [129]. Equivalent formulations of Hartree-Fock equations are described in Section 4.2.2, which are used in our linear scaling finite element solver [67]. Section 4.2.3 shows the finite element discretization of Hartree-Fock equations.

4.2.1 Hartree-Fock Equations

Restricted Hartree-Fock method assumes each occupied molecular spinorbital is occupied by two electrons with opposite spins. Suppose the electronic system has $2n$ electrons and n fully occupied molecular orbitals, we can denote the molecular spinorbitals by [129, 71]:

$$\Psi = \{\psi_1 \psi_2 \dots \psi_{2n}\}, \tag{4.2.1}$$

and molecular orbitals by [129, 71]:

$$\Phi = \{\phi_1 \phi_2 \dots \phi_n\}. \tag{4.2.2}$$

The relationship among molecular spinorbitals and molecular orbital are given by [129, 71]:

$$\psi_{2i-1} = \phi_i \alpha, \quad i = 1, \dots, n, \tag{4.2.3}$$

$$\psi_{2i} = \phi_i \beta, \quad i = 1, \dots, n. \tag{4.2.4}$$

And we have the orthonormality relationships [129, 71]:

$$\int \phi_i^*(\vec{x}) \phi_j(\vec{x}) d\vec{x} = \delta_{ij}, \quad i, j = 1, \dots, n. \tag{4.2.5}$$

The total energy functional of restricted Hartree-Fock is [129, 71]:

$$\mathcal{E} = T(\Psi) + V(\Psi) + J(\Psi) + K(\Psi). \tag{4.2.6}$$

The first term is the kinetic energy of the non-interacting electrons,

$$T(\Psi) = 2 \sum_{i=1}^{n} T_i = -\sum_{i=1}^{n} \int \phi_i^*(\vec{x}) \nabla^2 \phi_i(\vec{x}) d\vec{x}. \tag{4.2.7}$$

The second term is the electronic energy between electrons and the external field $V_{ext}(\vec{x})$,

$$V(\Psi) = 2\sum_{i=1}^{n} V_i = 2\sum_{i=1}^{n} \int \phi_i^*(\vec{x}) V_{ext}(\vec{x}) \phi_i(\vec{x}) d\vec{x}. \tag{4.2.8}$$

The third term is the electrostatic interaction energy among electrons,

$$J(\Psi) = 2\sum_{i,j=1}^{n} J_{ij} = 2\sum_{i,j=1}^{n} \int\int \frac{\phi_i^*(\vec{x})\phi_j^*(\vec{x}')\phi_i(\vec{x})\phi_j(\vec{x}')}{|\vec{x}-\vec{x}'|} d\vec{x} d\vec{x}'. \tag{4.2.9}$$

The last term is the exact exchange energy of electrons,

$$K(\Psi) = -\sum_{i,j=1}^{n} K_{ij} = -\sum_{i,j=1}^{n} \int\int \frac{\phi_i^*(\vec{x})\phi_j^*(\vec{x}')\phi_j(\vec{x})\phi_i(\vec{x}')}{|\vec{x}-\vec{x}'|} d\vec{x} d\vec{x}'. \tag{4.2.10}$$

By the method of the Lagrange multiplier, we have the diagonal effective one-electron eigenvalue equation [129, 71]:

$$\{\widehat{T}_i + \widehat{V}_i + \sum_{j=1}^{n} (2\widehat{J}_j - \widehat{K}_j)\}\phi_i = \epsilon_i \phi_i, i = 1, \ldots, n, \tag{4.2.11}$$

where ϵ_j Lagrange multipliers, and

$$\begin{aligned}
&\widehat{T}_i = -\frac{1}{2}\nabla^2, \\
&\widehat{V}_i = V_{ext}(\vec{x}), \\
&\widehat{J}_j\phi_i(\vec{x}) = \int \frac{\phi_j^*(\vec{x}')\phi_j(\vec{x}')}{|\vec{x}-\vec{x}'|} d\vec{x}' \phi_i(\vec{x}), \\
&\widehat{K_j}\phi_i(\vec{x}) = \int \frac{\phi_j^*(\vec{x}')\phi_i(\vec{x}')}{|\vec{x}-\vec{x}'|} d\vec{x}' \phi_j(\vec{x}).
\end{aligned} \tag{4.2.12}$$

Self-consistent iterative method is conducted to solve effective one-electron eigenvalue equation (4.2.11), and the ground state energy is achieved when the self-consistent solver gets convergent.

4.2.2 Equivalent Hartree-Fock Equations

Effective one-electron eigenvalue equation (4.2.11) could be rewritten as [55, 164, 71]

$$\{\widehat{T}_i - \epsilon_i\}\phi_i = \{-\widehat{V}_i - \sum_{j=1}^{n} (2\widehat{J}_j + \widehat{K}_j)\}\phi_i, i = 1, \ldots, n, \tag{4.2.13}$$

where we move the exact exchange operator $\widehat{K_j}$ to the right side of the effective one-electron equation, and we will treat this term as known in each step of Self-consistent solver, which will save the computational cost greatly.

Effective one-electron equation (4.2.13) will be the core equation in our linear scaling method, and the time cost of solving (4.2.13) will be discussed carefully in Section 4.3.

4.2.3 The Finite Element Discretization

Suppose m is the number of basis functions, and $\{\chi_\alpha\}_{\alpha=1}^m$ are the finite element basis functions. Then each molecular orbital could be written as [71]

$$\phi_i = \sum_{\alpha=1}^{m} c_{i\alpha}\chi_\alpha, i = 1, \ldots, n. \tag{4.2.14}$$

The finite element discretization of (4.2.13) is [71]

$$H_i c_i = v_i. \tag{4.2.15}$$

$c_i = \{c_{i1}, \ldots, c_{im}\}$ are the coefficients of molecular orbital in the expansions of finite element basis. H_i is the stiffness matrix of operator $(\widehat{T_i} - \epsilon_i)$, and the element of the stiffness matrix H_i is:

$$(H_i)_{\alpha,\beta} = \int_\Omega \left\{ \frac{1}{2}\nabla\chi_\alpha\nabla\chi_\beta - \epsilon_i\chi_\alpha\chi_\beta \right\} d\Omega, \alpha, \beta = 1, \ldots, n. \tag{4.2.16}$$

v_i is a constant vector, which could be calculated from the results of the Self-consistent solver in last step. The element of v_i is

$$(v_i)_\alpha = \int_\Omega \chi_\alpha \left\{ -\widehat{V}_i\phi_i - \sum_{j=1}^{n}(2\widehat{J}_j + \widehat{K_j})\phi_i \right\} d\Omega, \alpha = 1, \ldots, n. \tag{4.2.17}$$

The Helmholtz equation (4.2.15) could be solved in $O(m)$, and it will be analyzed in Section 4.3.

4.3 Algorithm

The algorithm [71] presented below is for restricted Hartree-Fock. Unrestricted Hartree-Fock models could be solved by the same process, except that we need to consider the spin-up and spin-down electrons separately.

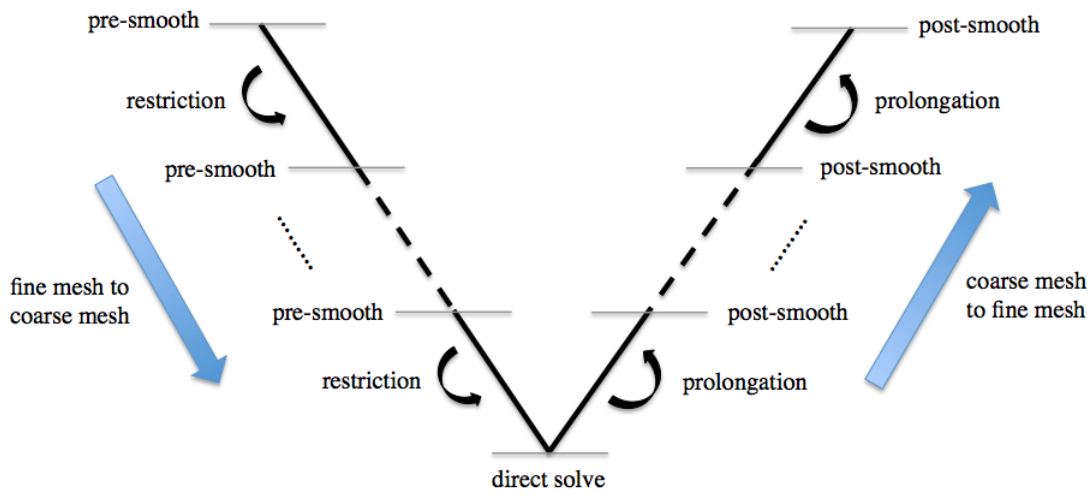

Figure 4.1: The V-cycle Multigrid Method

We introduce multigrid method in Section 4.2.1. The linear time cost analysis of solving equivalent Hartree-Fock equations is shown in Section 4.2.2, and the linear time cost of exact Fock matrix diagonalization step and energy correction step are discussed in Section 4.2.3. Section 4.2.4 shows the Self-consistent iteration process of our method.

4.3.1 Multigrid Method

Elliptic boundary value problems could be solved by multigrid method [5, 18, 54, 17, 21, 97]. The computational work of multigrid method is $O(m)$, where m is the number of bases. The approximate solution achieved by full multigrid method have comparable error bounds to theoretical error bounds in the finite element method. There are two main important steps of the multigrid method. One step is the smoothing step, which will reduce the high frequency part of the error. Then the smooth part of the residual error in the smoothing step will be corrected on the coarse mesh by the restriction step. Figure 4.1 [71] illustrates the V-cycle multigrid method. The V-cycle multigrid method only directly solves the problem in the most coarse mesh.

Consider the nested sequence of finite-element spaces $V_1 \subset V_2 \subset \cdots \subset V_n$. The prolongation operator I_k^{k-1} defines the mapping from V_{k-1} to V_k, and the restriction operator I_{k-1}^k defines the mapping from V_k to V_{k-1}. We denote by R_k the smoothing

operator. For the linear problem $Au = f$, an approximate coarse linear problem $A_k u_k = f_k$ needs to be constructed. Suppose $MG(k, u, f)$ is the approximate solution of the equation $A_k u = f$. Algorithm 1 [71] introduces the multigrid algorithm to solve the linear problem $Au = f$.

Algorithm 1 $MG(k, u, f)$

$MG(k, u, f)$ is the approximate solution of the equation $A_k u = f$.

Input $u_{k,0}$ is the initial guess for k step, TOL is the threshold for the smoothing step. m_1 and m_2 are the number of presmoothing steps and postsmoothing steps to reach the threshold TOL;

if $k = 1$ **then**

Direct solve: $u_1 = A_1^{-1} f_1$;

else

Presmoothing: $u_{k,i} = u_{k,i-1} + R_k^T(f_k - A_k u_{k,i-1})$, for $i = 1, \ldots, m_1$;

Error Correction: $u_{k,m_1+2} = u_{k,m_1+1} + I_{k-1}^k MG(k-1, 0, I_k^{k-1}[f_k - A_k u_{k,m_1+1}])$;

Postsmoothing: $u_{k,i} = u_{k,i} + R_k(f_k - A_k u_{k,i})$, for $i = m_1 + 2, \ldots, m_1 + m_2 + 1$;

Output $u_k = u_{k,m_1+m_2+1}$ is the solution of the kth level iteration.

Algebraic multigrid [20, 19, 97, 30] was introduced to solve linear systems based on multigrid principles, but only depends on the coefficients in the underlying matrix and needs no geometry knowledge of the problem. The matrix elements is the sole factor to determine coarse mesh, restriction and prolongation operators, and coarse mesh equations. Algebraic multigrid is a more general approach to algebraically enforce the variational conditions the variational conditions $A_{k-1} = I_k^{k-1} A_k I_{k-1}^k$ and $I_k^{k-1} = (I_{k-1}^k)^T$.

Algebraic multigrid can be used as a preconditioner, and combined with acceleration methods such as conjugate gradient method. It is more robust and efficient to use algebraic multigrid as a preconditioner, comparing to use it as a stand-alone solver. Algebraic multigrid method is used in our linear scaling finite element solver. The Helmholtz equation (4.2.15) in our Self-consistent solver are solved by conjugate gradient method with algebraic multigrid preconditioner. Also the Coulomb potentials $\widehat{J_j}$ and exchange potentials $\widehat{K_j}$ could be calculated by conjugate gradient method with algebraic multigrid preconditioner.

4.3.2 Equivalent Hartree-Fock Equations

In the Self-consistent solver, (4.2.13) could be rewritten as [71]

$$\begin{aligned} & (-\frac{1}{2}\nabla^2 - \epsilon_i^{(k-1)})\phi_i^{(k)}(\vec{x}) \\ = \quad & \sum_{j=1}^{n}\left[2V_{jj}^{(k-1)}(\vec{x})\phi_i^{(k-1)}(\vec{x}) - V_{ij}^{(k-1)}(\vec{x})\phi_j^{(k-1)}(\vec{x})\right] + V_{ext}(\vec{x})\phi_i^{(k-1)}(\vec{x}) \\ = \quad & f_i^{k-1}(\vec{x}), \end{aligned} \tag{4.3.1}$$

where $\{V_{ij}^{(k)}(\vec{x})\}_{i,j=1}^{n}$ could be calculated by solving Poisson equations,

$$-\frac{1}{4\pi}\nabla^2 V_{ij}^{(k)}(\vec{x}) = \phi_i^{*(k-1)}(\vec{x})\phi_j^{(k-1)}(\vec{x}), \tag{4.3.2}$$

where k, $(k-1)$ mean the kth, $(k-1)$th step of our self-consistent solver.

Because the bounded states energies $\{\epsilon_i\}_{i=1}^{n}$ are always negative and $-\nabla^2$ are positive definite, these make the operator $-\frac{1}{2}\nabla^2 - \epsilon_i^{(k-1)}$ is always positive definite. In our linear scaling finite element solver [71], (4.3.1) and (4.3.2) are solved by conjugate gradient method with algebraic multigrid preconditioner in $O(m)$.

4.3.3 Exact Fock Matrix Diagonalization and Energy Correction

We only update molecular orbitals in (4.3.1) and (4.3.2). In order to update energies, we need to include exact Fock matrix diagonalization step and energy correction step in the self-consistent solver. Harrison's paper [55] mentioned exact Fock matrix diagonalization step, and we will give a clear description of exact Fock matrix diagonalization step. The formulations of energy correction step followed Harrison's paper [55].

Exact Fock Matrix Diagonalization

Exact Fock matrix diagonalization is to diagonalize the Fock matrix in the space of occupied orbitals. Suppose updated molecular orbitals from last step is $\{\phi_i^{(k-1)}\}_{i=1}^{n}$,

and the Fock matrix $(F^{(k-1)})_{n*n}$ element is given [71]

$$\begin{aligned} F_{ij}^{(k-1)} &= \left(\phi_i^{(k-1)}(\vec{x}), -\frac{1}{2}\nabla^2\phi_j^{(k-1)}(\vec{x}) + V_{ext}(\vec{x})\phi_j^{(k-1)}(\vec{x})\right. \\ &\quad \sum_{\mu=1}^{m}\left[2\int \frac{\phi_\mu^{*(k-1)}(\vec{x}')\phi_\mu^{(k-1)}(\vec{x}')}{|\vec{x}-\vec{x}'|}d\vec{x}'\phi_j^{(k-1)}(\vec{x})\right. \\ &\quad \left.\left. -\int \frac{\phi_\mu^{*(k-1)}(\vec{x}')\phi_j^{(k-1)}(\vec{x}')}{|\vec{x}-\vec{x}'|}d\vec{x}'\phi_\mu^{(k-1)}(\vec{x})\right]\right), \\ &= \left(\phi_i^{(k-1)}(\vec{x}), -\frac{1}{2}\triangle^2\phi_j^{(k-1)}(\vec{x}) + \sum_{\mu=1}^{m}\left[2V_{\mu\mu}^{(k-1)}(\vec{x})\phi_j^{(k-1)}(\vec{x})\right.\right. \\ &\quad \left.\left. -V_{j\mu}^{(k-1)}(\vec{x})\phi_\mu^{(k-1)}(\vec{x}) + V_{ext}(\vec{x})\phi_j^{(k-1)}(\vec{x})\right]\right), \end{aligned} \tag{4.3.3}$$

where

$$-\frac{1}{4\pi}\nabla^2 V_{j\mu}^{(k-1)}(\vec{x}) = \phi_\mu^{*(k-1)}(\vec{x})\phi_j^{(k-1)}(\vec{x}). \tag{4.3.4}$$

All the potentials $\{V_{j\mu}^{(k-1)}(\vec{x})\}_{j,\mu=1}^{n}$ are calculated by conjugate gradient method with algebraic multigrid preconditioner in $O(m)$. V_{ext} is the nuclear potential. As a result of $m \ll n$, the time cost to get n^2 elements of Fock matrix is still $O(m)$.

After the diagonalization process of Fock matrix F, we get updated molecular orbitals $\{\phi_i^{(k)}\}_{i=1}^{n}$ and their corresponding energies $\epsilon_i^{(k)}$, which satisfy

$$\phi_i^{(k)}(\vec{x}) = \sum_{j=1}^{n} c_{ij}^{(k-1)}\phi_j^{(k-1)}(\vec{x}), \tag{4.3.5}$$

$$\sum_{\mu=1}^{n} c_{i\mu}^{(k-1)}c_{j\mu}^{(k-1)} = \delta_{ij}. \tag{4.3.6}$$

By the orthonormality of the occupied orbitals $\{\phi_i^{(k-1)}\}_{i=1}^{n}$, we have [71]

$$\begin{aligned} \int \phi_i^{(k)}(\vec{x})\phi_j^{(k)}(\vec{x})d\vec{x} &= \sum_{\nu,\mu=1}^{n} c_{i\nu}^{(k-1)}c_{j\mu}^{(k-1)}\int \phi_\nu^{(k-1)}(\vec{x})\phi_\mu^{(k-1)}(\vec{x})d\vec{x} \\ &= \sum_{\mu=1}^{n} c_{i\mu}^{(k-1)}c_{j\mu}^{(k-1)} = \delta_{ij}, \end{aligned} \tag{4.3.7}$$

which means $\{\phi_i^{(k)}\}_{i=1}^{n}$ is orthonormalized.

Energy Correction

Suppose the error between precise energy ϵ_i and calculated energy $\epsilon_i^{(k-1)}$ in $(k-1)$th step is $\delta\epsilon_i^{(k-1)}$. By (4.3.1), we have [55, 71]

$$\phi_i^{(k)}(\vec{x}) = -(-\frac{1}{2}\nabla^2 - \epsilon_i^{(k-1)} - \delta\epsilon_i^{(k-1)})^{-1} f_i^{(k-1)}(\vec{x}). \tag{4.3.8}$$

We expand the operator in a Taylor series and obtain [55, 71]

$$\begin{aligned}\phi_i^{(k)}(\vec{x}) = \quad & -(-\frac{1}{2}\nabla^2 - \epsilon_i^{(k-1)})^{-1} f_i^{(k-1)}(\vec{x}) - \delta\epsilon_i^{(k-1)}(-\frac{1}{2}\nabla^2 - \epsilon_i^{(k-1)})^{-2} f_i^{(k-1)}(\vec{x}) \\ & + O((\delta\epsilon_i^{(k-1)})^2).\end{aligned} \tag{4.3.9}$$

Left projection with $V\psi$ and rearrangement yield the following update for the energy:

$$\delta\epsilon_i^{(k-1)} = -\frac{\langle f_i^{(k-1)}(\vec{x}) | \phi_i^{(k-1)}(\vec{x}) - \widetilde{\phi}_i^{(k-1)}(\vec{x})\rangle}{\|\widetilde{\phi}_i^{(k-1)}(\vec{x})\|^2}, \tag{4.3.10}$$

where

$$\widetilde{\phi}_i^{(k-1)}(\vec{x}) = -(-\frac{1}{2}\nabla^2 - \epsilon_i^{(k-1)})^{-1} f_i^{(k-1)}(\vec{x}). \tag{4.3.11}$$

The time cost to calculate (4.3.10) and (4.3.11) is obviously $O(m)$. The updated process is [55, 71]

$$\begin{aligned}\psi^k &= \widetilde{\psi}^{(k-1)}, \\ \epsilon^{(k)} &= \epsilon^{(k-1)} + \delta\epsilon^{(k-1)}.\end{aligned} \tag{4.3.12}$$

4.3.4 Self-consistent Iteration

The time cost analysis in Section 4.3.2 and Section 4.3.3 shows the linear scaling $O(m)$ time cost of our algorithm. Because we never need to form a dense stiffness matrix for exact exchange operator, the memory cost is also linear [71]. During the numerical implementation, we use energy correction step in self-consistent solver to update the molecular orbitals for small systems, and we use exact Fock matrix diagonalization step in self-consistent solver to update the molecular orbitals for large systems [71]. In order to get the self-consistent solver convergent for large systems, we updated

molecular orbital energies only when the small difference is reached in two consecutive self-consistent iterations [71]. Algorithm 2 [71] shows the process of the self-consistent solver. An initial guess $(\phi_i^{(0)}, \epsilon_i^{(0)}), i = 1, \ldots, n$ is given to start the self-consistent solver. The self-consistent solver will stop when the total energy difference in two consecutive iterations is smaller than the tolerance TOL.

Algorithm 2 The Self-consistent Iteration

Input $(\phi_i^{(0)}, \epsilon_i^{(0)}), i = 1, \ldots, n$, TOL;

while $\|\epsilon_{total}^{(k)} - \epsilon_{total}^{(k-1)}\| >$TOL **do**

Evaluate potentials $V_{ij}^{(k)}, i, j = 1, \ldots, n$ in (4.3.2);

Evaluate $f_i^{(k)}(\vec{x}), i = 1, \ldots, n$ in (4.3.1);

Solve Helmholtz equation, and get updated $\{\phi_i^{(k+1)}, i = 1, \ldots, n\}$;

Exact Fock matrix diagonalization step/energy correction step, and get updated $\{\epsilon_i^{(k+1)}, i = 1, \ldots, n\}$;

k++;

Output $(\phi_i, \epsilon_i), i = 1, \ldots, n$.

4.4 Numerical Results

Our numerical implementations [71] are built based on Python FEniCS [34] package, which is a collection of free software with an extensive list of features for automated, efficient solution of differential equations. The source code of our linear scaling finite element solver [71] can be found [69]. Our calculation domain is a box, and we try to maintain precision independent of the size of the box. Because molecular orbitals decay exponentially from the position of the nucleus, we apply zero boundary condition to the molecular orbitals. The initial guess for the molecular orbitals are STO-3G molecular orbitals generated with NWChem [154, 77]. The singularities of nuclei potential are avoided by adding a small positive constant in the denominator.

We show in Section 4.4.1 the priori adaptive mesh constructed for calculation. Section 4.4.2 shows Hartree-Fock numerical results of He and Be atom. Section 4.4.3 shows Hartree-Fock numerical results of H_2, C_2 and BeF molecules. H_2 molecule

singlet/tripet are discussed in Section 4.4.4. Section 4.4.5 gives a simple example of numerical implementation of our linear scaling method on density functional theory models.

4.4.1 Priori Adaptive Mesh

We first construct a coarse triangulation mesh, with the nodes positioned at a high coarsening rate in the radial direction away from the nuclei. Then we increase the triangulations by successive uniform subdivisons. Figure 4.4 [71] is an example of priori adptive mesh for a single atom. In a multi-atom system, finite element method is flexible to give different amount of mesh to different atoms according to their nuclei potential decay rates. Figure 4.3 [71] is an example to show the mesh distribution for a two-atom system, which gives more mesh to the left nucleus than the right nucleus.

4.4.2 Hartree-Fock: Be, He

We use Be atom as an example to demonstrate the properties of our solver. Figure 4.4 [71] shows the convergence of self consistent solver for Be atom. Figure 4.5 [71] shows the convergence as a function of the number of subdivisions for Be atom . In order to verify convergence with respect to mesh size, we first construct a coarse priori adaptive triangulation mesh, and repeat the calculations on increasingly finer triangulations obtained by successive uniform subdivisons. Figure 4.6 [71] shows the computational cost scales linearly with the number of bases, which is consistent with the analysis in Section 4.3.2 and 4.3.3.

The calculated ground state energy of our Hartree-Fock solver performed upon the neural atoms He, Be are listed in Table 4.1 [71]. The accuracies of the total energies could reach $0.002\ a.u.$ for He atom and $0.02\ a.u.$ for Be atom.

4.4.3 Hartree-Fock: H_2, C_2, BeF

Ground state energy Hartree-Fock calculations are performed upon several molecules H_2 , C_2 and BeF are listed in Table 4.2 [71]. The bond length we used for H_2 , C_2 and BeF are $1.4\ a.u.$, $2.358\ a.u.$, and $2.386\ a.u.$ [55]. The accuracies of the

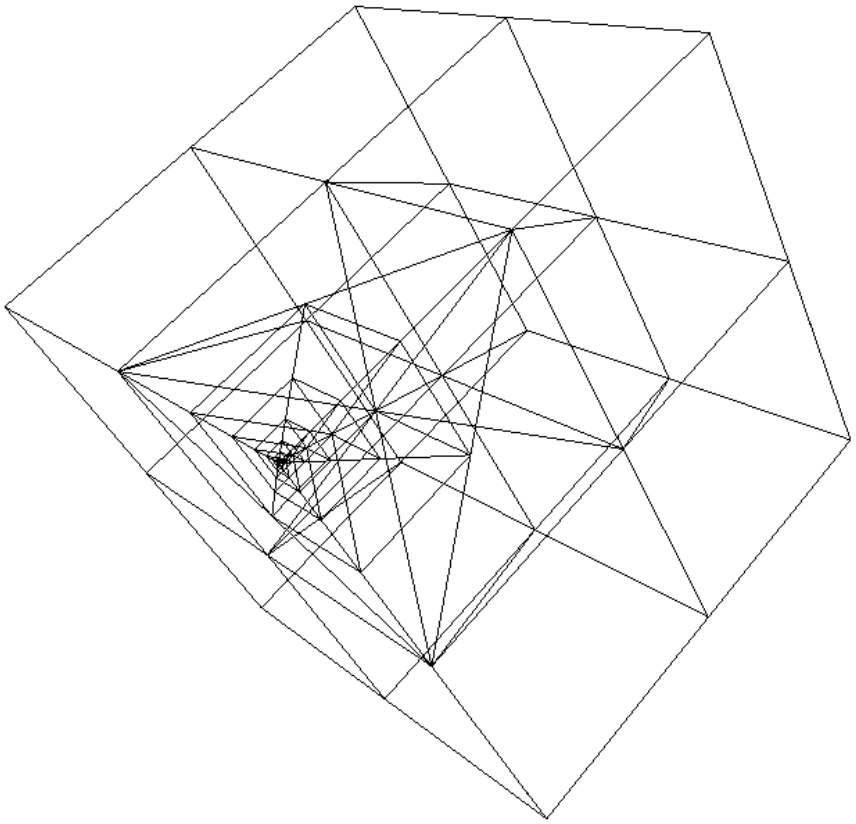

Figure 4.2: Priori Adaptive Mesh

Table 4.1: Hartree-Fock Ground State Energies (a.u.) for Be, He

Atom	Calculated total energy	Reference [80]
He	-2.859	-2.861
Be	-14.553	-14.573

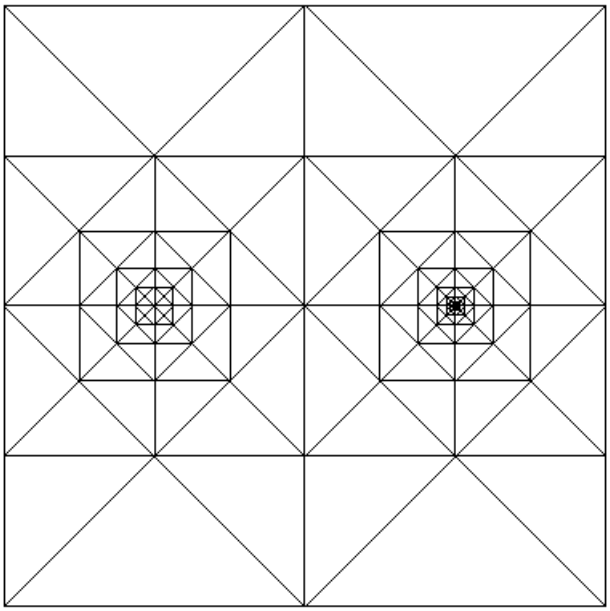

Figure 4.3: Priori Adaptive Mesh on a Two-atom System

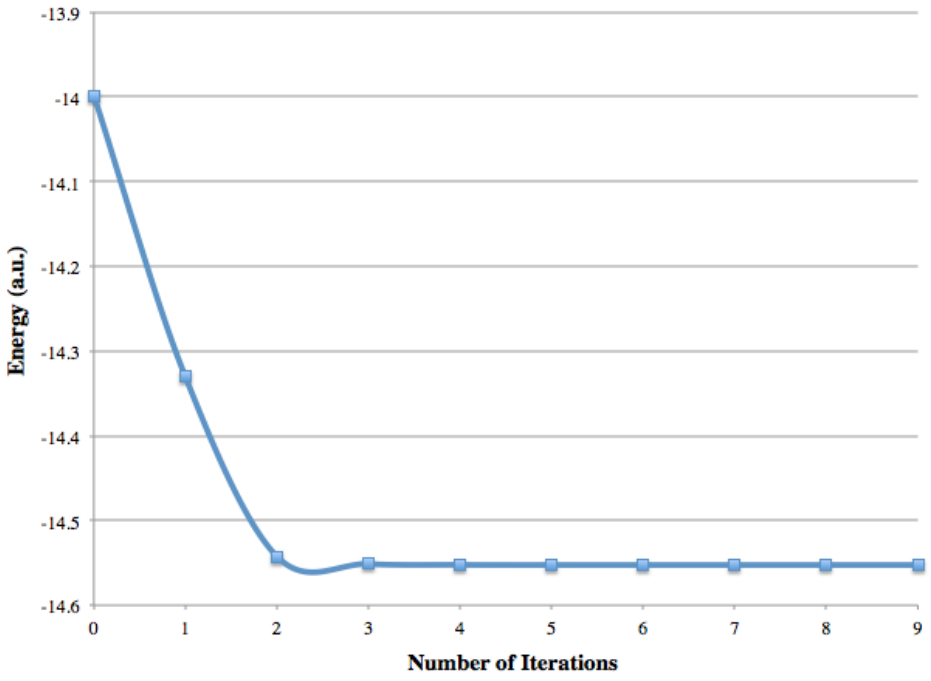

Figure 4.4: Convergence of Self Consistent Solver (Hartree-Fock: Be atom)

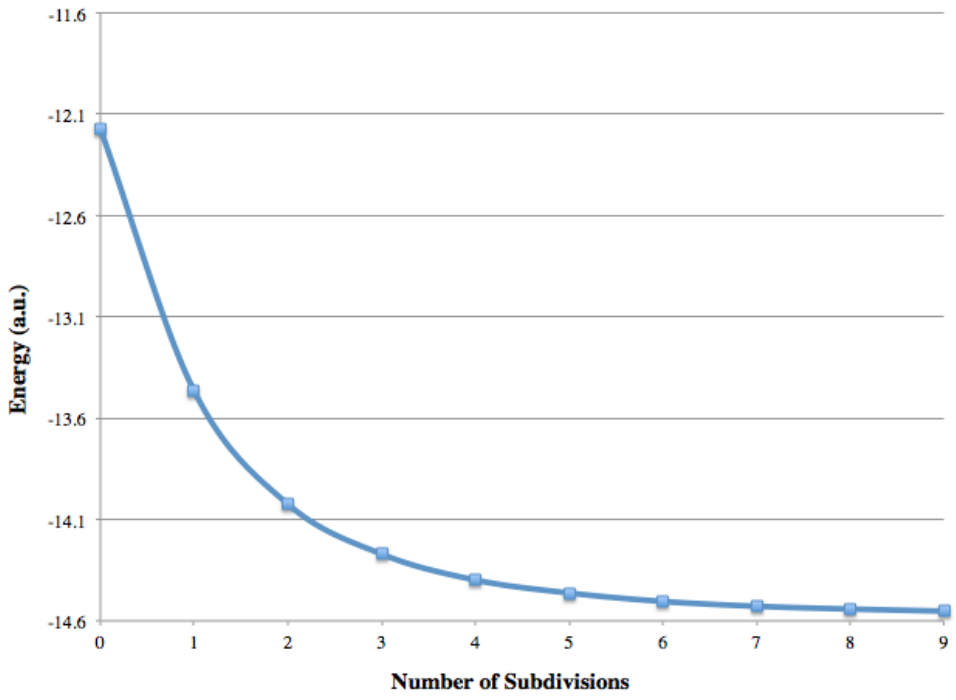

Figure 4.5: Convergence with respect to Mesh Size (Hartree-Fock: Be atom)

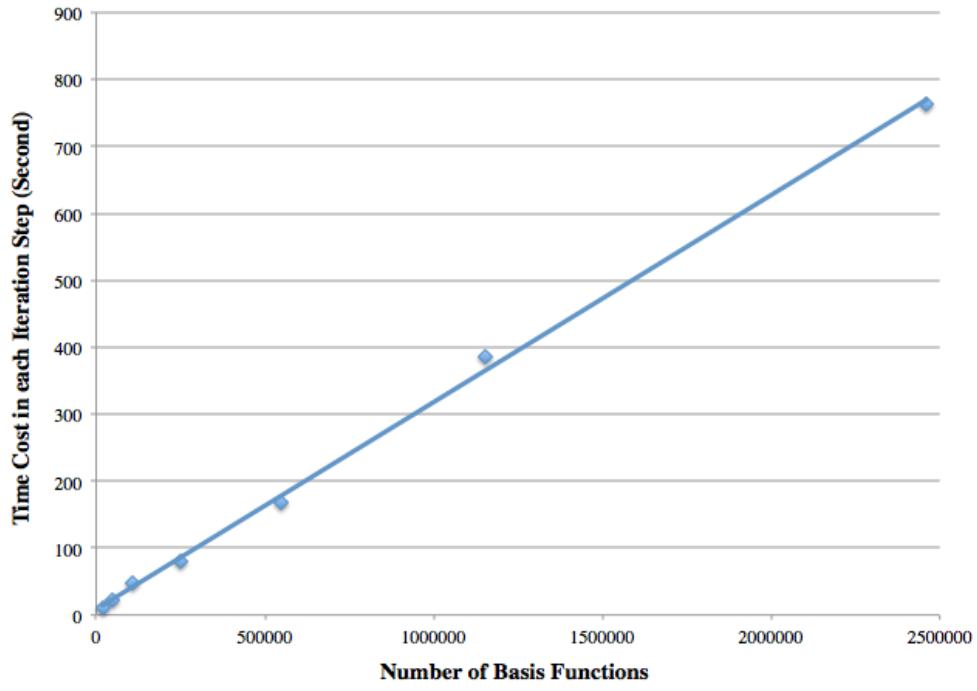

Figure 4.6: Linear Scaling Method (Hartree-Fock: Be atom)

Table 4.2: Hartree-Fock Ground State Energies (a.u.) for H_2, C_2, BeF

Molecule	Calculated total energy	Reference
H_2	-1.132	-1.134 [101]
C_2	-75.406	-75.407 [109]
BeF	-124.06	-124.17 [101]

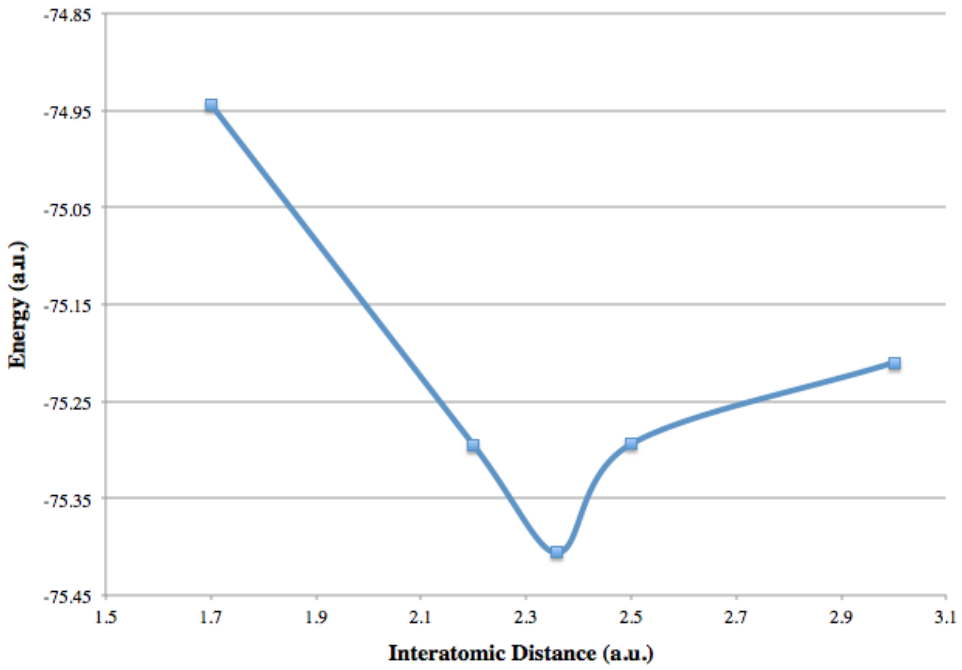

Figure 4.7: Hartree-Fock Ground State Energy of C_2 as a Function of Interatomic Distance

ground state energies could reach $0.002\ a.u.$ for H_2, $0.001\ a.u.$ for C_2 and $0.01\ a.u.$ for BeF. We also studied the ground state energies of C_2 at various interatomic distances. Figure 4.7 [71] shows our solver has a smooth energy curve, and our calculated bond length for C_2 is consistent with the reference [55].

4.4.4 Hartree-Fock: H_2 Singlet/Triplet

H_2 molecular triplet has exact exchange energy, thus it is a great example to apply our linear scaling method on the H_2 molecular triplet study. The interatomic dis-

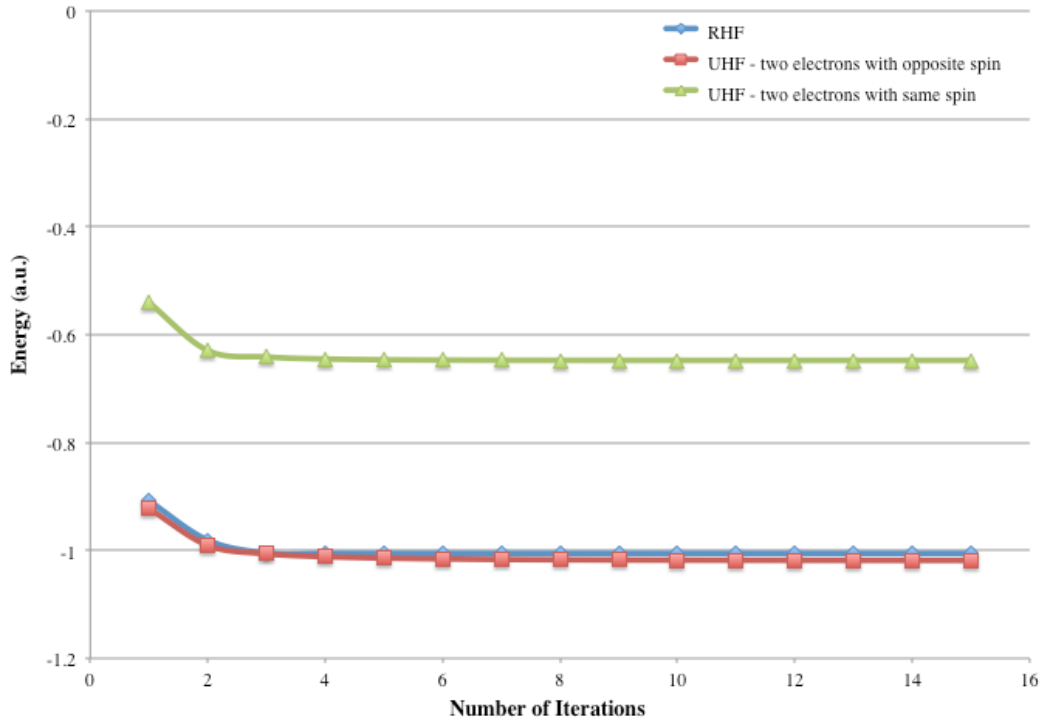

Figure 4.8: Convergence of Self Consistent Solver (H_2, bond length=$2.8\ a.u.$)

tance in this study is chosen as $2.8\ a.u.$, which is twice larger than the real H_2 molecular bond length. The large bond length will give different solutions for unrestricted Hartree-Fock singlet and restricted Hartree-Fock models single, which is a very interesting phenomenon and will be discussed very carefully in Chapter 5 [72]. Figure 4.8 [71] shows the convergence of self consistent solver for singlet restricted Hartree-Fock, singlet unrestricted Hartree-Fock (two electrons with same spin) and triplet unrestricted Hartree-Fock (two electrons with opposite spins) for H_2 with bond length $2.8\ a.u.$. The results in Figure 4.8 [71] is reasonable. Triplet unrestricted Hartree-Fock has the highest ground state energy, because two electrons have same spin and one electron has to occupy a relative high energy level molecular spinorbital. Singlet unrestricted Hartree-Fock (two electrons with same spin) has lower energy than singlet restricted Hartree-Fock, because the localized unrestricted solution is only local critical point.

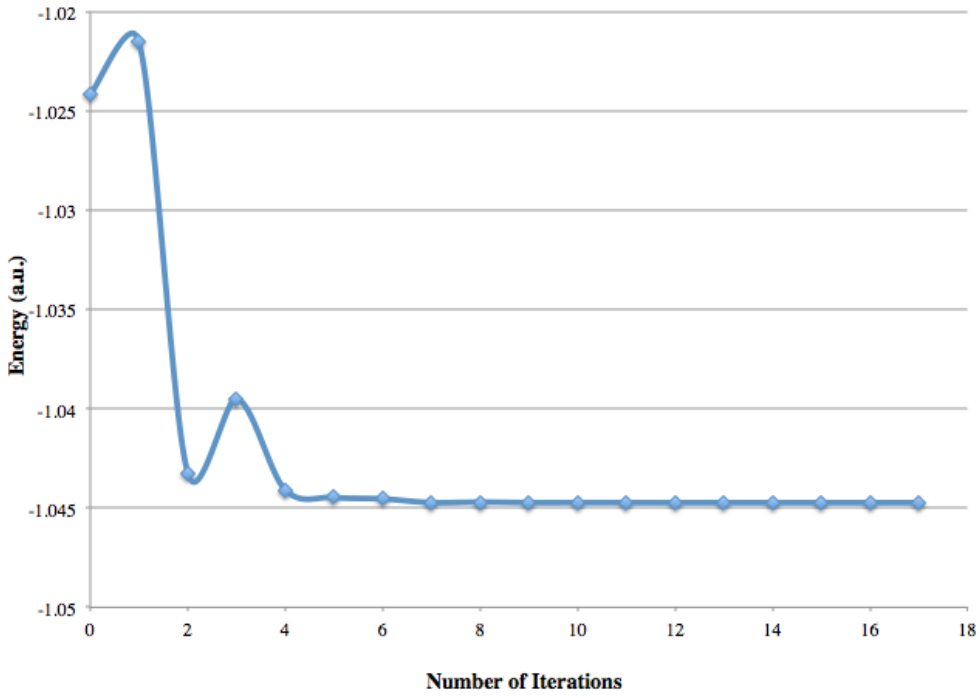

Figure 4.9: Convergence of LSDA Self Consistent Solver (H_2, bond length=$1.4\ a.u.$)

4.4.5 Density Functional Theory: H_2

Local spin density approximation has been implemented numerically with our linear scaling method. Figure 4.9 [71] shows a simple example of H_2. I didn't include the correlation energy term in the numerical implementation. The correlation energy is a small term and will not influence the convergence result of our solver. Figure 4.9 [71] examines the convergence behavior of our method. This shows our linear scaling method could also be applied to density functional theory models and hybrid models..

4.5 Conclusion

We successfully proposed and numerically implemented a linear time cost and memory cost method [71] with finite element basis, which could handle exact exchange operator. Our package is built based on Python FEniCS [34] package. The algebraic multigrid method is used in our algorithm in order to achieve linear scaling. The priori adaptive mesh is introduced to improve our solver. Several numerical experiments of

atoms and molecules have been conducted using our solver, and pretty good accuracies could be reached. The development of our method makes finite element basis become competitive in models with exact exchange energy. Our method could be generalized to density functional theory and hybrid models in a straightforward way. Each step in our self-consistent solver is well defined, and this makes the parallelization of our solver is feasible.

Chapter 4, is currently begin prepared for submission for publication of the material. Hu, H., Song, D., Weare, J., and Holst, M. The dissertation author was the primary investigator and author of this material.

Chapter 5

Stability Analysis of the Electronic Structure Models

This chapter introduces a systematic way to study the stability condition of Hartree-Fock and density functional theory models [72]. Section 5.1 shows the motivation to develop the method. Section 5.2 introduces the major theorem to analyze the stability condition. Section 5.3 shows the Hessian matrix derivations for Hartree-Fock and density functional theory models. Section 5.4 gives some details of numerical implementations. Section 5.5 shows some numerical results. Section 5.6 is our conclusion.

5.1 Motivation

Consider the zero set $\{(a, x) : \ f(a, x) = 0\}$. For a given parameter value a, if there are more than one solution in (a, x)-space, we call it bifurcation. Figure 5.1 gives an example of a bifurcation curve. If $\nabla_{a,x} f(a, x) \neq 0$, the Implicit Function Theorem guarantees that the system can be written as a smooth function of a. Otherwise, bifurcations can happen.

In Hartree-Fock and density functional theory solvers, molecular orbitals and orbital wavefunctions are solved by the Euler-Lagrange equation of the total energy functional. The Euler-Lagrange equation can be considered as $f(a, x)$ discussed above, where a can be interatomic distance R or nuclear charge Z, and x are molecular orbitals in Hartree-Fock and orbital wavefunctions in Density Functional Theory. We will show

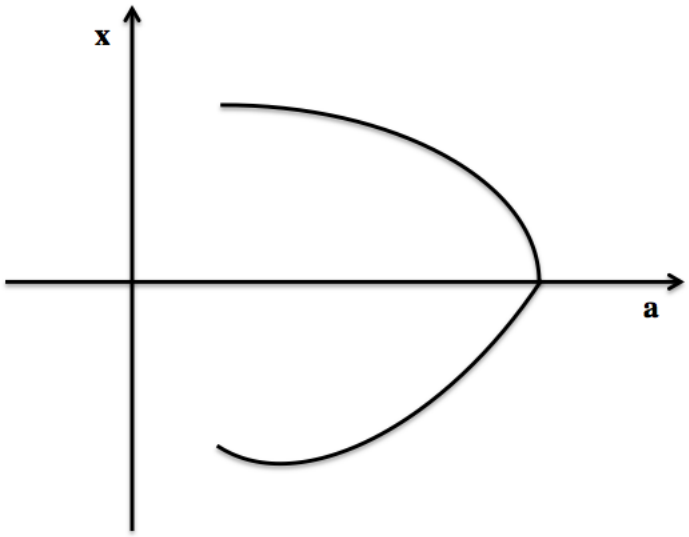

Figure 5.1: Bifurcation Curve

a few bifurcation figures below.

The solutions of the Euler-Lagrange equations are extremums. In order to establish whether that extremum is a maximizer, minimizer, or saddle point, shown in Figure 5.2, the second functional derivative must be analyzed. Few papers [74, 136, 110, 140] discussed the stability of Hartree-Fock and density functional theory results. In Chapter 6, we show a systematic way of stability study by Hessian matrix analysis [72].

5.2 Theorem

We will introduce the major theorem we use to analyze the stability condition in this section. Suppose the functional we want to minimize is $E(u)$, and let the constraints are

$$V = \{u : g_j(u) = 0, j = 1, \ldots, k\} \tag{5.2.1}$$

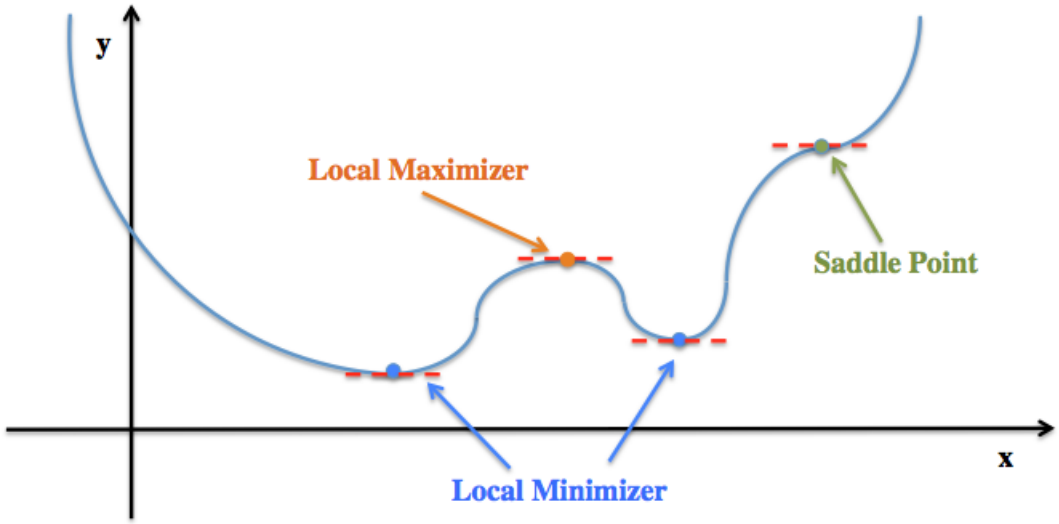

Figure 5.2: Minimum,Maximum, and Saddle Point

Let's define

$$L(u,\lambda) = E(u) + \sum_{j=1}^{m} \lambda_j g_j(u), \tag{5.2.2}$$

$$L'(u,\lambda) = E'(u) + \sum_{j=1}^{m} \lambda_j g_j'(u), \tag{5.2.3}$$

$$L''(u,\lambda) = E''(u) + \sum_{j=1}^{m} \lambda_j g_j''(u). \tag{5.2.4}$$

Then Theorem 5.2.1 holds.

Theorem 5.2.1 ([100, 155, 47]). *Suppose that* (u^*, λ^*) *satisfies* $L'(u^*, \lambda^*) = 0$*, and* $< L''(u^*, \lambda^*)h, h > > 0$ *for all* h *satisfying* $< g_j'(u^*), h >= 0$ *for every* $j = 1, \ldots, k$*. Then* (u^*, λ^*) *is a local minimizer of* E*.*

We call the vector space

$$V_{cons} = \{h :< g_j'(u^*), h >= 0, j = 1, \ldots, k\} \tag{5.2.5}$$

constraint manifold. If the eigenvalues λ_i on the constraint manifold are all positive, then u^* represents a local minimum. Correspondingly, if λ_i on the constraint manifold are all

negative, then u^* represents a local maximum. In the intermediate case of eigenvalues of mixed sign, we conclude that u^* corresponds to a saddle point of the functional.

Establishing whether a particular extremum is a global, rather than local, minimum or maximum is a more difficult problem in optimization theory that remains unsolved for arbitratry $E(u)$.

5.3 Formulation (Two-electrons' System – Hessian Matrix)

This section shows the Hessian matrix derivations for Hartree-Fock and density functional theory models. Section 5.3.1 and Section 5.3.2 show the formulations of restricted Hartree-Fock and unrestricted Hartree-Fock models. Section 5.3.3 and Section 5.3.4 show the formulations of restricted density functional theory and unrestricted density functional theory models.

People always choose molecular orbitals to be real-valued vectors, which means molecular orbital ϕ and its complex conjugate $\bar{\phi}$ are same. In our Hessian analysis, it will be equivalent if 1) we consider ϕ is a real-valued vector and the only variable; 2) ϕ and its complex conjugate $\bar{\phi}$ are two variables, but they need to satisfy $\phi = \bar{\phi}$. Thus, we assume molecular orbitals are real-valued vectors in our following analysis without any loss of generality.

5.3.1 Restricted Hartree-Fock

The restricted Hartree-Fock model for two-electrons' system has only one molecular orbital ϕ. Total energy functional is $E(\phi)$ is [72]

$$E(\phi) = \int |\nabla\phi(\vec{x})|^2 d\vec{x} + \int 2V_{ext}(\vec{x})|\phi(\vec{x})|^2 d\vec{x} + \int\int \frac{|\phi(\vec{x}')|^2|\phi(\vec{x})|^2}{|\vec{x}-\vec{x}'|} d\vec{x}' d\vec{x}, \quad (5.3.1)$$

where V_{ext} is the nuclear potential. Since we only study two electrons in our system, we don't have exchange potential term. We also have the constraint on ϕ

$$\int |\phi(\vec{x})^2| d\vec{x} = 1 \quad (5.3.2)$$

By Lagrange multiplier method, we have [72]

$$\begin{aligned} L(\phi,\epsilon) = \quad & \int |\nabla\phi(\vec{x})|^2 d\vec{x} + \int 2V_{ext}(\vec{x})|\phi(\vec{x})|^2 d\vec{x} \\ & + \int\int \frac{|\phi(\vec{x}')|^2|\phi(\vec{x})|^2}{|\vec{x}-\vec{x}'|} d\vec{x}' d\vec{x} - \epsilon(\int |\phi(\vec{x})^2| d\vec{x} - 1), \end{aligned} \tag{5.3.3}$$

where ϵ is the Lagrange multiplier.

In order to calculate the extremum of $L(\phi,\epsilon)$, the Euler-Lagrange equation of the total energy functional (5.3.3) with the constraint, which is the effective one-electron eigenvalue problem, is [72]

$$\frac{\delta L(\phi,\epsilon)}{\delta\phi} = F(\phi,\epsilon) = \left(-\frac{1}{2}\nabla^2 + V_{ext}(\vec{x}) + \int \frac{\phi(\vec{x}')^2}{|\vec{x}-\vec{x}'|} d\vec{x}'\right)\phi(\vec{x}) - \epsilon\phi(\vec{x}) = 0, \tag{5.3.4}$$

where ϕ needs to satisfy the constraint (5.3.2).

In order to establish whether that extremum is a maximizer, minimizer, or saddle point, the second order functional derivative (Hessian matrix) must be analyzed. Suppose the extremum (ϕ^*,ϵ^*) satisfy both (5.3.4) and (5.3.2). The Hessian analysis involves the solution of the eigenvalue problem [72]

$$\begin{aligned} \mathbf{H}w_i(\vec{x}') &= \int \frac{\delta L(\phi,\epsilon)}{\delta\phi(\vec{x})\delta\phi(\vec{x}')} w_i(\vec{x}') d\vec{x}' \Big|_{(\phi,\epsilon)=(\phi^*,\epsilon^*)} = \lambda_i w_i(\vec{x}) \\ &= DF(\phi,\epsilon)w_i = \frac{d}{dt}F(\phi + tw_i)|_{t=0} \\ &= -\frac{1}{2}\nabla^2 w_i + V_{ext}(\vec{x})w_i + \int \frac{2\phi^*(\vec{x}')w_i(\vec{x}')}{|\vec{x}-\vec{x}'|} d\vec{x}'\phi(\vec{x}) \\ &\quad + \int \frac{\phi^*(\vec{x}')^2}{|\vec{x}-\vec{x}'|} d\vec{x}' w_i(\vec{x}) - \epsilon^* w_i(\vec{x}), \end{aligned} \tag{5.3.5}$$

which includes the nonlocal exchange term. w_i needs to satisfy the constraint [72]

$$\int \phi^*(\vec{x}) w_i(\vec{x}) d\vec{x} = 0, \tag{5.3.6}$$

which means w_i are orthonormal to $\phi^*(\vec{x})$.

5.3.2 Unrestricted Hartree-Fock

Suppose (ϕ_1, ϕ_2) are the molecular orbitals for the unrestricted Hartree-Fock model. Total energy functional is $E(\phi)$ is [72]

$$\begin{aligned} E(\phi_1,\phi_2) = \quad & -\frac{1}{2}\int |\nabla\phi_1(\vec{x})|^2 d\vec{x} - \frac{1}{2}\int |\nabla\phi_2(\vec{x})|^2 d\vec{x} + \int V_{ext}(\vec{x})|\phi_1(\vec{x})|^2 d\vec{x} \\ & + \int V_{ext}(\vec{x})|\phi_2(\vec{x})|^2 d\vec{x} + \int\int \frac{|\phi_1(\vec{x}')^2||\phi_2(\vec{x})|^2}{|\vec{x}-\vec{x}'|} d\vec{x}' d\vec{x}, \end{aligned} \tag{5.3.7}$$

where V_{ext} is the nuclear potential. (ϕ_1, ϕ_2) also need to satisfy the constraints

$$\int |\phi_i(\vec{x})^2| d\vec{x} = 1, i = 1, 2. \tag{5.3.8}$$

By Lagrange multiplier method, we have [72]

$$\begin{aligned} L(\phi_1,\phi_2,\epsilon_1,\epsilon_2) = \quad & -\frac{1}{2}\int |\nabla\phi_1(\vec{x})|^2 d\vec{x} - \frac{1}{2}\int |\nabla\phi_2(\vec{x})|^2 d\vec{x} + \int V_{ext}(\vec{x})|\phi_1(\vec{x})|^2 d\vec{x} \\ & + \int V_{ext}(\vec{x})|\phi_2(\vec{x})|^2 d\vec{x} + \int\int \frac{|\phi_1(\vec{x}')^2||\phi_2(\vec{x})|^2}{|\vec{x}-\vec{x}'|} d\vec{x}' d\vec{x} \\ & -\epsilon_1\left(\int |\phi_1(\vec{x})^2| d\vec{x} - 1\right) - \epsilon_2\left(\int |\phi_2(\vec{x})^2| d\vec{x} - 1\right), \end{aligned} \tag{5.3.9}$$

where (ϵ_1, ϵ_2) are the Lagrange multipliers.

Suppose (ϵ_1, ϵ_2) are corresponding eigenvalues of molecular orbitals, and they are also the Lagrange multipliers. By the Euler-Lagrange equation of the total energy functional (5.3.9) with the constraints (5.3.8), we have the effective one-electron eigenvalue equations [72]

$$\begin{aligned} \left(-\frac{1}{2}\nabla^2 + V_{ext}(\vec{x}) + \int \frac{\phi_2(\vec{x}')^2}{|\vec{x}-\vec{x}'|} d\vec{x}'\right)\phi_1(\vec{x}) - \epsilon_1\phi_1(\vec{x}) = 0, \\ \left(-\frac{1}{2}\nabla^2 + V_{ext}(\vec{x}) + \int \frac{\phi_1(\vec{x}')^2}{|\vec{x}-\vec{x}'|} d\vec{x}'\right)\phi_2(\vec{x}) - \epsilon_2\phi_2(\vec{x}) = 0, \end{aligned} \tag{5.3.10}$$

where (ϕ_1, ϕ_2) need to satisfy the constraints (5.3.8).

Suppose the extremum (ϕ_1^*, ϕ_2^*) and $(\epsilon_1^*, \epsilon_2^*)$ satisfy both (5.3.8) and (5.3.10). We could get the Hessian matrix for unrestricted Hartree-Fock similarly as restricted Hartree-Fock [72],

$$\mathbf{H} = \begin{pmatrix} -\frac{1}{2}\nabla^2 + V_{ext} + \int \frac{\phi_2^*(\vec{x}')^2}{|\vec{x}-\vec{x}'|} d\vec{x}' - \epsilon_1^* & \int \frac{2\phi_2^*(\vec{x}')\cdot}{|\vec{x}-\vec{x}'|} d\vec{x}' \phi_1^*(\vec{x}) \\ \int \frac{2\phi_1^*(\vec{x}')\cdot}{|\vec{x}-\vec{x}'|} d\vec{x}' \phi_2^*(\vec{x}) & -\frac{1}{2}\nabla^2 + V_{ext} + \int \frac{\phi_1^*(\vec{x}')^2}{|\vec{x}-\vec{x}'|} d\vec{x}' - \epsilon_2^* \end{pmatrix}. \tag{5.3.11}$$

Suppose the finite element bases $\{\chi_\alpha\}_{\alpha=1}^m$ and the eigenvectors of $\mathbf{H}$ are $\{w_i\}_{i=1}^{2m}$. We could assume c_i is the coefficient of $w_i(\vec{x})$ expanded in the finite element bases. We can rewrite w_i as (w_{1i}, w_{2i}), where [72]

$$\begin{aligned} w_{1i} &= \sum_{j=1}^{m} c_{i,j}\chi_j, \\ w_{2i} &= \sum_{j=m+1}^{2m} c_{i,j}\chi_{j-m}, \end{aligned} \tag{5.3.12}$$

and they need to satisfy [72]

$$\begin{aligned} \int \phi_{1i}^*(\vec{x}) w_{1i}(\vec{x}) d\vec{x} = 0, \\ \int \phi_{2i}^*(\vec{x}) w_{2i}(\vec{x}) d\vec{x} = 0. \end{aligned} \tag{5.3.13}$$

5.3.3 Restricted Density Functional Theory

The restricted density functional theory for two-electrons' system has only one orbital wavefunction ϕ. The model we discussed below is local spin density approximation without correlation energy functional. Total energy functional is $E(\phi)$ is [72]

$$\begin{aligned} E(\phi) = \quad & \int |\nabla\phi(\vec{x})|^2 d\vec{x} + \int 2V_{ext}(\vec{x})|\phi(\vec{x})|^2 d\vec{x} + \int\int \frac{2|\phi(\vec{x}')|^2|\phi(\vec{x})|^2}{|\vec{x}-\vec{x}'|} d\vec{x}' d\vec{x}, \\ & -\frac{3}{2}\left(\frac{6}{\pi}\right)^{1/3} \int |\phi(\vec{x})|^{8/3} d\vec{x} \end{aligned} \tag{5.3.14}$$

where V_{ext} is the nuclear potential. The constraint on ϕ is

$$\int |\phi(\vec{x})^2| d\vec{x} = 1 \tag{5.3.15}$$

By Lagrange multiplier method, we have [72]

$$\begin{aligned} L(\phi, \epsilon) = \quad & \int |\nabla\phi(\vec{x})|^2 d\vec{x} + \int 2V_{ext}(\vec{x})|\phi(\vec{x})|^2 d\vec{x} + \int\int \frac{2|\phi(\vec{x}')|^2|\phi(\vec{x})|^2}{|\vec{x}-\vec{x}'|} d\vec{x}' d\vec{x} \\ & -\frac{3}{2}\left(\frac{6}{\pi}\right)^{1/3} \int |\phi(\vec{x})|^{8/3} d\vec{x} - \epsilon\left(\int |\phi(\vec{x})^2| d\vec{x} - 1\right), \end{aligned} \tag{5.3.16}$$

where ϵ is the Lagrange multiplier.

The Euler-Lagrange equation of the total energy functional (5.3.16) with the constraint, which is the effective one-electron eigenvalue problem, is [72]

$$\left(-\frac{1}{2}\nabla^2 + V_{ext}(\vec{x}) + \int \frac{2\phi(\vec{x}')^2}{|\vec{x}-\vec{x}'|}d\vec{x}' - \left(\frac{6}{\pi}\right)^{1/3}(\phi(\vec{x}))^{2/3}\right)\phi(\vec{x}) - \epsilon\phi(\vec{x}) = 0. \tag{5.3.17}$$

Suppose the extremum (ϕ^*, ϵ^*) satisfy both (5.3.17) and (5.3.15). The Hessian analysis, similar to the restricted Hartree-Fock analysis, involves the solution of the eigenvalue problem [72]

$$\begin{aligned}\mathbf{H}w_i(\vec{x}') &= -\frac{1}{2}\nabla^2 w_i + V_{ext}(\vec{x})w_i + \int \frac{2\phi^*(\vec{x}')w_i(\vec{x}')}{|\vec{x}-\vec{x}'|}d\vec{x}'\phi^*(\vec{x}) \\ &\quad + \int \frac{\phi^*(\vec{x}')^2}{|\vec{x}-\vec{x}'|}d\vec{x}'w_i(\vec{x}) - \frac{5}{3}\left(\frac{6}{\pi}\right)^{1/3}(\phi^*(\vec{x}))^{2/3}w_i(\vec{x}) - \epsilon^* w_i(\vec{x}),\end{aligned} \tag{5.3.18}$$

and w_i needs to satisfy the constraint [72]

$$\int \phi^*(\vec{x})w_i(\vec{x})d\vec{x} = 0, \tag{5.3.19}$$

which means w_i are orthonormal to $\phi^*(\vec{x})$.

5.3.4 Unrestricted Density Funtional Theory

The unrestricted density functional theory for two-electrons' system has two orbital wavefunctions (ϕ_1, ϕ_2). The model we discussed below is local spin density approximation without correlation energy functional. Total energy functional is $E(\phi)$ is [72]

$$\begin{aligned}E(\phi) = \quad & \frac{1}{2}\int |\nabla\phi_1(\vec{x})|^2 d\vec{x} + \frac{1}{2}\int |\nabla\phi_2(\vec{x})|^2 d\vec{x}, \\ & + \int V_{ext}(\vec{x})(|\phi_1(\vec{x})|^2 + |\phi_2(\vec{x})|^2)d\vec{x}, \\ & + \frac{1}{2}\int\int \frac{(|\phi_1(\vec{x})|^2 + |\phi_2(\vec{x})|^2)(|\phi_1(\vec{x}')|^2 + |\phi_2(\vec{x}')|^2)}{|\vec{x}-\vec{x}'|}d\vec{x}'d\vec{x}, \\ & - \frac{3}{4}\left(\frac{6}{\pi}\right)^{1/3}\int |\phi_1(\vec{x})|^{8/3}d\vec{x} - \frac{3}{4}\left(\frac{6}{\pi}\right)^{1/3}\int |\phi_2(\vec{x})|^{8/3}d\vec{x}\end{aligned} \tag{5.3.20}$$

where V_{ext} is the nuclear potential. The constraints on (ϕ_1, ϕ_2) are

$$\int |\phi_i(\vec{x})^2| d\vec{x} = 1, i = 1, 2 \tag{5.3.21}$$

By Lagrange multiplier method, we have [72]

$$\begin{aligned} L(\phi_1, \phi_2, \epsilon_1, \epsilon_2) = \quad & \frac{1}{2}\int (|\nabla\phi_1(\vec{x})|^2 + |\nabla\phi_2(\vec{x})|^2) d\vec{x} \\ & + \int V_{ext}(\vec{x})(|\phi_1(\vec{x})|^2 + |\phi_2(\vec{x})|^2) d\vec{x} \\ & + \frac{1}{2}\int\int \frac{(|\phi_1(\vec{x})|^2 + |\phi_2(\vec{x})|^2)(|\phi_1(\vec{x}')|^2 + |\phi_2(\vec{x}')|^2)}{|\vec{x} - \vec{x}'|} d\vec{x}' d\vec{x}, \\ & - \frac{3}{4}\left(\frac{6}{\pi}\right)^{1/3} \int |\phi_1(\vec{x})|^{8/3} d\vec{x} - \frac{3}{4}\left(\frac{6}{\pi}\right)^{1/3} \int |\phi_2(\vec{x})|^{8/3} d\vec{x}, \\ & - \epsilon_1 (\int |\phi_1(\vec{x})^2| d\vec{x} - 1) - \epsilon_2 (\int |\phi_2(\vec{x})^2| d\vec{x} - 1), \end{aligned} \tag{5.3.22}$$

, where (ϵ_1, ϵ_2) are the Lagrange multipliers.

The Euler-Lagrange equation of the total energy functional (5.3.22) with the constraint, which are the effective one-electron eigenvalue problems, are [72]

$$\begin{aligned} & \left(-\frac{1}{2}\nabla^2 + V_{ext}(\vec{x}) + \int \frac{(\phi_1(\vec{x}')^2 + \phi_2(\vec{x}')^2)}{|\vec{x} - \vec{x}'|} d\vec{x}' - \left(\frac{6}{\pi}\right)^{1/3} (\phi_i(\vec{x}))^{2/3} \right) \phi_i(\vec{x}) \\ = \quad & \epsilon_i \phi_i(\vec{x}), i = 1, 2, \end{aligned} \tag{5.3.23}$$

where (ϕ_1, ϕ_2) need to satisfy the constraint (5.3.21).

Suppose the extremum (ϕ_1^*, ϕ_2^*) and $(\epsilon_1^*, \epsilon_2^*)$ satisfy both (5.3.21) and (5.3.23). The Hessian analysis, similar to the unrestricted Hartree-Fock analysis, involves the solution of the eigenvalue problem [72]

$$\mathbf{H} = \begin{pmatrix} H_{11} & \int \frac{2\phi_2^*(\vec{x}')\cdot}{|\vec{x} - \vec{x}'|} d\vec{x}' \phi_1^*(\vec{x}) \\ \int \frac{2\phi_1^*(\vec{x}')\cdot}{|\vec{x} - \vec{x}'|} d\vec{x}' \phi_2^*(\vec{x}) & H_{22} \end{pmatrix}, \tag{5.3.24}$$

where

$$\begin{aligned} H_{11} &= -\frac{1}{2}\nabla^2 + V_{ext} + \int \frac{\phi_2^*(\vec{x}')^2}{|\vec{x} - \vec{x}'|} d\vec{x}' - \frac{5}{3}\left(\frac{6}{\pi}\right)^{1/3} (\phi_1^*(\vec{x}))^{2/3} - \epsilon_1^*, \\ H_{22} &= -\frac{1}{2}\nabla^2 + V_{ext} + \int \frac{\phi_1^*(\vec{x}')^2}{|\vec{x} - \vec{x}'|} d\vec{x}' - \frac{5}{3}\left(\frac{6}{\pi}\right)^{1/3} (\phi_2^*(\vec{x}))^{2/3} - \epsilon_2^*. \end{aligned} \tag{5.3.25}$$

Suppose the finite element bases $\{\chi_\alpha\}_{\alpha=1}^m$ and the eigenvectors of $\mathbf{H}$ are $\{w_i\}_{i=1}^{2m}$. We could assume c_i is the coefficient of $w_i(\vec{x})$ expanded in the finite element bases. We can rewrite w_i as (w_{1i}, w_{2i}), where [72]

$$\begin{aligned} w_{1i} &= \sum_{j=1}^{m} c_{i,j}\chi_j, \\ w_{2i} &= \sum_{j=m+1}^{2m} c_{i,j}\chi_{j-m}, \end{aligned} \tag{5.3.26}$$

and they need to satisfy [72]

$$\begin{aligned} \int \phi_{1i}^*(\vec{x}) w_{1i}(\vec{x}) d\vec{x} = 0, \\ \int \phi_{2i}^*(\vec{x}) w_{2i}(\vec{x}) d\vec{x} = 0. \end{aligned} \tag{5.3.27}$$

5.4 Numerical Implementation

This section describes describes how we implement it numerically. The most difficult part is the non-local exact exchange operator. Section 5.4.1 introduces the exact exchange operator. Section 5.4.2 gives some details on how we construct stiffness matrix of exact exchange operator and Hessian operator. Section 5.4.3 shows the general eigenvalue equations of Hartree-Fock and density functional theory Hessian analysis.

5.4.1 Exact Exchange Operator

Our analysis below is for the exact exchange operator of restricted Hartree-Fock model, and it could be generalized to other models in a straightforward way. Suppose the extremum (ϕ^*, ϵ^*) satisfy both (5.3.4) and (5.3.2), and $\{\chi_\alpha\}_{\alpha=1}^m$ are the finite element basis sets, then molecular orbital could be written as

$$\phi^* = \sum_{\alpha=1}^{m} c_\alpha \chi_\alpha \tag{5.4.1}$$

The exchange operator K can be expressed as [72]:

$$K\phi(\vec{x}) = \left(\int \frac{\phi_j^*(\vec{x}')\phi(\vec{x}')}{|\vec{x} - \vec{x}'|} d\vec{x}' \right) \phi_j^*(\vec{x}). \tag{5.4.2}$$

The element of exact exchange operator coefficient matrix is [72]:

$$\begin{aligned} K_{\alpha\beta} &= \int \chi_\alpha(\vec{x}) \left(\int \frac{\phi_j^*(\vec{x}')\chi_\beta(\vec{x}')}{|\vec{x}-\vec{x}'|} d\vec{x}' \right) \phi_j^*(\vec{x}) d\vec{x} \\ &= \sum_{\mu,\nu=1}^{m} \int \chi_\alpha(\vec{x}) \left(\int \frac{[c_\mu \chi_\mu(\vec{x}')]\chi_\beta(\vec{x}')}{|\vec{x}-\vec{x}'|} d\vec{x}' \right) [c_\nu \chi_\nu(\vec{x}')] d\vec{x} \\ &= \sum_{\beta',\alpha'} [c_{\beta'} c_{\alpha'}] \int \chi_\alpha(\vec{x}) \left(\int \frac{\chi_{\beta'}(\vec{x}')\chi_\beta(\vec{x}')}{|\vec{x}-\vec{x}'|} d\vec{x}' \right) \chi_{\alpha'}(\vec{x}) d\vec{x}. \end{aligned} \tag{5.4.3}$$

where $\chi_{\beta'}$, $\chi_{\alpha'}$ are the basis functions, which have some overlap with χ_β, χ_α. The summation $\sum_{\beta',\alpha'}$ have $O(1)$ term.

From (5.4.3), we could see exact exchange operator is a non-local operator. The stiffness matrix of the exact exchange operator will be dense and the time cost of solving the eigenvalue problem will be huge.

5.4.2 Constructing Stiffness Matrix of Hessian Operator

We construct stiffness matrix of Hessian operator based on Python FEniCS [34] package, which is a collection of free software with an extensive list of features for automated, efficient solution of differential equations. Here are the steps in our code [72]:

Step 1. Use FEniCS to generate adaptive mesh and form the stiffness matrix of all local operators.

Step 2. Write a subroutine to construct the stiffness matrix of the nonlocal operator.

Step 3. Combine the generated stiffness matrix in Step 1 and Step 2, and use scipy linalg module to get the spectrum of Hessian matrix.

Here are more details for Step 2. We consider the restricted Hartree-Fock model of two-electrons' system as an example.

Suppose (ϕ^*, ϵ^*) is the solution of the Euler-Lagrange equation, and $\{\chi_\alpha\}_{\alpha=1}^m$ are the finite element bases. We assume ϕ^* could be written as

$$\phi^* = \sum_{\alpha=1}^{m} c_\alpha \chi_\alpha \tag{5.4.4}$$

The element of the stiffness matrix for exchange operator [72]:

$$K_{i,j} = 2 \int \chi_i(\vec{x}) \left(\int \frac{\phi^*(\vec{x}')\chi_j(\vec{x}')}{|\vec{x} - \vec{x}'|} d\vec{x}' \right) \phi^*(\vec{x}) d\vec{x}. \tag{5.4.5}$$

Let T be a collection of finite elements: $T = \{E_k\}_{k=1}^{N_T}$. $K_{i,j}$ could be rewritten as [72]

$$K_{i,j} = 2 \sum_{k,k'=1}^{N_T} \int_{E_k} \chi_i(\vec{x}) \left(\int_{E_{k'}} \frac{\phi^*(\vec{x}')\chi_j(\vec{x}')}{|\vec{x} - \vec{x}'|} d\vec{x}' \right) \phi^*(\vec{x}) d\vec{x}. \tag{5.4.6}$$

We make an approximation to $|\vec{x} - \vec{x}'|$ is $R + \delta$, where $R_{k,k'}$ is the distance between the center of element E_k and $E_{k'}$ and δ is a small positive constant to avoid the singular condition. Then [72]

$$\begin{aligned} K_{i,j} &= 2 \sum_{k,k'=1}^{N_T} \frac{1}{R_{k,k'} + \delta} \int_{E_k} \chi_i(\vec{x}) \phi^*(\vec{x}) d\vec{x} \int_{E_{k'}} \chi_j(\vec{x}') \phi^*(\vec{x}') d\vec{x}', \\ &= 2 \sum_{k,k'=1}^{N_T} \sum_{\alpha,\beta=1}^{M} \frac{C_\alpha C_\beta}{R_{k,k'} + \delta} \int_{E_k} \chi_i(\vec{x}) \chi_\alpha(\vec{x}) d\vec{x} \int_{E_{k'}} \chi_j(\vec{x}') \chi_\beta(\vec{x}') d\vec{x}' \end{aligned} \tag{5.4.7}$$

Suppose $\{k_1, k_2, k_3, k_4\}$ and $\{k_1', k_2', k_3', k_4'\}$ are four vertices of element E_k and $E_{k'}$. Then [72]

$$K_{i,j} = 2 \sum_{k,k'=1}^{N_T} \sum_{\substack{\alpha=\{k_1,k_2,k_3,k_4\} \\ \beta=\{k_1',k_2',k_3',k_4'\}}} \frac{C_\alpha C_\beta}{R_{k,k'} + \delta} \int_{E_k} \chi_i(\vec{x}) \chi_\alpha(\vec{x}) d\vec{x} \int_{E_{k'}} \chi_j(\vec{x}') \chi_\beta(\vec{x}') d\vec{x}' \tag{5.4.8}$$

Algorithm 3 [72] describes the process to assemble exact exchange operator.

5.4.3 Generalized Eigenvalue Problem

This section shows the formulations of generalized eigenvalue problem for Hartree-Fock and density functional theory models.

Restricted Hartree-Fock/Restricted Density Functional Theory

The formulations of generalized eigenvalue problems for restricted Hartree-Fock and restricted density functional theory model are straightforward. We could discretize

Algorithm 3 Stiffness Matrix Assembly on Exact Exchange Operator

for E_k in T **do**
  for $E_{k'}$ in T **do**
    for i in $\{k_1, k_2, k_3, k_4\}$ **do**
      for j in $\{k'_1, k'_2, k'_3, k'_4\}$ **do**
        for α in $\{k_1, k_2, k_3, k_4\}$ **do**
          for β in $\{k'_1, k'_2, k'_3, k'_4\}$ **do**
            $K_{i,j} += 2\frac{C_\alpha C_\beta}{R_{k,k'}+\delta} \int_{E_k} \chi_i(\vec{x})\chi_\alpha(\vec{x})d\vec{x} \int_{E_{k'}} \chi_j(\vec{x}')\chi_\beta(\vec{x}')d\vec{x}'$

the eigenvalue problem [72]

$$\mathbf{H}w_i(\vec{x}') = \lambda_i w_i(\vec{x}), i = 1, \ldots, m, \tag{5.4.9}$$

with the finite element bases $\{\chi_\alpha\}_{\alpha=1}^m$. We will get following generalized eigenvalue problem [72]

$$Hc_i = \lambda_i M c_i, i = 1, \ldots, m, \tag{5.4.10}$$

where the matrices H and M are

$$\begin{aligned} H_{\alpha,\beta} &= (\chi_\alpha, \mathbf{H}\chi_\beta), \alpha, \beta = 1, \ldots, m, \\ M_{\alpha,\beta} &= (\chi_\alpha, \chi_\beta), \alpha, \beta = 1, \ldots, m, \end{aligned} \tag{5.4.11}$$

and c_i is the coefficient of $w_i(\vec{x})$ expanded in the finite element bases $\{\chi_\alpha\}_{\alpha=1}^m$, and it satisfies

$$w_i(\vec{x}) = c_{i,j}\chi_j, i, j = 1, \ldots, m, \tag{5.4.12}$$

Unrestricted Hartree-Fock/Unrestricted Density Functional Theory

The formulations of generalized eigenvalue problems for unrestricted Hartree-Fock and unrestricted density functional theory model are not trivial. The dimension of Hessian matrix H is $2m$, which is twice larger than the number of finite element basis sets m. Thus we need to increase the size of our basis sets twice, and new basis sets $\{\eta_\alpha\}_{\alpha=1}^{2m}$ are [72]

$$\left\{ \begin{pmatrix} \chi_1 \\ 0 \end{pmatrix}, \cdots, \begin{pmatrix} \chi_m \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \chi_1 \end{pmatrix}, \cdots, \begin{pmatrix} 0 \\ \chi_m \end{pmatrix} \right\}. \tag{5.4.13}$$

We could discretize the eigenvalue problem

$$\mathbf{H}w_i(\vec{x}') = \lambda_i w_i(\vec{x}), i = 1, \ldots, 2m, \tag{5.4.14}$$

with the new finite element bases in (5.4.13), and we will get following generalized eigenvalue problem

$$Hc_i = \lambda_i M c_i, i = 1, \ldots, 2m, \tag{5.4.15}$$

where the matrices H and M are [72]

$$\begin{aligned} H_{\alpha,\beta} &= (\eta_\alpha, \mathbf{H}\eta_\beta), \alpha, \beta = 1, \ldots, 2m, \\ M_{\alpha,\beta} &= (\eta_\alpha, \eta_\beta), \alpha, \beta = 1, \ldots, 2m, \end{aligned} \tag{5.4.16}$$

and c_i is the coefficient of $w_i(\vec{x})$ expanded in the finite element bases $\{\eta_\alpha\}_{\alpha=1}^m$, and it satisfies

$$w_i(\vec{x}) = c_{i,j}\eta_j, i, j = 1, \ldots, 2m, \tag{5.4.17}$$

5.5 Numerical Results

The source code of our numerical implementation [72] could be found [68]. We show in Section 5.5.1 the bifurcation figure of H_2 restricted Hartree-Fock and unrestricted Hartree-Fock models, and Hessian analysis results of extremums of restricted Hartree-Fock and unrestricted Hartree-Fock models at different bond lengths. Section 5.5.2 shows the bifurcation figure of H_2 restricted density functional theory and unrestricted density functional theory models. Section 5.5.3 shows the solutions of the restricted Hartree-Fock and unrestricted Hartree-Fock models of Two-electron atom.

5.5.1 Hartree-Fock: H_2

Figure 5.3 [72] shows the bifurcation figure of H_2 restricted Hartree-Fock and H_2 unrestricted Hartree-Fock. The y-axis is the total energy of Self-consistent solver, and the x-axis is the interatomic distance. From Figure 5.3, you could observe a threshold around $2.5\ a.u.$. If interatomic distance is smaller than the threshold, both H_2 restricted Hartree-Fock and H_2 unrestricted Hartree-Fock generate same delocalized solutions. If interatomic distance is larger than the threshold, H_2 unrestricted Hartree-Fock

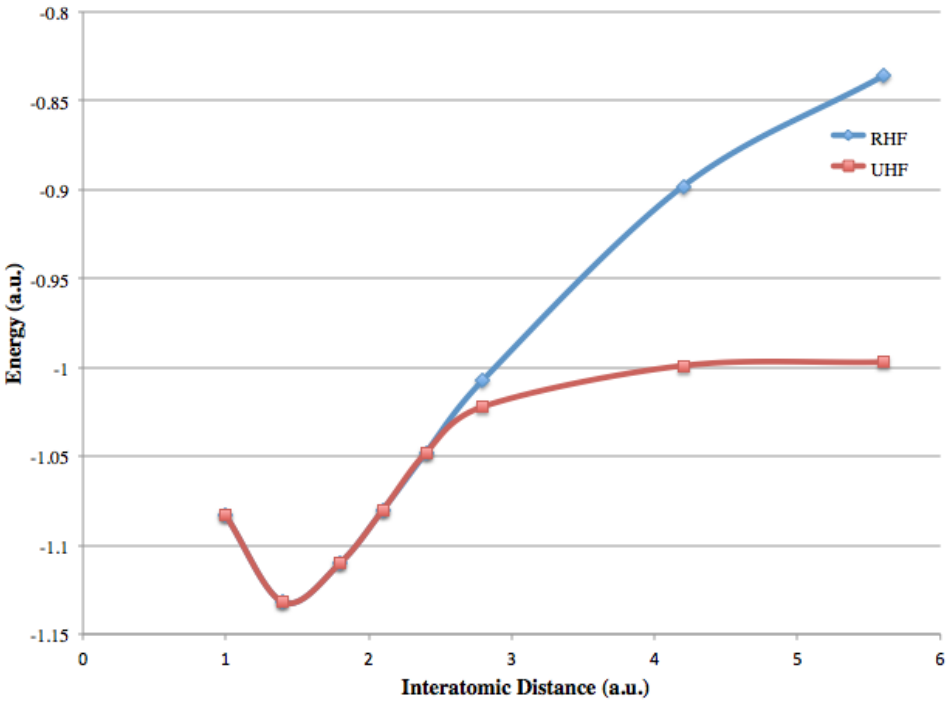

Figure 5.3: Bifurcation Figure of H_2 RHF and H_2 UHF

generate localized solutions and H_2 restricted Hartree-Fock still generate delocalized solutions. The H_2 unrestricted Hartree-Fock localized solutions is smaller than the H_2 restricted Hartree-Fock delocalized solutions.

Table 5.1 [72] shows restricted Hartree-Fock solutions are local minimizers when we consider them on the energy landscape of restricted Hartree-Fock.

Table 5.2 [72] shows delocalized unrestricted Hartree-Fock solutions are local minimizers, if the bond length is small. Localized unrestricted Hartree-Fock solutions are still local minimiers, but delocalized unrestricted Hartree-Fock solutions are saddle

Table 5.1: Hessian Analysis on RHF

Bond Length	Result	Details
1.4 a.u.	Local Minimizer	all evs on the constraint manifold are positive.
2.8 a.u.	Local Minimizer	all evs on the constraint manifold are positive.
5.6 a.u.	Local Minimizer	all evs on the constraint manifold are positive.

Table 5.2: Hessian Analysis on UHF

Bond Length	Solution	Result	Details
1.4 a.u.	delocalized solution	Local Minimizer	all evs on the constraint manifold are positive.
2.8 a.u.	delocalized solution	Saddle Point	1 ev on the constraint manifold is negative.
5.6 a.u.	delocalized solution	Saddle Point	1 ev on the constraint manifold is negative.
2.8 a.u.	localized solution	Local Minimizer	all evs on the constraint manifold are positive.
5.6 a.u.	localized solution	Local Minimizer	all evs on the constraint manifold are positive.

points, if the bond length is large.

5.5.2 Density Functional Theory: H_2

Figure 5.4 [72] shows the bifurcation figure of H_2 restricted local spin density approximation model and H_2 unrestricted local spin density approximation model. The y-axis is the total energy of Self-consistent solver, and the x-axis is the interatomic distance. We didn't include the correlation energy term in the numerical implementation. We could observe a threshold around $3.0\ a.u.$. If interatomic distance is smaller than the threshold, both H_2 restricted local spin density approximation and H_2 unrestricted local spin density approximation generate same delocalized solutions. If interatomic distance is larger than the threshold, H_2 unrestricted local spin density approximation generate localized solutions and H_2 restricted local spin density approximation still generate delocalized solutions. The H_2 unrestricted local spin density approximation localized solutions is smaller than the H_2 restricted local spin density approximation delocalized solutions.

Table 5.3 [72] shows solutions of restricted local spin density approximation model are local minimizers when we consider them on the energy landscape of restricted

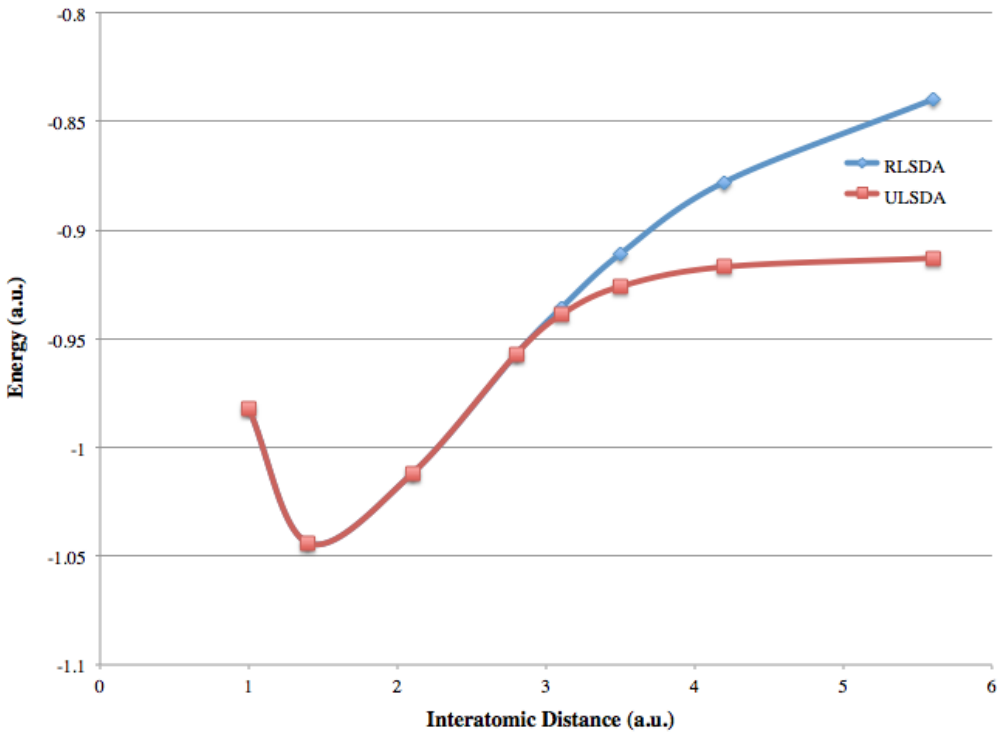

Figure 5.4: Bifurcation Figure of H_2 RLSDA and H_2 ULSDA

Table 5.3: Hessian Analysis on RLSDA

Bond Length	Result	Details
1.4 a.u.	Local Minimizer	all evs on the constraint manifold are positive.
5.6 a.u.	Local Minimizer	all evs on the constraint manifold are positive.
7.2 a.u.	Local Minimizer	all evs on the constraint manifold are positive.

Table 5.4: Hessian Analysis on ULSDA

Bond Length	Solution	Result	Details
1.4 a.u.	delocalized solution	Local Minimizer	all evs on the constraint manifold are positive.
5.6 a.u.	delocalized solution	Saddle Point	1 ev on the constraint manifold is negative.
7.2 a.u.	delocalized solution	Saddle Point	1 ev on the constraint manifold is negative.
5.6 a.u.	localized solution	Local Minimizer	all evs on the constraint manifold are positive.
7.2 a.u.	localized solution	Local Minimizer	all evs on the constraint manifold are positive.

local spin density approximation model.

Table 5.4 [72] shows delocalized unrestricted local spin density approximation solutions are local minimizers, if the bond length is small. Localized unrestricted local spin density approximation solutions are still local minimiers, but delocalized unrestricted local spin density approximation solutions are saddle points, if the bond length is large.

5.5.3 Hartree Fock: Two-electron atom

Figure 5.5 [72] shows the relationship between the total energy of a two-electron atom's unrestricted Hartree-Fock system and the nuclear charge of the system. If $Z \leq$

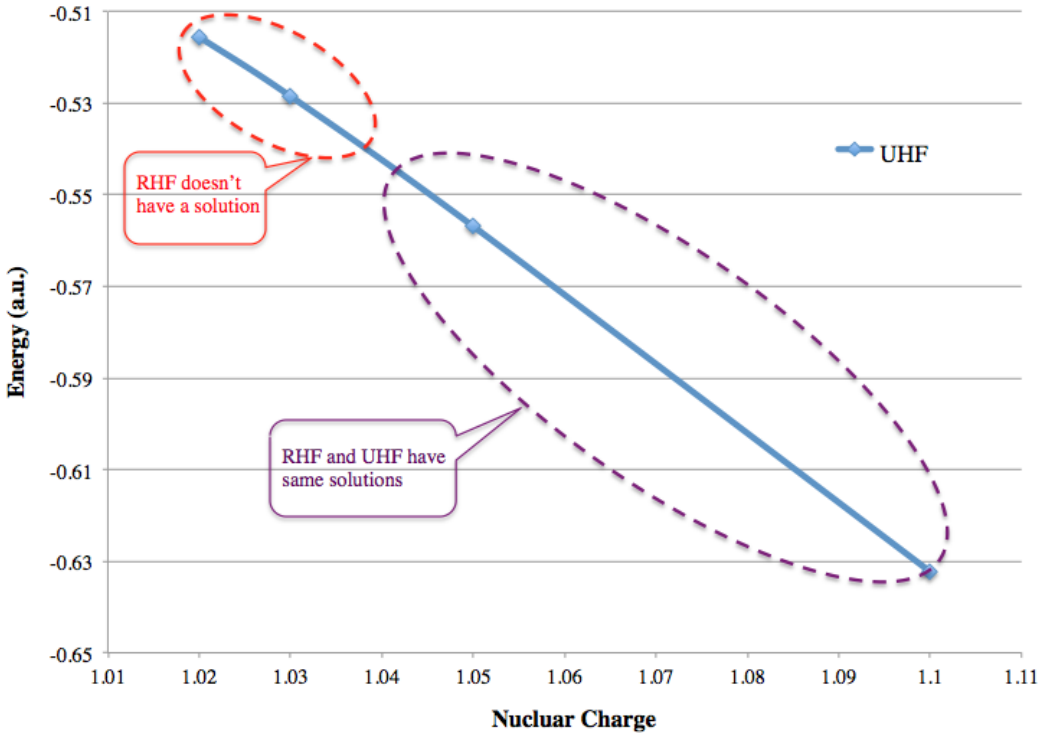

Figure 5.5: Two-electron Atom UHF Energy vs Nuclear Charge Z

1.03, restricted Hartree-Fock can't get convergent, and only unrestricted Hartree-Fock can get solutions. If $Z \geq 1.05$, both restricted Hartree-Fock and unrestricted Hartree-Fock generate same solutions.

From Table 5.5 [72], we find a threshold between $1.03 < Z < 1.05$ for unrestricted Hartree-Fock. If $Z \leq 1.03$, the first unrestricted Hartree-Fock orbital is a weakly bounded state that is close to H atom ground state energy, and the second unrestricted Hartree-Fock orbital is a bounded state. If $Z \geq 1.05$, the first unrestricted Hartree-Fock orbital is the same as the second unrestricted Hartree-Fock orbital.

5.6 Conclusion

We successfully proposed and numerically implemented a systematic way to study stability condition of Hartree-Fock and density functional theory models [72]. We generated a few bifurcation figures for Hartree-Fock and density functional theory.

Table 5.5: Two-electron Atom UHF Orbitals vs Nuclear Charge Z

Z	1st Orbital Energy	2nd Orbital Energy	Total Energy
1.02	-0.0048 a.u.	-0.2725 a.u.	-0.5155 a.u.
1.03	-0.0135 a.u.	-0.1888 a.u.	-0.5285 a.u.
1.05	-0.0688 a.u.	-0.0688 a.u.	-0.5569 a.u.
1.1	-0.0911 a.u.	-0.0911 a.u.	-0.6323 a.u.

Hessian analysis have also been conducted on extremums of Hartree-Fock and density functional theory models. A weakly bounded state is found when the nuclear charge of the two-electron atom gets really close to 1. It is a powerful tool to understand the stability of different solutions of Hartree-Fock and density functional theory models.

Chapter 5, is currently begin prepared for submission for publication of the material. Hu, H., Marzuola, J., Lu, J., Song, D., Weare, J., and Holst, M. The dissertation author was the primary investigator and author of this material.

Chapter 6

Conclusion

We developed finite-element formulations for Hartree-Fock and density functional theory models, including restricted Hartree-Fock, unrestricted Hartree-Fock, local spin density approximation, generalized gradient approximation, meta generalized gradient approximation and more generalized density functional theory models [67]. We probed the well-posedness, and the existence of minimizers for these models in a finite calculation domain for both the all-electron problem and pseudopotential approximations. We also established the convergence of the finite-element approximation, and the convergence of finite element approximation with numerical quadratures by Γ-convergence for both the all-electron problem and pseudopotential approximations. It will be useful for the development of studying Hartree-Fock and density functional theory models numerically by finite element method.

In order to handle exact exchange operator, we successfully proposed and numerically implemented a linear time cost and memory cost method [71] with finite element bases. Multigrid method is used in our algorithm in order to achieve linear scaling. A variety of numerical experiments for atoms and molecules demonstrate reliable precision and speed. Our method could be applied to all kinds of Hartree-Fock, density functional theory, hybrid models. The development of this method make finite element method become a competitive method in models with exact exchange energy.

Also we successfully proposed and numerically implemented a systematic way

to study stability condition of Hartree-Fock and density functional theory models [72]. We generated a few bifurcation figures for Hartree-Fock and density functional theory. Hessian analysis have also been conducted on extremums of Hartree-Fock and density functional theory models. A weakly bounded state is found when the nuclear charge of the two-electron atom gets really close to 1. It is a powerful tool to understand the stability of different solutions of Hartree-Fock and density functional theory models.

Appendix A

Plasmonic Dark Field Microscopy

This is my Ph. D. work [70] in another group during 2009-2010.

A.1 Abstract

We propose plasmonic dark field microscopy, which utilizes a chip-scale integrated plasmonic multilayered structure to substitute the bulky and expensive conventional condenser optics. Experimental results show that we can get high contrast image using the compact, low-cost and alignment free plasmonic dark field microscopy.

A.2 Content

The rapid progress in nanoscale science and technology demands new microscopy techniques that possess both high resolution and high contrast capabilities. Many techniques have been developed to improve resolution such as near-field scanning optical microscopy (NSOM),[147, 107, 125, 14, 106] stimulated emission depletion (STED) microscopy,[62, 79] photoactivated localization microscopy (PALM),[13, 64] stochastic optical reconstruction microscopy (STORM),[130] randomly adsorbed molecule microscopy (RAM),[163] the superlens,[37, 93] and hyperlens,,[94, 142] structured illumination microscopy (SIM),[53, 52] etc. Few techniques, however, have been developed for high contrast imaging. Dark field (DF) microscopy is widely used to view the object that has low contrast in bright-field microscopy.

In the conventional DF microscopy, the central part of the illumination light which ordinarily passes through and around the sample is blocked by a light stop, allowing only oblique rays to strike the sample on the microscope slide, as shown in Figure (A.1)(a). This is of great help when the objects have refractive indices very close to those of their surroundings and are difficult to image in the conventional bright field microscopy. While the conventional DF microscopy can achieve high contrast imaging, its resolution may also be improved using a high numerical aperture (NA) configuration of the condenser/objective pair. However, the NA of the objective can not be larger than that of the condenser to avoid oblique illuminating rays entering the objective. Also the high NA condensers, such as the cardioid condenser,[126] are very sensitive

to alignment and thus must be accurately positioned and aligned to the very sharp cone of illumination, making it very hard to use. In addition, the illumination light in such a high NA arrangement must be very strong due to the sharp illumination cone, which is wasteful of energy. Thus conventional DF microscopy is instrumentally complex, costly and bulky.

In order to resolve the aforementioned limitations associated with conventional DF microscopy, we propose another DF technology. The proposed technology utilizes a chip-scale integrated plasmonic multilayered structure as the substitute of the bulky and expensive conventional condenser optics. We term this integrated structure as a plasmonic condenser (PC) and this imaging technology as plasmonic dark field (PDF) microscopy. The PDF microscopy is schematically shown in Figure (A.1)(b). The PC is the critical element in PDF microscopy, which uses surface plasmons (SPs) to illuminate the sample. The effective NA of the PC is the ratio of the wavevetors (k-vectors) of the SPs to the photons in free space. Because the SPs, which are surface electromagnetic waves formed by collective oscillation of electrons at a metal-dielectric interface, may possess k-vectors much higher than those of the free-space photons at the same frequencies, large effective NAs can be achieved by the PCs. In addition, the SPs are evanescent on the metal-dielectric interface and do not propagate to the far field, so they can not be detected by an objective lens in the far field. When objects are brought to the vicinity of the metal surface, the SPs can be converted into free-space photons, which can propagate to the far field where they can be detected. Meanwhile the SPs will remain evanescent in the areas without the objects. Therefore, a high contrast DF image of the samples in the far field can be formed. The essential feature of the PC is the integration of a thin layer of metal with a SP excitation mechanism in the very proximity of the metal to excite SPs using the near field coupling. The SP excitation mechanisms can be quantum dots, photoluminescent and electroluminescent materials, etc. In this work, we design and demonstrate the PDF concept using a chip-scale integrated active PC with fluorescent Rhodamine 6G as the active material.

The active PC is schematically shown in Figure (A.1)(c). It is a three-layer structure, i.e., an active layer of the mixture of polymethyl methacrylate (PMMA) and Rhodamine 6G molecules sandwiched between a glass substrate and a thin layer of sil-

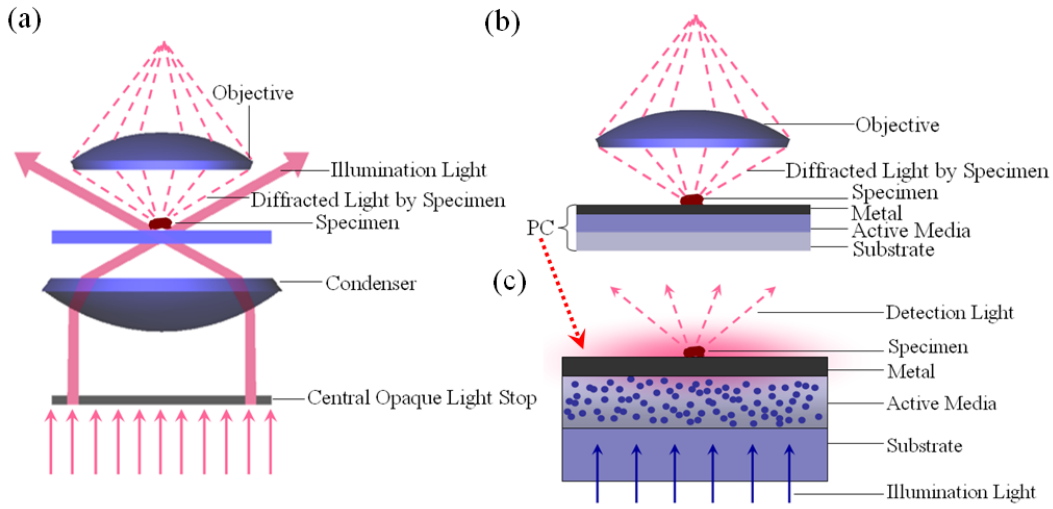

Figure A.1: Schematic configurations of (a) conventional dark field microscopy, (b) plasmonic dark field microscopy and (c) plasmonic condenser.

ver. When the Rhodamine 6G molecules are illuminated by the incident light from the glass side, the emitted fluorescence will excite the SPs on the silver-air interface by near field coupling. This active PC can be easily fabricated using standard microfabrication techniques. The Rhodamine 6G molecules was first mixed with PMMA at the concentration around 10-4 mol/L. Then the mixture was spincoated on a cover glass substrate. After a 2 minutes soft bake process, a 200-nm thick layer mixture as the active layer was obtained. Finally, a 60-nm thick silver film was deposited on top of the PMMA using the E-beam evaporation method.

To verify the PDF idea, we tested the fabricated fluorescent active PC using aggregations of 2-m diameter polystyrene beads. A two-dimensional (2D) hexagonally close packed lattice of the polystyrene beads was fabricated on top of the silver film using a self-assembly method. The sample was first examined by using a standard optical microscope dark field objective (EC Epiplan-Neofluar, 50X, NA = 0.8). The reflection mode dark field image is shown in Figure (A.2)(a) with a green filter (56010 nm, bandpass) added in the light path. As a comparison, the same sample was examined by PDF microscopy and the image is shown in Figure (A.2)(b). Note that the same objective was used for both techniques, however, in the PDF microscopy, a pair of band pass filters were used for the excitation (53010 nm, bandpass) and detection light (56010 nm,

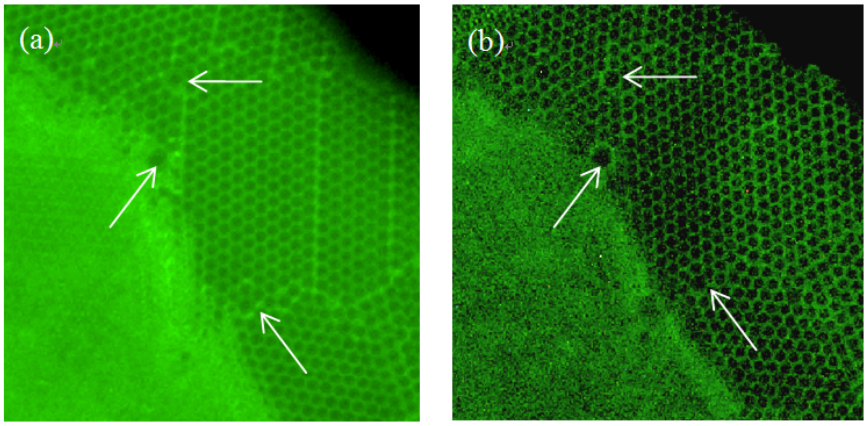

Figure A.2: Images of the monolayer polystyrene bead lattice, obtained using (a) the conventional dark field microscopy and (b) the plasmonic dark field microscopy.

bandpass), respectively, based on the properties of the dye. The images obtained using the conventional DF microscopy and the PDF microscopy are shown in Figures (A.2)(a) and (b), respectively. It can be seen from Figure (A.2) that the contrast of the PDF microscopy is better than that of the conventional DF microscopy. This is especially true for the contrast between the center parts and the boundaries of the polystyrene beads, as marked by the arrows.

Because the illumination of PDF microscopy only exists at the interface of the PC in the form of SPs, the depth of field and the sensitivity of the PDF microscopy along the direction normal to the surface of the PC are solely determined by the decay property of the SPs. Due to the differences between the SPs and the conventional illumination light, the image information obtained using PDF microscopy and conventional DF microscopy are significantly different. The sample was examined using another standard optical microscope dark field objective (EC Epiplan-Neofluar, 20X, NA = 0.22). Figures (A.3)(a) and (b) show the images of the sample, which is two layers of 2-m polystyrene beads, obtained using conventional DF microscopy and PDF microscopy, respectively. It can be seen that conventional DF microscopy mainly provides contrast at the top layer; while PDF image reveals the bottom layer with much improved contrast. Notice the additional layer of particles doesnt affect the image quality much in PDF microscopy.

As shown above, we have experimentally demonstrated the PDF concept for high contrast imaging. The image quality of PDF microscopy is strongly related to the near

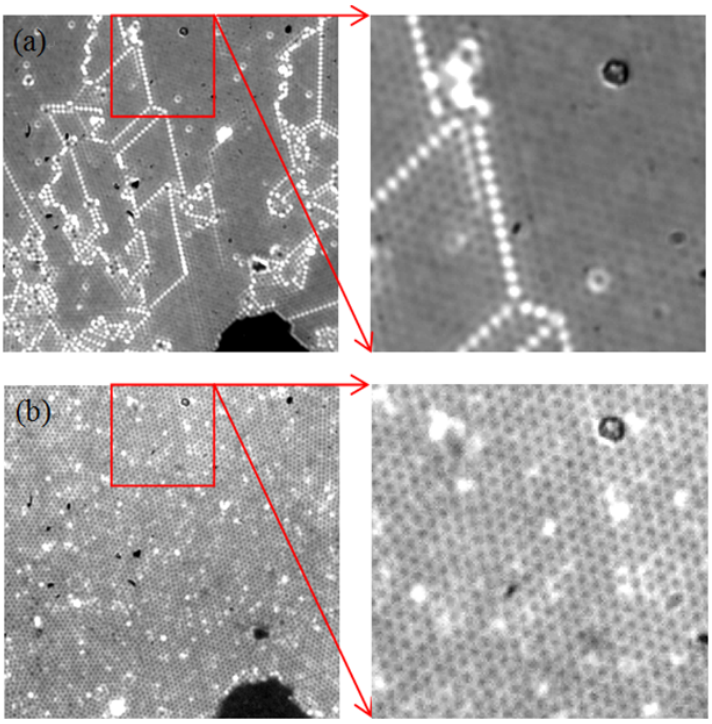

Figure A.3: Images of the dual-layer polystyrene bead lattice (diameter 2 um), obtained using (a) the conventional dark field microscopy and (b) the plasmonic dark field microscopy.

field coupling between the active medium and the SPs on the metal-insulator interfaces in the PC; this has been intensively investigated.[25, 7, 158] One of the well-known models is to study the coupling of a dipole in the vicinity of a metallic interface to SPs, as the fluorescent dyes can be modeled as dipoles with isotropic orientations. When the thickness of the silver film is 60 nm, about 40% of the energy can be transferred to the SPs.[7, 158] The remaining 60% of the light is attenuated by the 60-nm silver layer, thus resulting in very low transmission. So in the far field only the free-space photons arising from SP scattering by the object can be detected. The surface roughness, in addition to the object, can also scatter the SPs into free space photons, thus contributing to the background. In the demonstrated PC, the surface roughness is less than 2 nm in terms of root mean square (RMS). Therefore, the scattered light from the random roughness of such a smooth interface is very weak, resulting in the high contrast as shown in Figure (A.2)(b).

In conclusion, we have demonstrated the PDF microscopy concept using a chip-scale integrated multilayered fluorescent active PC. It can be easily extended to other types of PCs as mentioned above. The PDF microscopy idea can also be applied to fluorescence microscopy, which is similar to the extension of total internal reflection (TIR) microscopy to total internal reflection fluorescence (TIRF) microscopy.[4] Because the PCs can have large effective NA, high resolution beyond the diffraction limit may be obtained with the compact and low-cost PDF microscopy. It is worth noting that the field of view of the PDF microscopy may be large, because it is dependent on the planar area of the PC, rather than the relatively short propagation length of the SPs. The proposed PDF technology shows better contrast imaging capabilities for thin samples than the conventional DF microscopy. Finally, because the PCs are planar, the PDF microscopy does not require very sensitive alignment as in the conventional DF.

Appendix A, in full, is a reprint of the material as it appears in Applied Physics Letter 2010. Hu, H., Ma, C., and Liu Z. The dissertation author was the primary investigator and author of this material.

Bibliography

[1] R. A. Adams and J. J. Fournier. *Sobolev spaces*, volume 140. Access Online via Elsevier, 2003.

[2] R. Alizadegan, K. Hsia, and T. Martinez. A divide and conquer real space finite-element hartree–fock method. *The Journal of chemical physics*, 132(3):034101, 2010.

[3] A. Anantharaman and E. Cancès. Existence of minimizers for kohn–sham models in quantum chemistry. In *Annales de l'Institut Henri Poincare (C) Non Linear Analysis*, volume 26, pages 2425–2455. Elsevier, 2009.

[4] D. Axelrod, T. P. Burghardt, and N. L. Thompson. Total internal reflection fluorescence. *Annual review of biophysics and bioengineering*, 13(1):247–268, 1984.

[5] R. E. Bank and T. Dupont. An optimal order process for solving finite element equations. *Mathematics of Computation*, 36(153):35–51, 1981.

[6] G. Bao, G. Hu, and D. Liu. An h-adaptive finite element solver for the calculation of the electronic structures. *Journal of Computational Physics*, 231:4867–4979, 2012.

[7] W. Barnes. Fluorescence near interfaces: the role of photonic mode density. *journal of modern optics*, 45(4):661–699, 1998.

[8] T. L. Beck. Real-space mesh techniques in density-functional theory. *Reviews of Modern Physics*, 72(4):1041, 2000.

[9] A. Becke and M. Roussel. Exchange holes in inhomogeneous systems: A coordinate-space model. *Physical Review A*, 39(8):3761, 1989.

[10] A. D. Becke. Density-functional exchange-energy approximation with correct asymptotic behavior. *Physical Review A*, 38(6):3098, 1988.

[11] A. D. Becke. Density-functional thermochemistry. iii. the role of exact exchange. *The Journal of Chemical Physics*, 98(7):5648–5652, 1993.

[12] A. D. Becke. A new mixing of hartree–fock and local density-functional theories. *The Journal of Chemical Physics*, 98(2):1372–1377, 1993.

[13] E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, and H. F. Hess. Imaging intracellular fluorescent proteins at nanometer resolution. *Science*, 313(5793):1642–1645, 2006.

[14] E. Betzig, J. Trautman, T. Harris, J. Weiner, and R. Kostelak. Breaking the diffraction barrier: optical microscopy on a nanometric scale. *Science*, 251(5000):1468–1470, 1991.

[15] D. Bowler, R. Choudhury, M. Gillan, and T. Miyazaki. Recent progress with large-scale ab initio calculations: the conquest code. *physica status solidi (b)*, 243(5):989–1000, 2006.

[16] A. Braides. *Gamma-convergence for Beginners*, volume 22. Oxford University Press, 2002.

[17] J. H. Bramble. *Multigrid methods*, volume 294. CRC Press, 1993.

[18] J. H. Bramble and J. E. Pasciak. New convergence estimates for multigrid algorithms. *Mathematics of computation*, 49(180):311–329, 1987.

[19] A. Brandt. Algebraic multigrid theory: The symmetric case. *Applied mathematics and computation*, 19(1):23–56, 1986.

[20] A. Brandt, S. McCoruick, and J. Huge. Algebraic multigrid (amg) f0r sparse matrix equati0ns. *Sparsity and its Applications*, page 257, 1985.

[21] S. C. Brenner and L. R. Scott. *The mathematical theory of finite element methods*, volume 15. Springer, 2008.

[22] E. J. Bylaska, M. Holst, and J. H. Weare. Adaptive finite element method for solving the exact kohn- sham equation of density functional theory. *Journal of Chemical Theory and Computation*, 5(4):937–948, 2009.

[23] A. Castro, H. Appel, M. Oliveira, C. A. Rozzi, X. Andrade, F. Lorenzen, M. Marques, E. Gross, and A. Rubio. octopus: a tool for the application of time-dependent density functional theory. *physica status solidi (b)*, 243(11):2465–2488, 2006.

[24] D. Ceperley and B. Alder. Exchange-correlation potential and energy for density-functional calculation. *Phys. Rev. Lett*, 45:567–581, 1980.

[25] R. Chance, A. Prock, and R. Silbey. Molecular fluorescence and energy transfer near interfaces. *Adv. Chem. Phys*, 37(1):65, 1978.

[26] J. R. Chelikowsky, N. Troullier, and Y. Saad. Finite-difference-pseudopotential method: Electronic structure calculations without a basis. *Physical review letters*, 72(8):1240, 1994.

[27] H. Chen, X. Dai, X. Gong, L. He, and A. Zhou. Adaptive finite element approximations for kohn-sham models. *arXiv preprint arXiv:1302.6896*, 2013.

[28] P. G. Ciarlet. *The finite element method for elliptic problems*. Access Online via Elsevier, 1978.

[29] J. Čížek. On the correlation problem in atomic and molecular systems. calculation of wavefunction components in ursell-type expansion using quantum-field theoretical methods. *The Journal of Chemical Physics*, 45(11):4256–4266, 1966.

[30] F. R. D. An introduction to algebraic an introduction to algebraic multigrid.

[31] G. Dal Maso. *An Introduction to [gamma]-convergence*, volume 8. Springer, 1993.

[32] P. A. M. Dirac. The principles of quantum mechanics. *The International Series of Monographs on Physics, Oxford: Clarendon Press, 1947*, 1, 1947.

[33] J. Dobson. Spin-density functionals for the electron correlation energy with automatic freedom from orbital self-interaction. *Journal of Physics: Condensed Matter*, 4(39):7877, 1992.

[34] T. Dupont, J. Hoffman, C. Johnson, R. C. Kirby, M. G. Larson, A. Logg, and L. R. Scott. *The FEniCS project*. Chalmers Finite Element Centre, Chalmers University of Technology, 2003.

[35] L. C. Evans. Partial differential equations. graduate studies in mathematics. *American Mathematical Society*, 2, 1998.

[36] J. Fang, X. Gao, and A. Zhou. A kohn–sham equation solver based on hexahedral finite elements. *Journal of Computational Physics*, 231(8):3166–3180, 2012.

[37] N. Fang, H. Lee, C. Sun, and X. Zhang. Sub–diffraction-limited optical imaging with a silver superlens. *Science*, 308(5721):534–537, 2005.

[38] J.-L. Fattebert and M. B. Nardelli. Finite difference methods for ab initio electronic structure and quantum transport calculations of nanostructures. 2002.

[39] E. Fermi. Un metodo statistico per la determinazione di alcune priorieta dell'atome. *Rend. Accad. Naz. Lincei*, 6(602-607):32, 1927.

[40] E. Fermi. A statistical method for the determination of some atomic properties and the application of this method to the theory of the periodic system of elements. *Z. Phys*, 48(73):29, 1928.

[41] E. Fermi. Sulla deduzione statistica di alcune proprietá dell'atomo. applicazione alla teoria del sistema periodico degli elementi. *Rend. Lincei*, 7:342, 1928.

[42] J. Flores, E. Clementi, and V. Sonnad. A p-version finite-element approach for atomic hartree-fock calculations. *Chemical physics letters*, 163(2):198–201, 1989.

[43] J. Flores, E. Clementi, and V. Sonnad. Solution of atomic hartree–fock equations with the p version of the finite element method. *The Journal of chemical physics*, 91(11):7030–7038, 1989.

[44] J. B. Foresman and A. Frisch. Exploring chemistry with electronic structure methods, 1996. *Gaussian Inc, Pittsburgh, PA*, pages 98–99, 1996.

[45] V. Gavini, K. Bhattacharya, and M. Ortiz. Quasi-continuum orbital-free density-functional theory: A route to multi-million atom non-periodic dft calculation. *Journal of the Mechanics and Physics of Solids*, 55(4):697–718, 2007.

[46] V. Gavini, J. Knap, K. Bhattacharya, and M. Ortiz. Non-periodic finite-element formulation of orbital-free density functional theory. *Journal of the Mechanics and Physics of Solids*, 55(4):669–696, 2007.

[47] I. M. Gelfand, I. M. Gelfand, and S. V. Fomin. *Calculus of variations*. Courier Dover Publications, 2000.

[48] L. Genovese, B. Videau, M. Ospici, T. Deutsch, S. Goedecker, and J.-F. Méhaut. Daubechies wavelets for high performance electronic structure calculations: The bigdft project. *Comptes Rendus Mécanique*, 339(2):149–164, 2011.

[49] D. Gilbarg and N. Trudinger. Elliptic pde of second order, 1983.

[50] M. Goeppert-Mayer and A. L. Sklar. Calculations of the lower excited levels of benzene. *The Journal of Chemical Physics*, 6:645, 1938.

[51] X. Gonze, J.-M. Beuken, R. Caracas, F. Detraux, M. Fuchs, G.-M. Rignanese, L. Sindic, M. Verstraete, G. Zerah, F. Jollet, et al. First-principles computation of material properties: the abinit software project. *Computational Materials Science*, 25(3):478–492, 2002.

[52] M. G. Gustafsson. Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy. *Journal of microscopy*, 198(2):82–87, 2000.

[53] M. G. Gustafsson. Nonlinear structured-illumination microscopy: wide-field fluorescence imaging with theoretically unlimited resolution. *Proceedings of the National Academy of Sciences of the United States of America*, 102(37):13081–13086, 2005.

[54] W. Hackbusch. *Multi-grid methods and applications*, volume 4. Springer-Verlag Berlin, 1985.

[55] R. J. Harrison, G. I. Fann, T. Yanai, Z. Gan, and G. Beylkin. Multiresolution quantum chemistry: Basic theory and initial applications. *The Journal of chemical physics*, 121(23):11587–11598, 2004.

[56] W. A. Harrison. *Pseudopotentials in the Theory of Metals*, volume 201. WA Benjamin New York, 1966.

[57] W. J. Hehre, R. F. Stewart, and J. A. Pople. Self-consistent molecular-orbital methods. i. use of gaussian expansions of slater-type atomic orbitals. *The Journal of Chemical Physics*, 51:2657, 1969.

[58] V. Heine. The pseudopotential concept. *Solid state physics*, 24:1–36, 1970.

[59] D. Heinemann, B. Fricke, and D. Kolb. Accurate hartree-fock-slater calculations on small diatomic molecules with the finite-element method. *Chemical physics letters*, 145(2):125–127, 1988.

[60] D. Heinemann, B. Fricke, and D. Kolb. Solution of the hartree-fock-slater equations for diatomic molecules by the finite-element method. *Physical Review A*, 38(10):4994, 1988.

[61] D. Heinemann, A. Rosen, and B. Fricke. Solution of the hartree-fock equations for atoms and diatomic molecules with the finite element method. *Physica Scripta*, 42(6):692, 1990.

[62] S. W. Hell and J. Wichmann. Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. *Optics letters*, 19(11):780–782, 1994.

[63] H. Hellmann. A new approximation method in the problem of many electrons. *Journal of Chemical Physics*, 3:61, 1935.

[64] S. T. Hess, T. P. Girirajan, and M. D. Mason. Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. *Biophysical journal*, 91(11):4258–4272, 2006.

[65] J. Heyd, G. E. Scuseria, and M. Ernzerhof. Hybrid functionals based on a screened coulomb potential. *The Journal of Chemical Physics*, 118(18):8207–8215, 2003.

[66] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Physical review*, 136(3B):B864, 1964.

[67] M. Holst, H. Hu, and Y. Zhu. A mathematical framework of models in calculating electronic structures of materials. *to be submitted*, 2014.

[68] H. Hu. Source code of "a systematic way to study stability conditions of models in calculating electronic structures of materials".

[69] H. Hu. Source code of "the linear scaling finite element solver for hartree-fock".

[70] H. Hu, C. Ma, and Z. Liu. Plasmonic dark field microscopy. *Applied Physics Letters*, 96(11):113107, 2010.

[71] H. Hu, D. Song, J. Weare, and M. Holst. A linear scaling finite element solver for hartree-fock. *to be submitted*, 2014.

[72] H. Hu, D. Song, J. Weare, and M. Holst. A systematic way to study stability conditions of models in calculating electronic structures of materials. *in preparation*, 2014.

[73] A. C. Hurley. *Electron correlation in small molecules*. Academic Press New York, 1976.

[74] S. Ivanov, R. Lopez-Boada, A. Görling, and M. Levy. Closed-form expression relating the second-order component of the density functional theory correlation energy to its functional derivative. *Journal of Chemical Physics*, 109:6280–6286, 1998.

[75] F. Jensen. *Introduction to computational chemistry*. Wiley. com, 2007.

[76] F. Jones. *Lebesgue integration on Euclidean space*. Jones & Bartlett Learning, 2001.

[77] R. A. Kendall, E. Aprà, D. E. Bernholdt, E. J. Bylaska, M. Dupuis, G. I. Fann, R. J. Harrison, J. Ju, J. A. Nichols, J. Nieplocha, et al. High performance computational chemistry: An overview of nwchem a distributed parallel application. *Computer Physics Communications*, 128(1):260–283, 2000.

[78] K. Kim and K. Jordan. Comparison of density functional and mp2 calculations on the water monomer and dimer. *The Journal of Physical Chemistry*, 98(40):10089–10094, 1994.

[79] T. A. Klar and S. W. Hell. Subdiffraction resolution in far-field fluorescence microscopy. *Optics letters*, 24(14):954–956, 1999.

[80] T. Koga, S. Watanabe, K. Kanayama, R. Yasuda, and A. J. Thakkar. Improved roothaan–hartree–fock wave functions for atoms and ions with n 54. *The Journal of chemical physics*, 103(8):3000–3005, 1995.

[81] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Physical Review*, 140(4A):A1133, 1965.

[82] G. Kresse and J. Furthmüller. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Physical Review B*, 54(16):11169, 1996.

[83] D. C. Langreth and M. Mehl. Beyond the local-density approximation in calculations of ground-state electronic properties. *Physical Review B*, 28(4):1809, 1983.

[84] B. Langwallner, C. Ortner, and E. Süli. Existence and convergence results for the galerkin approximation of an electronic density functional. *Mathematical Models and Methods in Applied Sciences*, 20(12):2237–2265, 2010.

[85] C. Le Bris. Computational chemistry from the perspective of numerical analysis. *Acta Numerica*, 14:363–444, 2005.

[86] C. Lee, W. Yang, and R. G. Parr. Development of the colle-salvetti correlation-energy formula into a functional of the electron density. *Physical Review B*, 37(2):785, 1988.

[87] L. Lehtovaara, V. Havu, and M. Puska. All-electron density functional theory and time-dependent density functional theory with high-order finite elements. *The Journal of chemical physics*, 131:054103, 2009.

[88] I. N. Levine. *Quantum chemistry*, volume 5. Prentice Hall Upper Saddle River, NJ, 2000.

[89] Z. H. Levine and J. W. Wilkins. An energy-minimizing mesh for the schrödinger equation. *Journal of Computational Physics*, 83(2):361–372, 1989.

[90] E. H. Lieb and B. Simon. The hartree-fock theory for coulomb systems. *Communications in Mathematical Physics*, 53(3):185–194, 1977.

[91] L. Lin, J. Lu, L. Ying, et al. Adaptive local basis set for kohn–sham density functional theory in a discontinuous galerkin framework i: Total energy calculation. *Journal of Computational Physics*, 231(4):2140–2154, 2012.

[92] P.-L. Lions. Solutions of hartree-fock equations for coulomb systems. *Communications in Mathematical Physics*, 109(1):33–97, 1987.

[93] Z. Liu, S. Durant, H. Lee, Y. Pikus, N. Fang, Y. Xiong, C. Sun, and X. Zhang. Far-field optical superlens. *Nano Letters*, 7(2):403–408, 2007.

[94] Z. Liu, H. Lee, Y. Xiong, C. Sun, and X. Zhang. Far-field optical hyperlens magnifying sub-diffraction-limited objects. *science*, 315(5819):1686–1686, 2007.

[95] N. H. March. Self-consistent fields in atoms. 1975.

[96] A. Masud and R. Kannan. B-splines and nurbs based finite element methods for kohn–sham equations. *Computer Methods in Applied Mechanics and Engineering*, 241:112–127, 2012.

[97] S. F. McCormick. *Multigrid methods*, volume 3. siam, 1987.

[98] R. McWeeny and B. T. Sutcliffe. *Methods of molecular quantum mechanics*, volume 2. Academic press London, 1989.

[99] A. Messiah. Quantum mechanics, vol. ii. *English Edition)(North Holland: Amster*, 1962.

[100] E. Miersemann. Calculus of variations i.

[101] A. V. Mitin. Exact solution of the hartree-fock equation for the h 2 molecule in the linear-combination-of-atomic-orbitals approximation. *Physical Review A*, 62(1):010501, 2000.

[102] N. Modine, G. Zumbach, and E. Kaxiras. Adaptive-coordinate real-space electronic-structure calculations for atoms, molecules, and solids. *Physical Review B*, 55(16):10289, 1997.

[103] C. Møller and M. S. Plesset. Note on an approximation treatment for many-electron systems. *Physical Review*, 46(7):618, 1934.

[104] P. Motamarri, M. Iyer, J. Knap, and V. Gavini. Higher-order adaptive finite-element methods for orbital-free density functional theory. *Journal of Computational Physics*, 231(20):6596–6621, 2012.

[105] P. Motamarri, M. R. Nowak, K. Leiter, J. Knap, and V. Gavini. Higher-order adaptive finite-element methods for kohn-sham density functional theory. *arXiv preprint arXiv:1207.0167*, 2012.

[106] L. Novotny and B. Hecht. *Principles of nano-optics*. Cambridge university press, 2012.

[107] J. A. O'Keefe. Resolving power of visible light. *Journal of the Optical Society of America*, 46(5):359, 1956.

[108] C. on Facilitating Interdisciplinary Research. *Facilitating Interdisciplinary Research*. The National Academies Press, 2004.

[109] F. A. Pahl and N. C. Handy. Plane waves and radial polynomials: a new mixed basis. *Molecular Physics*, 100(20):3199–3224, 2002.

[110] J. Paldus and J. Čížzek. Stability conditions for the solutions of the hartree–fock equations for atomic and molecular systems. ii. simple open-shell case. *The Journal of Chemical Physics*, 52(6):2919–2936, 1970.

[111] R. G. Parr. *The quantum theory of molecular electronic structure: a lecture-note and reprint volume*. WA Benjamin New York, 1963.

[112] R. G. Parr and W. Yang. *Density-functional theory of atoms and molecules*, volume 16. Oxford university press, 1989.

[113] J. Pask, B. Klein, C. Fong, and P. Sterne. Real-space local polynomial basis for solid-state electronic-structure calculations: A finite-element approach. *Physical Review B*, 59(19):12352, 1999.

[114] J. Pask, B. Klein, P. Sterne, and C. Fong. Finite-element methods in electronic-structure theory. *Computer physics communications*, 135(1):1–34, 2001.

[115] J. Pask and P. Sterne. Finite element methods in ab initio electronic structure calculations. *Modelling and Simulation in Materials Science and Engineering*, 13(3):R71, 2005.

[116] J. P. Perdew. Accurate density functional for the energy: Real-space cutoff of the gradient expansion for the exchange hole. *Physical review letters*, 55(16):1665, 1985.

[117] J. P. Perdew. Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Physical Review B*, 33(12):8822, 1986.

[118] J. P. Perdew, K. Burke, and M. Ernzerhof. Generalized gradient approximation made simple. *Physical review letters*, 77(18):3865, 1996.

[119] J. P. Perdew, J. Chevary, S. Vosko, K. A. Jackson, M. R. Pederson, D. Singh, and C. Fiolhais. Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation. *Physical Review B*, 46(11):6671, 1992.

[120] J. P. Perdew, M. Ernzerhof, and K. Burke. Rationale for mixing exact exchange with density functional approximations. *The Journal of Chemical Physics*, 105(22):9982–9985, 1996.

[121] J. P. Perdew, S. Kurth, A. Zupan, and P. Blaha. Accurate density functional with correct formal properties: A step beyond the generalized gradient approximation. *Physical Review Letters*, 82(12):2544, 1999.

[122] J. P. Perdew and Y. Wang. Accurate and simple analytic representation of the electron-gas correlation energy. *Physical Review B*, 45(23):13244, 1992.

[123] J. P. Perdew and W. Yue. Accurate and simple density functional for the electronic exchange energy: Generalized gradient approximation. *Physical Review B*, 33(12):8800, 1986.

[124] W. E. Pickett. Pseudopotential methods in condensed matter applications. *Computer Physics Reports*, 9(3):115–197, 1989.

[125] D. Pohl, W. Denk, and M. Lanz. Optical stethoscopy: Image recording with resolution $\lambda/20$. *Applied physics letters*, 44(7):651–653, 1984.

[126] S. Prince and W. McGuigan. Alignment and tolerancing of a cardioid condenser. In *Optical Engineering+ Applications*, pages 66760K–66760K. International Society for Optics and Photonics, 2007.

[127] E. Proynov, S. Sirois, and D. Salahub. Extension of the lap functional to include parallel spin correlation. *International Journal of Quantum Chemistry*, 64(4):427–446, 1997.

[128] R. Radovitzky and M. Ortiz. Error estimation and adaptive meshing in strongly nonlinear dynamic problems. *Computer Methods in Applied Mechanics and Engineering*, 172(1):203–240, 1999.

[129] C. C. J. Roothaan. New developments in molecular orbital theory. *Reviews of modern physics*, 23(2):69, 1951.

[130] M. J. Rust, M. Bates, and X. Zhuang. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (storm). *Nature methods*, 3(10):793–796, 2006.

[131] J. J. Sakurai. Modern quantum mechanics. *Reading, MA: Addison Wesley,—c1985, edited by Tuan, San Fu*, 1, 1985.

[132] V. Schauer and C. Linder. All-electron kohn-sham density functional theory on hierarchic finite element spaces. *Journal of Computational Physics*, 2013.

[133] E. Schwegler and M. Challacombe. Linear scaling computation of the hartree–fock exchange matrix. *The Journal of chemical physics*, 105(7):2726–2734, 1996.

[134] E. Schwegler and M. Challacombe. Linear scaling computation of the fock matrix. iv. multipole accelerated formation of the exchange matrix. *The Journal of chemical physics*, 111(14):6223–6229, 1999.

[135] E. Schwegler, M. Challacombe, and M. Head-Gordon. Linear scaling computation of the fock matrix. ii. rigorous bounds on exchange integrals and incremental fock build. *The Journal of chemical physics*, 106(23):9708–9717, 1997.

[136] R. Seeger and J. A. Pople. Self-consistent molecular orbital methods. xviii. constraints and stability in hartree–fock theory. *The Journal of Chemical Physics*, 66(7):3045–3050, 1977.

[137] M. Segall, P. J. Lindan, M. Probert, C. Pickard, P. Hasnip, S. Clark, and M. Payne. First-principles simulation: ideas, illustrations and the castep code. *Journal of Physics: Condensed Matter*, 14(11):2717, 2002.

[138] I. Shavitt and R. J. Bartlett. *Many-body methods in chemistry and physics: MBPT and coupled-cluster theory*. Cambridge University Press, 2009.

[139] C. D. Sherrill and H. F. Schaefer III. The configuration interaction method: Advances in highly correlated approaches. *Advances in quantum chemistry*, 34:143–269, 1999.

[140] J. Silverman, O. Platas, and F. Matsen. Simple configuration-interaction wave functions. i. two-electron ions: A numerical study. *The Journal of Chemical Physics*, 32(5):1402–1406, 1960.

[141] C.-K. Skylaris, P. D. Haynes, A. A. Mostofi, and M. C. Payne. Introducing onetep: Linear-scaling density functional simulations on parallel computers. *The Journal of chemical physics*, 122:084119, 2005.

[142] I. I. Smolyaninov, Y.-J. Hung, and C. C. Davis. Magnifying superlens in the visible frequency range. *Science*, 315(5819):1699–1701, 2007.

[143] J. M. Soler, E. Artacho, J. D. Gale, A. García, J. Junquera, P. Ordejón, and D. Sánchez-Portal. The siesta method for ab initio order-n materials simulation. *Journal of Physics: Condensed Matter*, 14(11):2745, 2002.

[144] I. Stakgold and M. J. Holst. *Green's functions and boundary value problems*, volume 99. Wiley. com, 2011.

[145] P. Stephens, F. Devlin, C. Chabalowski, and M. J. Frisch. Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields. *The Journal of Physical Chemistry*, 98(45):11623–11627, 1994.

[146] P. Suryanarayana, V. Gavini, T. Blesgen, K. Bhattacharya, and M. Ortiz. Non-periodic finite-element formulation of kohn–sham density functional theory. *Journal of the Mechanics and Physics of Solids*, 58(2):256–280, 2010.

[147] E. Synge. Xxxviii. a suggested method for extending microscopic resolution into the ultra-microscopic region. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 6(35):356–362, 1928.

[148] A. Szabo and N. S. Ostlund. *Modern quantum chemistry: introduction to advanced electronic structure theory*. DoverPublications. com, 1989.

[149] J. Tao, J. P. Perdew, V. N. Staroverov, and G. E. Scuseria. Climbing the density functional ladder: Nonempirical meta–generalized gradient approximation designed for molecules and solids. *Physical Review Letters*, 91(14):146401, 2003.

[150] L. H. Thomas. The calculation of atomic fields. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 23, pages 542–548. Cambridge Univ Press, 1927.

[151] E. Tsuchida and M. Tsukada. Electronic-structure calculations based on the finite-element method. *Physical Review B*, 52(8):5573, 1995.

[152] E. Tsuchida and M. Tsukada. Adaptive finite-element method for electronic-structure calculations. *Physical Review B*, 54(11):7602, 1996.

[153] E. Tsuchida and M. Tsukada. Large-scale electronic-structure calculations based on the adaptive finite-element method. *Journal of the Physics Society Japan*, 67(11):3844–3858, 1998.

[154] M. Valiev, E. J. Bylaska, N. Govind, K. Kowalski, T. P. Straatsma, H. J. Van Dam, D. Wang, J. Nieplocha, E. Apra, T. L. Windus, et al. Nwchem: a comprehensive and scalable open-source solution for large scale molecular simulations. *Computer Physics Communications*, 181(9):1477–1489, 2010.

[155] B. Van Brunt. *The calculus of variations*, volume 1. Springer, 2004.

[156] T. Van Voorhis and G. E. Scuseria. A novel form for the exchange-correlation energy functional. *The Journal of chemical physics*, 109:400, 1998.

[157] S. H. Vosko, L. Wilk, and M. Nusair. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Canadian Journal of Physics*, 58(8):1200–1211, 1980.

[158] W. Weber and C. Eagen. Energy transfer from an excited dye molecule to the surface plasmons of an adjacent metal. *Optics Letters*, 4(8):236–238, 1979.

[159] M. Weissbluth. *Atoms and molecules*. Elsevier, 1978.

[160] S. R. White, J. W. Wilkins, and M. P. Teter. Finite-element method for electronic structure. *Physical Review B*, 39(9):5819, 1989.

[161] J. M. Wills and B. R. Cooper. Synthesis of band and model hamiltonian theory for hybridizing cerium systems. *Physical Review B*, 36(7):3809, 1987.

[162] S. Wilson. *Electron correlation in molecules*. Clarendon Press Oxford, 1984.

[163] D. Wu, Z. Liu, C. Sun, and X. Zhang. Super-resolution imaging by random adsorbed molecule probes. *Nano letters*, 8(4):1159–1162, 2008.

[164] T. Yanai, G. I. Fann, Z. Gan, R. J. Harrison, and G. Beylkin. Multiresolution quantum chemistry in multiwavelet bases: Hartree–fock exchange. *The Journal of chemical physics*, 121(14):6680–6688, 2004.

[165] J. Zeng. Quantum mechanics, vol. ii. *Science Press, Beijing*, 2000.

[166] D. Zhang, L. Shen, A. Zhou, and X.-G. Gong. Finite element method for solving kohn–sham equations based on self-adaptive tetrahedral mesh. *Physics Letters A*, 372(30):5071–5076, 2008.