FAST TREATMENT OF π-STACKING USING DENSITY FUNCTIONAL THEORY.

by

A. B. Sharapov

B.S. Saint-Petersburg State University, 2005

M.S. Saint-Petersburg State University, 2008

M.S. University of Pittsburgh, 2008

M.S. University of Pittsburgh, 2009

M.S. University of Pittsburgh, 2012

Submitted to the Graduate Faculty of

the Department of Physics and Astronomy in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2013

UNIVERSITY OF PITTSBURGH

DEPARTMENT OF PHYSICS AND ASTRONOMY

This dissertation was presented

by

A. B. Sharapov

It was defended on

April 17, 2013

and approved by

Geoffrey Hutchison, Assistant Professor, Department of Chemistry

David Jasnow, Professor, Department of Physics and Astronomy

Adam Leibovich, Associate Professor,Department of Physics and Astronomy

Gurudev Dutt, Assistant Professor, Department of Physics and Astronomy

Dissertation Advisors: Geoffrey Hutchison, Assistant Professor, Department of Chemistry,

Anthony Duncan, Professor, Department of Physics and Astronomy

FAST TREATMENT OF π-STACKING USING DENSITY FUNCTIONAL THEORY.

A. B. Sharapov, PhD

University of Pittsburgh, 2013

In this work, we concentrate on various unconventional Density Functional Theory approaches for calculations of electronic structure of molecules. In particular, we are motivated by gaining substantial speed-up of such calculations. Firstly, explore the Orbital-Free Density Functional Theory and show how one can compute the energies of dimers by performing large-scale optimization with more than a million variables. We also show that this approach does not give satisfactory results. Secondly, we investigate tight-binding methods and in particular show that Harris approximation gives good results when applied to calculations of energies of dimers. In the last section, we apply Harris approximation to molecular packing optimization,and show that the SPSA algorithm is capable of finding all minima of the energy surface.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1.0 INTRODUCTION.

The desire to predict the crystal structure of molecular compounds from the molecular geometry alone has triggered a substantial amount of research in the past and is motivated by the wide range of potential scientific and technological applications. Most solid-state properties such as density, hardness, color, morphology, or solubility crucially depend on the crystal packing. The ability to predict crystal structures would open up the door to the computation of solid-state properties with nothing more than the molecular geometry as a starting point, thus making it possible to select molecules for certain solid-state applications before they are even synthesized. The state of the art in the field of polymorph prediction has been assessed by a series of blind tests organized by the Cambridge Crystallographic Data Center [1, 2, 3, 4, 5, 6]. The blind tests reveal a fundamental shortcoming of all current approaches, which is the inability to calculate lattice energy differences with an accuracy that is higher than or at least comes close to typical energy differences of about 0.01-0.1 kcal/mol found between the low-energy crystal structures of a given molecule. Even though crystallization kinetics may complicate things further, it can be expected that in most cases only the stable polymorph and some of the meta-stable polymorphs with slightly higher lattice energies can actually be crystallized.

As a consequence, polymorph prediction is bound to fail if the energy ranking is too inaccurate to allow for the reliable identification of the most stable crystal structures. Current methods used for lattice energy calculation can be roughly divided into empirical, semiempirical, and *ab initio* approaches. Empirical approaches use force fields which consist of a set of functional forms and parameters that have been fitted to experimental data and/or high-level *ab initio* calculations. If well parameterized, force fields can be appropriate for polymorph prediction [8, 9], but in general sufficiently accurate parameters are not readily

available for most molecules. In contrast to force fields, *ab initio* calculations do not need to be parameterized, but involve the numerical solution of Schrodinger's equation for the electronic motion, thus requiring significantly more CPU time and computer memory. Since the exact solution of Schrodinger's equation is impossible for all but the simplest cases, various approximate techniques have been developed of which density functional theory (DFT) calculations are of particular interest, as they offer a good compromise between speed and accuracy for crystal structures with cell volumes of $100 - 1000$ Å^3 and above. Because DFT calculations make use of the independent electron approximation, they fail to incorporate long-range dispersive interactions (van der Waals interactions) which result from electron correlation effects and play an important role in molecular crystals. Accordingly, pure DFT calculations are not appropriate for the structure optimization of molecular crystals.

Several attempts have been reported to overcome this problem. Some authors have tried to incorporate van der Waals interactions in the DFT formalism [10, 11, 12, 13]. Although promising, these approaches either are currently too time-consuming for practical applications or suffer from other limitations. A more pragmatic approach, already explored by several authors [12, 13] and also used in this work, is to combine DFT calculations with an empirical van der Waals correction. The empirical correction is defined as the sum over atom-atom pair potentials with each pair potential for two atoms A and B being the product of an asymptotic $C_{6,A,B}/r^6$ term and a damping function $d_{A,B}(r)$. The damping function is required to counteract the divergence of the $C_{6,A,B}/r^6$ term at short interatomic distances r. At large interatomic distances, the damping function is equal to 1. This approach has already been reviewed in many details by Marcus [14].

Although dispersion-corrected DFT(DFT+D) or hybrid approaches are quite common nowadays, they can be computationally expensive. The running time of such methods can be days and sometimes weeks of calculations on multiple cores. Indeed, these methods have a decent accuracy but may take very long run-times. On the other side, we have force fields methods which are based on certain parametrization of the underlying equations. For example, one can model bonds as springs and try to find the coefficient of proportionality by means of the available experimental data. This is an appealing approach and indeed, it is very fast in the computational sense. The only problem is that it is not *ab initio* and

the accuracy suffers considerably when applied outside the range of the parametrization, especially with electrostatic interactions. The overall conclusion that can be drawn from the above reasoning is that there is a lack of an *ab initio* method which is computationally fast and reasonably accurate compared to the high-level methods. Such a method can be tentatively applied not only for geometry optimization by also for a quick scan over thousands of molecules in order to study general trends.

The goal of the current project was to find a new electronic structure calculation method which is reasonably accurate and at the same time possesses good computational performance. We would like to apply such a method for geometry optimization purposes and be able to find the global minimum of a system consisting of a number of molecules. As we have mentioned above, the field of the electronic structure calculations can be subdivided into two extremes. On one side, we have fast but questionable methods based in empirical force field. These methods are computationally very fast (geometry optimization takes time on the order of minutes) but their accuracy wishes much to be desired. Placing atoms in different environments can have unpredictable effects on the accuracy. On the other side, we have very accurate but at the same time very slow high-level wave functional based methods. One can qualify them as the state-of-the-art modern approaches. Unfortunately, the applicability of such methods to geometry optimization may seem infeasible for the reason of their low computational performance. As an example, the conventional Kohn-Sham Density Functional Theory calculation scales cubically with the number of atoms (N), MP2 method scales as $O(N^5)$, and some of the more advanced method have an even higher complexity. This simple analysis shows that in order to run the geometry optimization using high-level methods, one will have to wait for days, if not weeks, in order to obtain results.

In the work described in this thesis, we present a method which is reasonably accurate (i.e. the method gives results close to those of high-level, computationally intensive methods) and relatively fast. Moreover, we are mostly concerned with geometry optimization.For this reason, we quantify the desired accuracy of our method in terms of the proximity of the binding curves obtained using our method to the binding curves obtained by means of high-level methods as well as the proximity of locations of the minima of those binding curves. In terms of computational performance, we would like to obtain the optimized geometry within

several hours. Since our method is an approximation, the overall strategy for any geometry optimization should be divided into two parts. Firstly, we run our method in order to find an approximate location of the minimum. Secondly, one can now run a high-level method in order to adjust the geometry.

In chapter 2 of this work, we investigate the performance and the accuracy of the Orbital-Free Density Functional Theory. This is a very attractive method of the electronic structure calculation because its complexity scales linearly with the number of atoms as opposed to the cubic scaling of the conventional Kohn-Sham DFT. This scaling is achieved by completely eliminating orbitals from the calculation. Instead, the electron density is discretized in a rectangular grid and the total energy functional is optimized with respect to the values of the electron density on each site. The total energy consists of five major parts:

$$E_{tot}[\rho] = T_s[\rho] + E_{ext}[\rho] + E_{hart}[\rho] + E_{xc}[\rho] + E_{NN} \tag{1.1}$$

The electron kinetic energy $T_s[\rho]$ approximated by the sum of the Thomas-Fermi and von Weizsacker terms; the electron-nucleus interaction $E_{ext}[\rho]$;the Hartree term $E_{hart}[\rho]$;the nucleus-nucleus interaction E_{NN}, and the exchange-correlation term $E_{xc}[\rho]$. The major problem of the OF-DFT approach is related to the kinetic energy approximation used: the Thomas-Fermi term is only valid for a homogeneous electron gas and the von Weizsacker term is only exact for systems with no more than two electrons. Other, non-local terms have been recently proposed [62],[51]. We make use of the periodic boundary conditions in order to reduce the computational complexity of the Hartree term from $O(M^2)$ to $O(M \log M)$, where M is the number of grid points. At the same time, the imposition of the periodic boundary conditions will require us to have a very big simulation region in order to prevent the electrostatic images of molecules from feeling each other. This will cause the number of grid point to rise and so will prevent us from using the non-local kinetic energy terms, which otherwise can make the calculation computationally infeasible. This leaves us just two contributions to the kinetic energy of electrons: Thomas-Fermi and von Weizsacker terms:

$$T_s[\rho] = C_{TF} \int \rho(\mathbf{r})^{5/3} d\mathbf{r} + \frac{\lambda}{8} \int \frac{|\nabla \rho(\mathbf{r})|^2}{\rho(r)} d\mathbf{r} \tag{1.2}$$

Following [93], we investigate the performance of the model

$$T_s[\rho] = a \times C_{TF} \int \rho(\mathbf{r})^{5/3} d\mathbf{r} + b \times \frac{1}{8} \int \frac{|\nabla \rho(\mathbf{r})|^2}{\rho(r)} d\mathbf{r} \tag{1.3}$$

to the benzene and thiophene dimers, where a and b are parameters taking values $a = 1.0, 1.1, 1.2$ and $b = \frac{1}{3}, \frac{1}{6}, \frac{1}{9}$. The introduction of these parameters may compensate for the lack of non-local terms and approximate the binding curves up the a reasonable accuracy. We compute the binding curves for the benzene and thiophene dimers and show that one can find a pair of parameters such that the resulting binding curve is able to reasonably approximate the benchmark results obtained by means of M06-2X functional. Unfortunately, we do not find any transferability of any given model from molecule to molecule: the optimal parameters vary in an unpredictable way as the molecular structure changes. We can summarize the results as follows:

- Results show no transferability.
- Certain trends can be recovered. As $a \uparrow$ or $b \uparrow$ the minima occur at longer distances.
- Further investigation is needed to access the applicability of OF-DFT to molecular systems.
- Minima occur quite far from the benchmark results.
- We do not intend to further apply this method

In chapter 3, we take a completely different approach. It is motivated by the Harris Functional Approximation [28] of the Kohn-Sham DFT. The conventional way of solving Kohn-Sham equations is called the Self Consistent Field (SCF). The equations are non-linear, so the only way of finding a solution is by using an iterative approach. The Harris approximation is the first order expansion of the Kohn-Sham energy around the electron density of non-interacting fragments and so it avoids iterations completely, increasing performance by a double digit factor. Historically, the Harris approximation has used atoms as fragments.In the current work, we proposed to use molecules as fragments and compute the Harris energy based on that assumption. The molecular electron density can be easily pre-computed using the standard SCF technique and then reused multiple times. We also proposed to augment the Harris energy by an empirical correction [15]. We investigate

the performance of the augmented Harris approximation with LDA (local density approximation) and GGA (generalized gradient approximation) and compare the results to two benchmark high-level methods - M06-2X and MP2. We explore the accuracy of the Harris based methods by computing the biding curves of the benzene and thiophene dimer and comparing them to the benchmark curves obtained by MO6-2X [21] and MP2 [26]. The results can be summarized as follows:

- Each energy curve obtained with Harris + Dispersion approximation gives binding.
- LDA Harris + Dispersion performs somewhat worse than GGA Harris + Dispersion in comparison to M06-2X and MP2.
- GGA Harris + Dispersion method performs very well compared to M06-2X and MP2. Both methods give minima which coincide up to one tenth of an angstrom (may be better).Shapes of binding curves are almost identical.
- The method can be accelerated using openMP and gives a substantial speed-up over the full SCF.
- The method is suitable for geometry optimization but details remain to be tested.

The results show that Harris+GGA+D method compares very well to the results of the benchmark method, with the results obtained at a fraction of the cost. This suggests that this method can be successfully applied to geometry optimization.

In chapter 4, we apply the Harris+GGA+D method to the geometry optimization of the benzene dimer. The drawback of the Harris energy expression is the lack of the explicit formulae for the gradient of the energy with respect to the nuclei positions. One can approximate the gradient by means of the finite central differences. Although possible in theory, such approximation quickly becomes computationally intractable, since for each degree of freedom one has to perform two energy evaluations. Unlike in a conventional approach where all nuclei positions are optimized, we constraint our system to only 6 degrees of freedom.In the present situation we optimize 3 translational coordinates associated with the centroid of a molecules and 3 rotational coordinates associated with rotations around that centroid, yielding 6 degrees of freedom and so 12 energy evaluations per iteration. This makes it computationally very expensive. Instead, we propose to use the Simultaneous Perturbation

Stochastic Approximation (SPSA) algorithm [112]. It needs only 2 energy evaluations per iteration and is very attractive in the present context. Furthermore, this algorithm is a global optimizer [113],which will help us find the configuration with the minimal energy.

The most important component of the algorithm is the gain sequence which controls the magnitude of a step toward the minimum at each iteration. We perform a very careful investigation of the choice of the gain sequence and show that different values may yield different final configurations. In particular, we show that one choice of gain sequence makes the algorithm find two minima with very close energy values: starting sandwiched configuration becomes a t-shaped final configuration and starting parallel-displaced configuration stays one but at a different distance. Another choice of the gain sequence makes the algorithm converge to the final t-shaped configuration independently of the choice of the initial orientation. The algorithm provide a certain level of flexibility giving researchers a way of finding either the overall global minimum (in this case t-shaped configuration) or converging to a local minimum. The result show that SPSA algorithm finds the global minimum of the benzene dimer and is also able to locate the local minima, if needed. GGA Harris + Dispersion method may indeed be used as a starter for geometry optimization. Current implementation scales almost linearly with the number of cores and provides a double-digit speedup over the conventional full SCF methods.The method is efficient and accurate enough to be able to predict the molecular packing of the benzene dimer. The results can be summarized as follows:

- GGA Harris + Dispersion method may indeed be used as a starter for geometry optimization.
- Current implementation scales almost linearly with the number of cores and provides a double-digit speedup over the conventional full SCF methods.
- The method is efficient and accurate enough to be able to predict the molecular packing of the benzene dimer.
- Current implementation is able to find all minima of the energy surface in multiple dimensions.

2.0 ORBITAL-FREE DENSITY FUNCTIONAL THEORY AND ITS APPLICATION TO MODELING OF π-STACKING.

In this section, we explain the orbital free Density Functional Theory (OF-DFT) approach to modeling of molecular systems. We start by introducing the reader to the basic concepts of this methodology. We derive the expressions for some basic energy functionals and give the motivation for why it is advantageous to use this scheme for real calculations. We also pinpoint various drawbacks and possible solutions which have been recently proposed. Next, we apply the OF-DFT approach to modeling of π-stacking in molecules such as benzene and thiophene. We explain our computational approach and point out its advantages and disadvantages. Finally, we present the results of our calculations.

2.1 APPLICATION OF ORBITAL FREE DENSITY FUNCTIONAL THEORY TO MODELING OF COVALENT INTERACTIONS.

OF-DFT is a first principle quantum mechanics method that can be formulated to scale linearly with system size. By contrast, Kohn-Sham DFT, which will be discussed in the next chapter, scales cubically with size initially and then can be made to scale linearly asymptotically. Currently, OF-DFT can be used to study samples consisting of tens of thousands of atoms on a single processor and hundreds of thousands of atoms on tens of processors if parallel implementation is used.

OF-DFT eliminates the Kohn-Sham orbitals and instead relies on approximation to the kinetic energy that depends explicitly in the electron density. The viability of OF-DFT thus depends heavily on accuracy of the kinetic energy functional chosen. Due to limitation

in the accuracy of state-of-the-art kinetic limitations in the accuracy of the kinetic energy functionals, at present OF-DFT is only accurate as KS-DFT for main group metals, as well as for some properties of semiconductors.

Currently, the best approximations available are those that explicitly account for the exact linear response of a uniform electron gas density subject to small perturbations in the potential. We gave an example of such a functional which was pioneered by Wang and Teter [51], modified by Perrot [52], and generalized by Wang, Govind and Carter (WGC) [58].

Unfortunately, the latter functionals suffer from a series of drawbacks when applied to molecules. The Wang and Teter [51] functional is purely nonlocal which leads to quadratic scaling $O(M^2)$, where M is the discretization of the integration grid. In practice, it is possible to use the periodic boundary conditions and reduce the complexity of the calculation of the Coulomb term from $O(M^2)$ to $O(M \log M)$. This is not possible for the Wang and Teter functional [51]. Wang, Govind and Carter [58] were able to generalize the Wang and Teter functional, and by means of Taylor expansion around the average electron density they were able to reduce the complexity to $O(M \log M)$. The main issue here is that it is hard to define an average electron density for a complex molecule.

2.1.1 Orbital-free energy functional.

We now examine each term in the OF-DFT individually. The OF-DFT total energy can be written as

$$E_{tot}[\rho] = T_s[\rho] + E_{xc}[\rho] + E_{ext}[\rho] + E_{hart}[\rho] + E_{ii}[\mathbf{r}] \tag{2.1}$$

where $T_s[\rho]$ is the electronic kinetic energy, $E_{ext}[\rho]$ is the electron-ion potential energy, $E_{hart}[\rho]$ is the Coulomb repulsion energy between electrons, $E_{xc}[\rho]$ is the exchange-correlation energy, and $E_{ii}[\mathbf{r}]$ is the ion-ion repulsion energy. In our calculation we decided to place all the nuclei on the grid. This is one of the reason why our grid should be very fine. Taking that into account we will simply combine the last three terms in our energy expression into one and call it $J[\rho]$.

In the following we present individual expressions for the functionals and the methods of their calculation.

2.1.2 Coulomb term.

The Coulomb energy for the charge density including nuclei at a point $\mathbf{r}$ interacting with the electron density at another point $\mathbf{r}^{'}$ is expressed as

$$J[\rho] = \frac{1}{2}\int\int \frac{\rho(\mathbf{r})\rho(\mathbf{r}^{'})}{|\mathbf{r}-\mathbf{r}^{'}|}d\mathbf{r}d\mathbf{r}^{'} \tag{2.2}$$

Under the periodic boundary conditions, it is possible to perform this convolution in reciprocal space with $O(M \ln M)$ scaling to obtain the Coulomb energy as

$$J[\rho] = \frac{V}{2}\sum_{\mathbf{g}\neq 0}\frac{4\pi}{|\mathbf{g}|^2}\rho(\mathbf{g})\rho(-\mathbf{g}) \tag{2.3}$$

where V is the volume of the periodic cell and $\rho(\mathbf{g})$ is obtained as $\rho(\mathbf{g}) = \hat{F}(\rho(\mathbf{r}))$, where $\hat{F}$ is a forward fast Fourier transform, defined as

$$\hat{F}(f(\mathbf{r})) = f(\mathbf{g}) = \frac{1}{N}\sum_{\mathbf{r}} f(\mathbf{r})e^{-i\mathbf{g}\cdot\mathbf{r}} \tag{2.4}$$

The potential in real space is computed as

$$\frac{\partial J[\rho]}{\partial \rho}(\mathbf{r}) = \hat{F}^{'}\left(\frac{4\pi}{|\mathbf{g}|^2}\rho(\mathbf{g})\right) \tag{2.5}$$

where $\hat{F}^{'}$denotes the reverse fast Fourier transform

$$\hat{F}^{'}(f(\mathbf{g})) = f(\mathbf{r}) = \sum_{\mathbf{g}} f(\mathbf{g})e^{i\mathbf{g}\cdot\mathbf{r}} \tag{2.6}$$

2.1.3 Thomas-Fermi kinetic energy density functional.

The Thomas-Fermi kinetic energy functional [64, 65] is the simplest one available. It is completely local, and depends only on the density. It is the LDA for the kinetic energy. It has the form

$$T_{TF}[\rho] = C_{TF} \int \rho(\mathbf{r})^{5/3} d\mathbf{r} \tag{2.7}$$

However, this functional by itself predicts no shell structure for atoms and no molecular binding whatsoever [82, 83], and therefore is clearly inadequate on it own for real materials.

The Thomas-Fermi potential is given by

$$\frac{\partial T_{TF}[\rho]}{\partial \rho(\mathbf{r})} = \frac{5}{3} C_{TF} \rho(\mathbf{r})^{2/3} \tag{2.8}$$

2.1.4 von Weizsacker kinetic energy density functional.

The von Weizsacker (vW) kinetic energy functional [84] is exact for any single orbital system, i.e. for up to two singlet-coupled fermions or any number of bosons. The vW functional by itself is a lower bound to the true kinetic energy, since it neglects spin. Imposition of the Pauli Principle for more than two electrons introduces nodes in the many-body wave function, which increases the kinetic energy. Additionally, the functional $T_s[\rho] = T_{TF}[\rho] + T_{vW}[\rho]$ has been shown to be an upper bound to the true kinetic energy for non-interacting particles in one dimension [85, 86].

The vW functional has the form

$$T_{vW}[\rho] = \int \sqrt{\rho(\mathbf{r})} \left(-\frac{1}{2} \nabla^2 \right) \sqrt{\rho(\mathbf{r})} d\mathbf{r} \tag{2.9}$$

This integral is simply the standard Hamiltonian form of the kinetic energy when the wave function is precisely the square root of the electron density. This functional also can be written as [87]

$$T_{vW}[\rho] = \frac{1}{8} \int \frac{|\nabla \rho(\mathbf{r})|^2}{\rho(\mathbf{r})} d\mathbf{r} \tag{2.10}$$

This can be proved by the chain rule:

$$T_{vW}[\rho] = \int \sqrt{\rho(\mathbf{r})}\left(-\frac{1}{2}\nabla^2\right)\sqrt{\rho(\mathbf{r})}d\mathbf{r} = \int -\frac{1}{2}\sqrt{\rho(\mathbf{r})}\nabla(\nabla\sqrt{\rho(\mathbf{r})})d\mathbf{r} = \tag{2.11}$$

$$= \int -\frac{1}{2}\sqrt{\rho(\mathbf{r})}\nabla\left(\frac{\nabla\rho(\mathbf{r})}{2\sqrt{\rho(\mathbf{r})}}\right)d\mathbf{r} = \int -\frac{1}{4}\sqrt{\rho(\mathbf{r})}\nabla\left(\frac{\nabla\rho(\mathbf{r})}{2\sqrt{\rho(\mathbf{r})}}\right)d\mathbf{r} =$$

$$= \int -\frac{1}{4}\sqrt{\rho(\mathbf{r})}\left(\frac{\sqrt{\rho(\mathbf{r})}\nabla^2\rho(\mathbf{r}) - \frac{\nabla\rho(\mathbf{r})\cdot\nabla\rho(\mathbf{r}))}{2\sqrt{\rho(\mathbf{r})}}}{\rho(\mathbf{r})}\right)d\mathbf{r} = \int\left(-\frac{1}{4}\nabla^2\rho(\mathbf{r}) + \frac{1}{8}\frac{|\nabla\rho(\mathbf{r})|^2}{\rho(\mathbf{r})}\right)d\mathbf{r}$$

Since the value of $\int -\frac{1}{4}\nabla^2\rho(\mathbf{r})d\mathbf{r}$ is zero under the periodic boundary conditions, we recover formula (2.10) for a periodic system. The expression (2.9) for the vW kinetic energy is favored because it gives more stable convergence to the minimum. Expression (2.10) becomes ill-behaved in the regions far from the nuclei when the electron density is very small.

The functional derivative with respect to ρ of the vW kinetic energy functional is given by

$$\frac{\partial T_{vW}[\rho]}{\partial\rho(r)} = -\frac{1}{4}\frac{\nabla^2\rho}{\rho} + \frac{1}{8}\frac{\nabla\rho\cdot\nabla\rho}{\rho^2} = -\frac{1}{2}\frac{\nabla^2\sqrt{\rho}}{\sqrt{\rho}} \tag{2.12}$$

When using an optimization algorithm, it was found that keeping the electron density positive is a very challenging problem unless we introduce a new variable $\chi = \sqrt{\rho}$ and minimize the total energy with respect to χ [88, 89, 90]. Using this variable transformation, the first expression of the vW functional becomes

$$T_{vW}[\chi] = \int \chi(\mathbf{r})\left(-\frac{1}{2}\nabla^2\right)\chi(\mathbf{r})d\mathbf{r} \tag{2.13}$$

and the functional derivative becomes

$$\frac{\partial T_{vW}}{\partial\chi(\mathbf{r})} = -\nabla^2\chi(\mathbf{r}) \tag{2.14}$$

2.1.5 Exchange-correlation functionals.

The expression for exchange-correlation functionals have already been presented above in abundance. For the present work, we do not derive any expressions for the gradient of the exchange-correlation functionals either. We prefer to use the well-established library of functionals libxc [72].

2.1.6 Optimization method for OF-DFT.

The second Hohenberg-Kohn [50] theorem states that for a trial density $\rho(r)$, in which $\rho(r) \geq 0$ at all point and $\int \rho(r)dr = N_e$, with N_e the number of electrons, $E[\tilde{\rho}(t)] \geq E_{ground}$, where E_{ground} is the exact ground state energy.

Thus, given a configuration of ions that create an external potential and a trial electron density $\tilde{\rho}(r)$, we can recover the ground state electron density and energy by minimizing the total energy of the system with respect to the electron density, subject to the two constraints that the number of electrons be conserved and that the density be nonnegative everywhere. More simply, we wish to solve the nonlinear, multidimensional problem

$$E_{ground}^{OF-DFT} = \min \left\{ E^{OF-DFT}[\rho] - \lambda \left(\int \rho(r)dr - N_e \right) ; \rho \geq 0 \right\} \tag{2.15}$$

where λ is a Lagrange multiplier used to enforce the constraint that the total number of electrons is conserved.

A variety of technique are available for minimizing (2.15). Ideally, an algorithm must be robust, fast, and use $O(M)$ amount of memory. After trying a number of algorithms we have focused on the augmented lagrangian method coupled with low storage BFGS optimizer [75, 76, 77, 78].

A brief description of these algorithms is given below.

The problem we are trying to solve is generally expressed in the following form:

$$\min_{x \in \mathbb{R}^N} f(x) \tag{2.16}$$

subject to

$$c(x) = 0$$

To solve this problem we compose the augmented Lagrangian function:

$$\Phi(x, u, \mu) = f(x) - u^T c(x) + \frac{1}{2\mu} ||c(x)||_2^2 \tag{2.17}$$

where u and μ are auxiliary parameters. By making such a transformation, we perform the so-called convexification of the Lagrangian function. The goal is to adjust u and μ to encourage convergence.

The algorithm usually does the following:

Given $\mu_0 > 0$ and u_0, set $k = 0$.

Until convergence iterate:

- Starting from x_k^s, use an unconstrained minimization algorithm (in our case BFGS) to find an approximate minimizer x_k of $\Phi(x, u_k, \mu_k)$ for which $||\nabla_x \Phi(x_k, u_k, \mu_k)|| \leq \epsilon_k$.
- If $||c(x_k)|| \leq \eta_k$, set $u_{k+1} = u_k - \frac{c(x_k)}{\mu_k}$ and $\mu_{k+1} = \mu_k$
- Otherwise set $u_{k+1} = u_k$ and $\mu_{k+1} \leq \tau \mu_k$
- Set suitable ϵ_{k+1} and η_{k+1} and increase k by 1.

2.1.6.1 BFGS Method Given starting point x_0, convergence tolerance $\epsilon > 0$ inverse Hessian approximation H_0;

- set $k \leftarrow 0$

while $||\nabla f_k|| > \epsilon$; Compute search direction

$$p_k = -H_k \nabla f_k \tag{2.18}$$

- set $x_{k+1} = x_k + \alpha_k p_k$ where α_k is computed from a line search procedure to satisfy the Wolfe conditions;
- Define $s_k = x_{k+1} - x_k$ and $y_k = \nabla f_{k+1} - \nabla f_k$;
- Compute H_{k+1} by means of (2.19) and set $k \leftarrow k + 1$;

$$H_{k+1} = (I - \rho_k s_k y_k^T) H_k (I - \rho_k y_k s_k^T) + \rho_k s_k s_k^T \tag{2.19}$$

where

$$\rho_k = \frac{1}{y_k^T s_k} \tag{2.20}$$

end(while)

2.1.7 Results and discussions.

In the previous section we have attempted to outline some major components of our OF-DFT code/approach. We have specified individual pieces of the total energy and ways to numerically approximate them. In the following, we further explain some of the finer details of our approach and give the results of our simulation. In particular, we choose to study the sandwiched configuration of benzene and thiophene dimers. This can give us an idea of how well those systems are explained by the OF-DFT approach.

In order to evaluate the above formula for the total energy of the system, we discretize the space onto a fine grid of points and approximate the integrals by finite sums. In particular, we impose periodic boundary conditions in order to be able to use Fast Fourier Transform routines implemented by fftw3 library [94] to evaluate the Coulomb energy. One possible drawback of imposing the periodic boundary conditions is the fact that we have to make our simulation region very big in order to prevent images of dimers from feeling each other as indicated in Fig.2.1. In particular, in case of the benzene dimer we use a simulation box of the size of $18 \times 18 \times 18$ Å and the lattice of $113 \times 113 \times 113$ with the lattice spacing of about 0.16Å. For the case of the thiophene dimer we use a simulation region of $20 \times 20 \times 20$ Å and the lattice of $100 \times 100 \times 100$ with the lattice spacing of about 0.2Å. These parameters can be changed by the user. In fact, the finer discretization is more favorable in this case. One should remember though that such grids produce optimization problems with more than one million of variables. For example, in the present situation in case of the benzene dimer with the monomers at the distance of 3.5Å, we can get an optimization problem with 1.5 million variables. We attempt to place the nuclei on the lattice sites here. By placing the nuclei on the grid we avoid the usage of local pseudo potentials and approximate techniques for evaluation of Coulomb energy in periodic systems such as Ewald sums [73], etc. On the other hand, by shifting the nuclei we introduce unrealistic perturbations into the system. We can only hope that those perturbations are negligible. In the later section, we will discuss alternative ways to evaluate ion-ion and ion-electron Coulomb energy.

One important ingredient of the optimization routine for the electron density which we have evaded discussing so far is the initial guess. This is a very important piece of the

puzzle since gradient based optimization routines are intrinsically local which allows the possibility to converge to a wrong local minimum. Another point of concern is the speed of convergence. In other words, starting from a sensible initial guess can significantly reduce the amount of time needed to get close to the minimum. In the current study we are primarily interested in the calculation of binding curves of molecules such as benzene, thiophene,etc. We consider only two molecules for the current simulation. We suggest using the initial guess for the conglomerate/dimer in the form of the overlap of the electron densities of each individual molecule pre-computed by means of any suitable quantum chemical software. In the current study we use Gaussian 09[30] software and generate the electron density on the grid by means of the utility function cubegen [30]. In order to obtain the electron density we run a series of Gaussian calculations with B3LYP[25] exchange-correlation functional and cc-pVDZ [34, 35, 36] basis set. Subsequently, we create the initial guess for the dimer by means of a simple overlap of the two electron densities of each monomer. Once we create the initial guess we proceed to the optimization routine. In the current work we use nlopt [72] optimization library.

We run our OF-DFT calculation with two types of functionals. We use LDA approximation to the exchange-correlation functional in the form Dirac exchange [17] and Perdew and Wang correlation [18]. We also use the GGA approximation in the form of PBE exchange-correlation functional [19]. In order to evaluate them we make use of the libxc library [72]. We only evaluate the stacked configuration with the present approach. Finally, we compare our results with M062X [21, 22, 23, 24] exchange-correlation functional and use it as a benchmark in this type of calculations. We again use the Gaussian 09 [30] software to compute the binding curves of the molecules under study.

It has been mentioned above that models of the $TFD\lambda W$ have been quite popular in the literature, where the useful values of λ such as $1/3, 1/6, 1/9$ have been reported to give good approximations of atomic energies [92]. Knowing that the kinetic energy in the OF-DFT approximation is not known (it is only known for the homogeneous electron gas and one-electron system), we introduce an additional parameter following [93]. Our kinetic energy

functional looks like

$$T[\rho] = a \times C_{TF} \int \rho(\mathbf{r})^{5/3} d\mathbf{r} + b \times \frac{1}{8} \int \frac{|\nabla \rho(\mathbf{r})|^2}{\rho(\mathbf{r})} d\mathbf{r} \tag{2.21}$$

We use parameter b instead of λ. We try several values of parameter $a = 1.0, 1.1, 1.2$ and $b = 1/3, 1/6, 1/9$. We hope to see trends or/and transferability of models. The results are given below in Fig. 2.2, 2.3. The locations of minima of binding curves are given in Tab.2.1,2.2. To obtain the minima locations we perform the spline smoothing of the binding curves and find the minima of those by using simple one-dimensional minimization routines. We intentionally do not present results produced by the LDA approximation because of their incorrectness and non-smoothness of curves. On the other hand, GGA gives very smooth curves and results which are somewhat close to the benchmark. The good quality of the GGA result may be explained by the presence of the gradient dependence of the GGA functionals, which in turn enforces the smoothness of the curves.

Let us first examine the binding curves computed for the benzene dimer. Figures 2(a), 2(b), 2(c) show the results. We intensionally avoided presenting the curves with the parameter $b = 1/3$ because it creates very deep potential wells. One can clearly see from these figures that curves with parameter $b = 1/6$ significantly overbind by creating very deep potential wells as well as by having minima at a short distance. In other words, the increase of parameter b creates more and more overbinding. On the other hand, the increase of parameter a does not have any influence on the depth of the potential wells but shifts the minima to longer separations. These trends can be summarized in Tab.2.1. The minima obtained from the OF-DFT calculations are quite far from the benchmark result obtained by M06-2X where the minimum is at 3.73Å.

Next, we examine the binding curves of the thiophene dimer Fig. 3(a), 3(b), 3(c). We observe similar trends where the increase of parameter b introduces overbinding and very deep potential wells,while the increase of parameter a simply increases shifts the minima to longer distances. These trends can be seen from Tab.2.2.

After examining the binding curves and the corresponding minima, we can see that as the coefficient a increases, the distance increases as well. On the other hand, the decrease of the b coefficient makes the distance increase. This trend is clearly presented in both cases.

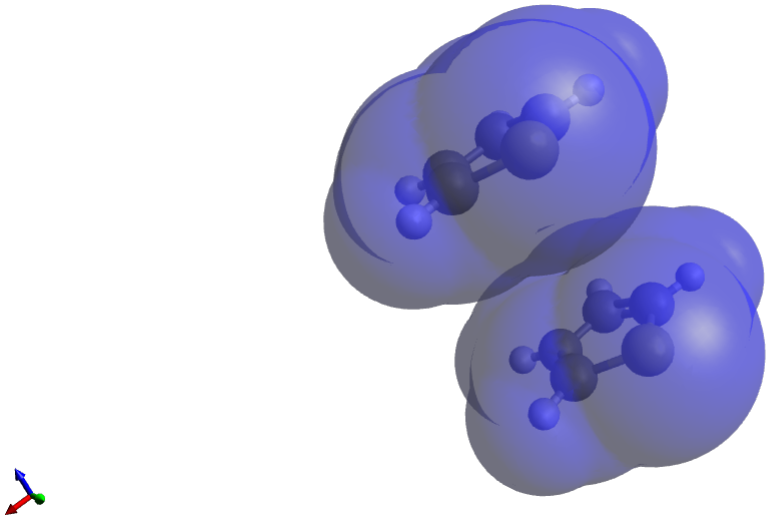


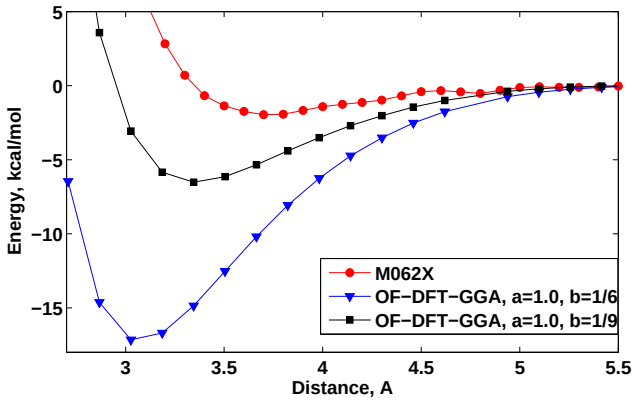
Figure 2.1: Sandwiched configuration of thiophene

We get the best model for benzene as $a = 1.2, b = 1/9$ and the best model for thiophene as $a = 1.1, b = 1/9$. We do not observe transferability across the models. In fact, the best binding curve for the benzene dimer is off by 0.2Å from the benchmark result. On the other hand, the best model for the thiophene dimer is quite accurate. If we pick a model with parameters $a = 1.1$ and $b = 1/9$ the results obtained for benzene are a little worse but are reasonable.

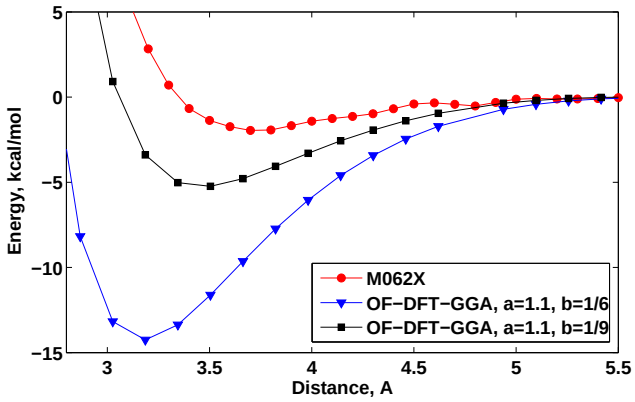
The current behavior the OF-DFT model can be explained by its inadequacy. Indeed, we tried to approximate the molecular electron density by the uniform electron gas. In order to improve the performance we should include non-local terms. Unfortunately, this may yield the model computationally infeasible. In the next section, we suggest possible improvements for the model.

2.1.8 Possible improvement of the model.

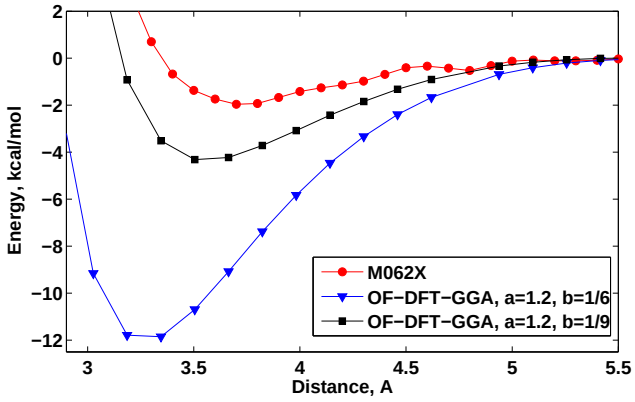
2.1.8.1 Kinetic energy. The main drawback of OF-DFT approach is the lack of the kinetic energy functional. So far, we have used the Thomas-Fermi term and the von Weiz-



(a) Stacked configuration $a = 1.0$

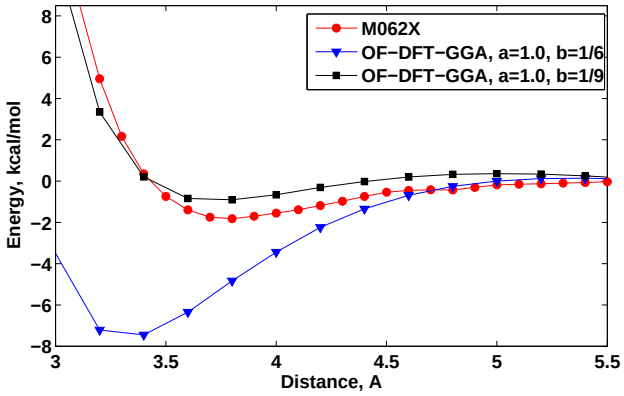


(b) Stacked configuration $a = 1.1$

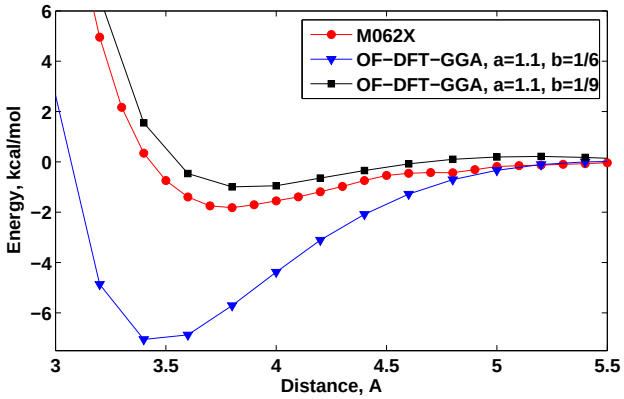


(c) Stacked configuration $a = 1.2$

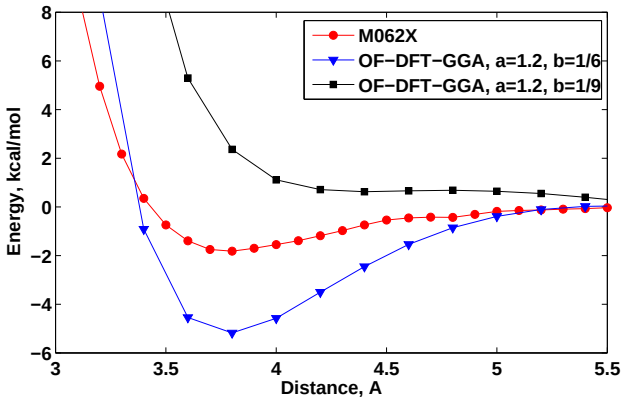
Figure 2.2: Calculation of stacked configuration of benzene binding curve with different parameters.



(a) Stacked configuration $a = 1.0$



(b) Stacked configuration $a = 1.1$



(c) Stacked configuration $a = 1.2$

Figure 2.3: Calculation of stacked configuration of thiophene binding curve with different parameters.

Table 2.1: Minima locations (Å) of benzene binding curves vs M062X minimum at 3.73Å

a/b	1	1.1	1.2
1/6	3.09	3.18	3.26
1/9	3.34	3.45	3.55

Table 2.2: Minima locations (Å) of thiophene binding curves vs M062X minimum at 3.78Å

a/b	1	1.1	1.2
1/6	3.33	3.46	3.75
1/9	3.7	3.8	none

sacker term together to approximate the kinetic energy. The von Weizsacker term originates from the expansion of the kinetic energy with respect to local inhomogeneities of the density. One can use that expansion to the fourth order, and can write

$$T[\rho] = T_{TF}[\rho] + T_2[\rho] + T_4[\rho] \tag{2.22}$$

where $T_{TF}[\rho]$ is the Thomas-Fermi expression, T_2 is the von Weizsacker term, and fourth term looks like:

$$T_4[\rho] = \frac{(3\pi^2)^{-2/3}}{540} \int \rho^{1/3} \left[\left(\frac{\nabla^2 \rho}{\rho} \right)^2 - \frac{9}{8} \frac{\nabla^2 \rho}{\rho} \left(\frac{|\nabla \rho|^2}{\rho} \right)^2 + \frac{1}{3} \left(\frac{|\nabla \rho|^2}{\rho} \right)^4 \right] d\mathbf{r} \tag{2.23}$$

As usual with perturbative expansions, the convergence with increasing number of terms is not guaranteed. In this case, the second order in known to improve over Thomas-Fermi, while the inclusion of T_4 does not represent a significant advance over T_2. In addition, this expansion cannot account for rapidly varying density such as those occurring in vacuum regions (outside molecules and surfaces). In fact, for atoms the density calculated with the truncated gradient expansion still diverges at long distances.

The correct second-order correction for one-electron systems, and also for two electron systems where the two electrons occupy the same spatia orbital is represented by the Weizsacker expression. Any approximate functional should be able to retrieve this expression in those cases, but should also reproduce the gradient expansion in the limit of slowly varying densities. Several advances in this direction have been proposed along the years, and have become more reliable in recent times.

One possibility is to propose an effective re-summation schemes similar to GGA for exchange:

$$T_s[\rho] = c_0 \int \rho(\mathbf{r})^{5/3} P(s[\rho, \nabla\rho]) d\mathbf{r} \tag{2.24}$$

with $s = c_2/c_0 |\nabla\rho/\rho^{4/3}|^2$. The function $P(s)$ should observe the two limits $P(s) \to 1 + s$ for small s, and $P(s) \to 9s$ for $s \to \infty$, in order to respect the behavior of the energy in the two known limits. Besides this, there is much freedom in how to do this. It was shown that the gradient contributions from the kinetic energy and the exchange terms are closely related, and proposed a functional form consistent with Becke's exchange. This improves the energetics of atomic and molecular systems, but does not lead to the correct ground state density by minimization of the energy functional. This issue was addressed by Wang at al. (2001), who derived a differential equation for $P(s)$ in terms of the Kohn-Sham potential, for spherically symmetric systems. They concluded that in order for a kinetic functional to achieve this goal, it must be at least of the meta-GGA type, i.e. including the Laplacian of the density. King and Handy (2001) explored a similar road, but using

$$T_s[\rho] = c_0 \int \rho(\mathbf{r})^{5/3} \left(\Phi(\mathbf{r}) + \frac{1}{8} \frac{|\nabla\rho|^2}{\rho} \right) d\mathbf{r} \tag{2.25}$$

where $\Phi(\mathbf{r})$ is a modular function that goes to zero at long distances. Interestingly, using this approach they obtained the correct atomic shell structure.

A cleaner, although computationally heavier, route to the shell structure is to abandon semi-local approximations in favor of a fully non-local approach to the kinetic term. This has been done in the spirit of the weighted density approximation for exchange. The idea here is to divide the kinetic energy into a Weizsacker term T_{vW} and a non-local contribution T_{nl} written as

$$T_{nl}[\rho] = \frac{8}{5} \int \rho(\mathbf{r}) t[\tilde{\rho}(\mathbf{r})] dr - \frac{3}{5} T_{TF}[\rho] \tag{2.26}$$

where $t[\rho]$ is the kinetic energy density in the LDA (Thomas-Fermi), and

$$\tilde{\rho}(\mathbf{r}) = \int \rho(\mathbf{r}')\omega(\xi(\mathbf{r},\mathbf{r}'),|\mathbf{r}-\mathbf{r}'|)d\mathbf{r}' \tag{2.27}$$

is a non-spherical average density defined by the weight function ω. This latter depends on the local density though a function $\xi(\mathbf{r},\mathbf{r}'))$ which, importantly is symmetric in $\mathbf{r}$ and $\mathbf{r}'$. A universal weight function is determined by imposing that the linear response of the homogeneous electron gas is reproduced.

The idea of enforcing the linear response has also been pursued by other authors, but in a slightly different form. The kinetic functional is written as $T_s = T_{TF} + T_{vW} + T_{nl}$, with

$$T_{nl}[\rho] = \int\int \rho^{\alpha}(\mathbf{r})\omega_{\alpha\beta}(\xi(\mathbf{r},\mathbf{r}'),|\mathbf{r}-\mathbf{r}'|)\rho^{\beta}(\mathbf{r}')d\mathbf{r}d\mathbf{r}' \tag{2.28}$$

and $\omega_{\alpha\beta}$ is again determined by imposing the correct linear response. There have also been proposals of incorporating the second order response into this approach. Unfortunately, the feasibility of this scheme relies on Taylor expansion around some uniform density, and is thus of little applicability to molecular systems.

2.1.8.2 Coulomb interaction. As we have mentioned above, we placed the nuclei on the lattice sites by moving them by a small margin. This may cause bad consequences because now the nuclei are located at the wrong positions. This problem can be alleviated by introducing a finer grid. At the same time doubling the dimensions of the lattice will cause the increase in the number of variables by a factor of 8, which may results in a prohibitively long calculation or an excessive amount of computer memory.

A more realistic approach is to split the Coulomb energy into three parts: the electron-electron, ion-ion, and electron-ion parts. The first part can be computed by using the Fourier Transform method described above. The ion-ion Coulomb energy can be evaluated using the Ewald summation [73] or fast multipole method [111]. The electron-ion interaction can be computed by means of the local pseudo potential. We chose to avoid the usage of pseudo potentials because of the lack of parameters used to parameterize the pseudo potential.

Further improvements can be introduced by partitioning the space into two types of regions: the regions around the nuclei enclosed by a certain cutoff radii (usually used with

pseudo potentials) and the regions outside of the cutoff radii of each nucleus. The grid within the first type regions should be made fine enough to approximate the quickly changing electron density and the grid in the second type regions can be made coarse for reason that the electron density in the outskirts of a molecule take small values and varies just slightly.

The problem with this approach can arise from the fact that when we compute the electron-electron Coulomb energy, the interaction energy of the inner regions with the outer regions should be calculated. This is not possible using the Fourier Transform because the standard Fourier Transform routines require a uniform grid. In fact this can be solved by setting up a unique fine grid for the whole system but by placing a zero charge at all second, third, fourth grid site in the outer region. This way we can reduce the number of variables quite substantially and take into account the local pseudo potential.

3.0 HARRIS APPROXIMATION METHOD FOR MODELING OF COVALENT INTERACTION.

In this section, we take a conventional approach used in the Density Functional Theory (DFT) based on Kohn-Sham (KS) equation and orbitals. We first give a derivation of the KS method and show how one can implement it using Gaussian basis sets. We will also present some other functionals used in DFT but which can only used within the KS framework. Next, we introduce the Harris approximation and explain its advantages and disadvantages.

3.1 INTRODUCTION TO KOHN-SHAM METHOD.

3.1.1 Introduction to orbitals and Kohn-Sham equations.

It is indeed appealing that the ground-state energy of a many-electron system can be obtained as the minimum of the energy functional

$$E[\rho] = \int \rho(\mathbf{r}) v(\mathbf{r}) d\mathbf{r} + F[\rho] \tag{3.1}$$

where

$$F[\rho] = T[\rho] + V_{ee}[\rho] \tag{3.2}$$

all terms having been defined in the previous section. The ground state electron density is the density that minimizes $E[\rho]$ and hence satisfies the Euler equation

$$\mu = v(\mathbf{r}) + \frac{\delta F[\rho]}{\delta \rho(\mathbf{r})} \tag{3.3}$$

where μ is the Lagrange multiplier associated with the constraint

$$\int \rho(\mathbf{r})d\mathbf{r} = N \tag{3.4}$$

Among all possible solutions of (3.3), one takes that which minimizes $E[\rho]$.

We have seen in the previous section how one can proceed to approximate implementation by making certain assumptions. How can we do better? How can we avoid the great loss in accuracy associated with the Thomas-Fermi model and its derivative models?

Thomas-Fermi and related models constitute a direct approach, where one constructs explicit approximate forms for $T[\rho]$ and $V_{ee}[\rho]$. This produces a nice simplicity, because the equations involve electron density alone. Unfortunately, however, there are incredible difficulties in going beyond the crude level of approximation. In a trade of simplicity for accuracy, Kohn and Sham invented an indirect approach to the kinetic energy functional $T[\rho]$, the Kohn-Sham (KS) method.

Kohn and Sham proposed to introduce orbitals into the problem in such a way that the kinetic energy can be computed simply to good accuracy, leaving a small residual correction that is handled separately. To understand what is involved and what Kohn and Sham did, it is convenient to begin with the exact formula for the ground-state kinetic energy,

$$T = \sum_{i=1}^{N} n_i < \psi_i| -\frac{1}{2}\nabla^2|\psi_i > \tag{3.5}$$

where the ψ_i and n_i are, respectively, natural spin orbitals and their occupation numbers. The Pauli principle requires that $0 \leq n_i \leq 1$. From the Hohenberg-Kohn theory this T is a functional of the total electron density

$$\rho(\mathbf{r}) = \sum_{i=1}^{N} n_i \sum_{s} |\psi_i(\mathbf{r}, s)| \tag{3.6}$$

For any interacting system of interest, there is an infinite number of terms in (3.5) and (3.6). Kohn and Sham showed that one can build a theory using simpler formulas namely

$$T_s[\rho] = \sum_{i=1}^{N} n_i < \psi_i| -\frac{1}{2}\nabla^2|\psi_i > \tag{3.7}$$

and

$$\rho(\mathbf{r}) = \sum_{i=1}^{N} \sum_{s} |\psi_i(\mathbf{r}, s)| \tag{3.8}$$

Equations (3.7) and (3.8) are special cases of (3.5) and (3.6) having $n_i = 1$ for N orbitals and $n_i = 0$ for the rest. This representation of kinetic energy and density holds true for the determinantal wave function that exactly describes N noninteracting electrons.

Any nonnegative, continuous, and normalized density ρ is N- representable and always can be decomposed according to (3.8). But given a $\rho(\mathbf{r})$, how can we have a unique decomposition in terms of orbitals so as to give a unique value to $T_s[\rho]$ through (3.7)?

In analogy with the Hohenberg-Kohn definition of the universal functional $F_{HK}[\rho]$, Kohn and Sham invoked a corresponding noninteracting reference system, with the Hamiltonian

$$\hat{H}_s = \sum_{i=1}^{N} \left(-\frac{1}{2}\nabla_i^2 \right) + \sum_{i=1}^{N} v_s(\mathbf{r}) \tag{3.9}$$

in which there are no electron-electron repulsion terms, and for which the ground-state electron density is exactly ρ. For this system there will be an exact determinantal ground-state wave function

$$\Psi_s = \frac{1}{\sqrt{N!}} \det[\psi_1 \psi_2 \cdots \psi_N] \tag{3.10}$$

where the ψ_i are the N lowest eigenvalues of the one-electron Hamiltonian $\hat{h}_s$:

$$\hat{h}_s \psi_i = \left[-\frac{1}{2}\nabla^2 + v_s(r) \right] \psi_i = \epsilon_i \psi_i \tag{3.11}$$

The kinetic energy is $T_s[\rho]$, given by (3.7)

$$T_s[\rho] = < \Psi_s | \sum \left(-\frac{1}{2}\nabla^2 \right)_{i=1}^{N} | \Psi_s > = \sum_{i=1}^{N} < \psi_i | -\frac{1}{2}\nabla^2 | \psi_i > \tag{3.12}$$

and the density is decomposed as in (3.8).

The definition of $T_s[\rho]$ leaves an undesirable restriction on the density - it needs to be noninteracting v-representable. That is, there must exist a noninteracting ground state with the given $\rho(\mathbf{r})$. In the Kohn-Sham approach this restriction can be lifted, and $T_s[\rho]$ of the form (3.7) can be defined for any density derived from an antisymmetric wave function.

The quantity $T_s[\rho]$, although uniquely defined for any density, is still not the exact kinetic energy functional $T[\rho]$. The very clever idea of Kohn and Sham is to set up a problem of interest in such a way that $T_s[\rho]$ is its kinetic energy component, exactly. The resultant theory turns out to be of independent particle form.

To produce the desired separation out of $T_s[\rho]$ as the kinetic energy component, rewrite (3.2) as

$$F[\rho] = T_s[\rho] + J[\rho] + E_{xc}[\rho] \tag{3.13}$$

where

$$E_{xc}[\rho] = T[\rho] - T_s[\rho] + V_{ee}[\rho] - J[\rho] \tag{3.14}$$

The defined quantity $E_{xc}[\rho]$ is called the exchange-correlation energy.

The Euler equation now becomes

$$\mu = v_{eff}(\mathbf{r}) + \frac{\delta T_s[\rho]}{\delta \rho(\mathbf{r})} \tag{3.15}$$

where the KS effective potential is defined by

$$v_{eff}(\mathbf{r}) = v(\mathbf{r}) + \frac{\delta J[\rho]}{\delta \rho(\mathbf{r})} + \frac{\delta E_{xc}[\rho]}{\delta \rho(\mathbf{r})} = v(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + v_{xc}(\mathbf{r}) \tag{3.16}$$

with the exchange-correlation potential

$$v_{xc}(\mathbf{r}) = \frac{\delta E_{xc}[\rho]}{\delta \rho(\mathbf{r})} \tag{3.17}$$

The Kohn-Sham procedure runs as follows. Equation (3.15) with the constraint (3.4) is precisely the same equation that one obtains from conventional density functional theory when one applies it to a system of noninteracting electrons moving in the external potential $v_s(\mathbf{r}) = v_{eff}(\mathbf{r})$. Therefore, for a given $v_{eff}(\mathbf{r})$, one obtains the $\rho(\mathbf{r})$ that satisfies (3.15) simply by solving the N one electron equations

$$\left[-\frac{1}{2}\nabla^2 + v_{eff}(\mathbf{r}) \right] \psi_i = \epsilon_i \psi_i \tag{3.18}$$

and setting

$$\rho(\mathbf{r}) = \sum_{i=1}^{N} \sum_{s} |\psi_i(\mathbf{r}, s)| \tag{3.19}$$

3.1.2 Derivation of the Kohn-Sham equations.

We here express the Hohenberg-Kohn variational problem (3.3) in terms of the Kohn-Sham orbitals appearing in (3.10). The energy functional (3.1) can be rewritten as

$$E[\rho] = T_s[\rho] + J[\rho] + E_{xc}[\rho] + \int v(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} = \tag{3.20}$$

$$= \sum_{i=1}^{N}\sum_{s}\int \psi_i^*(\mathbf{r})\left(-\frac{1}{2}\nabla^2\right)\psi_i(\mathbf{r})dr + J[\rho] + E_{xc}[\rho] + \int v(\mathbf{r})\rho(\mathbf{r})d\mathbf{r}$$

and the electron density as

$$\rho(r) = \sum_{i=1}^{N}\sum_{s}|\psi_i(\mathbf{r},s)|^2 \tag{3.21}$$

We thus have the energy expressed in terms of N orbitals.

If the N orbitals are allowed to vary over the space of functions that are continuous - to have finite kinetic energy, and be square integrable - to guarantee normalization, then the density ρ covers all N-representable densities, the appropriate domain of definition of $E[\rho]$ in (3.20). That is to say, variational search for the minimum of $E[\rho]$ can be equivalently effected in the space of orbitals $\{\psi_i\}$. In doing this one must actually constrain the orbitals to be orthogonal, namely,

$$\int \psi_i^*(\mathbf{x})\psi_j(\mathbf{x})d\mathbf{x} = \delta_{ij} \tag{3.22}$$

because otherwise the kinetic energy formula of (3.7) would not be valid. Note that (3.22) implies that $\rho(\mathbf{r})$ remains normalized as required by (3.4).

Define the functional of the N orbitals

$$\Omega[\{\psi_i\}] = E[\rho] - \sum_{i=1}^{N}\sum_{j=1}^{N}\epsilon_{ij}\int \psi_i^*(\mathbf{x})\psi_j(\mathbf{x})d\mathbf{x} \tag{3.23}$$

where $E[\rho]$ is the functional of the ψ_i expressed in (3.20) and the ϵ_{ij} are Lagrange multipliers for the constraints (3.22). For $E[\rho]$ to be a minimum,it is necessary that

$$\delta\Omega[\{\psi_i\}] = 0 \tag{3.24}$$

which leads to the equations

$$\hat{h}_{eff}\psi_i = \left[-\frac{1}{2}\nabla^2 + v_{eff}\right]\psi_i = \sum_{j=1}^{N}\epsilon_{ij}\psi_j \tag{3.25}$$

with the effective potential $v_{eff}(\mathbf{r})$ determined from the density through (3.16). $\hat{h}_{eff}$ is a Hermitian operator, hence ϵ_{ij} is a Hermitian matrix, and can be diagonalized by a unitary transformation of the orbitals. Such a transformation leaves invariant the determinant of (3.10), the density (3.21), and hence the Hamiltonian of (3.25). The Kohn-Sham orbital equations are thus obtained in their canonical form:

$$\left[-\frac{1}{2}\nabla^2 + v_{eff}\right]\psi_i = \epsilon_i\psi_i \tag{3.26}$$

$$v_{eff}(\mathbf{r}) = v(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}' + v_{xc}(\mathbf{r}) \tag{3.27}$$

$$\rho(\mathbf{r}) = \sum_{i=1}^{N}\sum_{s}|\psi_i(\mathbf{r},s)|^2 \tag{3.28}$$

In (3.25) and (3.26), the solutions ψ_i can be different. These equations are nonlinear and be solved iteratively. The total energy can be determined from the resultant density from the formula

$$E = \sum_{i=1}^{N}\epsilon_i - \frac{1}{2}\int\int\frac{\rho(\mathbf{r})\rho(\boldsymbol{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}d\mathbf{r}' + E_{xc}[\rho] - \int v_{xc}(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} \tag{3.29}$$

3.2 EXCHANGE-CORRELATION FUNCTIONALS AND OTHER METHODS.

3.2.1 Hybrid HF-KS approaches.

The observation that LDA and GGA trends are opposite to those of Hartree-Fock motivated the development of approximations which combine these two approaches. These involve a DFT correlation with a combination of DFT and Hartree-Fock exchange:

$$E_{xc}^{hyb} = \alpha E_X^{HF} + (1-\alpha)E_X^{DFT} + E_C^{DFT} \tag{3.30}$$

where the coefficient α is either chose to assume a specific value such as $1/2$, or is fitted to some properties of a molecular database.

As example of the latter is the approximation known as B3LYP. This reproduces the geometries and binding energies of molecular systems, to the same accuracy of low-level correlated quantum chemistry approaches like second order Moller-Plesset perturbation theory (MP2) [26], and sometimes even at a higher level such as coupled clusters or CI methods. This idea is appealing and physically sensible, but the approach of fitting the coefficient α to molecular data is not compatible with an *ab initio* approach.

Interestingly, it has been shown that hybrid schemes have a rigorous formal justification within the exact DFT scheme, when this is formulated as a generalized Kohn-Sham approach. The reference system is still described by a Slater determinant, but, at variance with the usual Kohn-Sham approach, it does not correspond to the minimization of the expectation value of the kinetic energy $T_s =< \Psi_s|T|\Psi_s >$. Instead, it is obtained by minimizing a different functional under the constraint that the wave function can be written as a Slater determinant. If this functional is chosen to be

$$T_{HF-KS} =< \Psi_s|T + V_{ee}|\Psi_s >= T_s + E_H[\Psi] + E_X[\Psi] \tag{3.31}$$

the scheme resembles the Hartree-Fock method, but it contains an unknown, formally exact correlation term that is absent in standard HF. This approach is known as the Hartree-Fock-Kohn-Sham scheme (HF-KS). Another possibility is to construct a functional where the exchange term is replaced by a screened exchange, with the bare Coulomb interaction replaced with a statically screened interaction of the Yukawa type.

3.2.2 Van der Waals (dispersion) interactions.

The issue of van der Waals of dispersion interactions is a difficult benchmark is many-body theory that ultimately any correlation functional should address. The origin of the van der Waals interaction between two non-chemically bonded fragments is the coupling of the electric field generated by fluctuations in the electronic density of one fragment with the density of the other fragment. This is a dynamical correlation effect that the usual local and semi-local functionals such LDA and GGA cannot capture, and is not related to the exchange. At long distances the van der Waals interaction should approach the classical dipole-dipole

interaction, which decays as $E_{vdW} = -C_6/R^6$. Typical van der Waals systems are dimers of closed-shell atoms. Most theoretical works concentrated precisely on reproducing this long-range behavior by separating the electron-electron interaction into a short-range and a long-range part. The long-range part is responsible for the van der Waals interaction, and can be typically represented by an effective interaction of the form

$$U_{ee}^{lr}(\mathbf{r}, \mathbf{r}^{'}) = \frac{A(\mathbf{r}, \mathbf{r}^{'})}{|\mathbf{r} - \mathbf{r}^{'}|^6} \tag{3.32}$$

where $A(\mathbf{r}, \mathbf{r}^{'})$ depends on the choice of an effective density for the exchange-correlation linear response kernel, like $\rho_{eff} = \sqrt{\rho(\mathbf{r})\rho(\mathbf{r}^{'})}$ or $\rho_{eff} = \{\sqrt{\rho(\mathbf{r})\rho(\mathbf{r}^{'})}[\sqrt{\rho(\mathbf{r})} + \sqrt{\rho(\mathbf{r}^{'})}]\}^{2/3}$. With this, the long-range contribution to the XC energy from two well-separated fragments in volumes V_1 and V_2 is given by

$$E_{XC}^{lr} = -\frac{3}{\pi}\int_0^\infty du \int_{V_1} d\mathbf{r}_1 \int_{V_2} d\mathbf{r}_2 \frac{\chi_1^z(iu)\chi_2^z(iu)}{|\mathbf{r}_1 - \mathbf{r}_2|^6} \tag{3.33}$$

where $\chi_i^z(\omega)$ is the density response of a uniform 5electron gas of density $\rho(\mathbf{r}_i)$ to a perturbation in the direction of the bond. A more general approach that calculates the density response using real-time propagation has been proposed by Kohn. This finally leads to the well known long-range behavior of the van der Waals interaction between two fragments separated by a distance R:

$$E_{vdW}(R) = -\left(\frac{3}{\pi}\Im\int_0^\infty \chi_1^z(\omega)\chi_2^z(\omega)d\omega\right)\frac{1}{R^6} \tag{3.34}$$

3.2.3 Becke-Johnson approach to dispersion interaction.

The dispersion interaction between molecules is a weak attraction attributed to an instantaneous dipole moment in one molecule inducing a dipole moment in another molecule. The resulting dipole-induced-dipole interaction leads, in the limit of large intermolecular separation, to a dispersion energy of the form

$$E_{disp} = -\frac{C_6}{R^6} \tag{3.35}$$

where the dispersion coefficient C_6 is a constant depending on the molecules involved.

While dispersion interactions are very important in chemistry, efficient modeling of dispersion remains a hard computational problem. Correlated *ab initio* methods with large basis sets provide an accurate treatment of dispersion but are computationally expensive. There is considerable interest in modifying the more computationally affordable methods like Hartree-Fock or DFT to approximately account for dispersion.

Such modifications usually involve addition of a dispersion term of the following form to the HF or DFT energy:

$$E_{disp} = -\sum_{i>j} \frac{C_6}{R_{ij}^6} \tag{3.36}$$

The summation is over all atom pairs and C_6 coefficients depend on the atoms i and j.

But how does one determine the values of the interatomic C_6 coefficients? Interatomic C_6's are commonly obtained from empirical fits. An example is the Slater-Kirkwood [106, 107] approach in which C_6's are obtained from atomic polarizabilities fit to molecular polarizability data [108]. The atomic polarizabilities are dependent on the molecular environment and involve explicit atom types. Wu and Yang [109] have suggested the interatomic C_6's be directly fit to a reference set of intermolecular C_6 data. Intermolecular C_6's can be obtained from experimental dipole oscillator strengths and can be rigourously calculated from frequency dependent polarizabilities. Wu and Yang's fitted C_6's were averaged over atom types to obtain a more general parameter set and used in more extensive calculations by Grimme [15, 16].

The Becke-Johnson model [98, 99, 100, 101, 102] uses the *ab initio* approach to calculate these coefficients. They consider the exchange hole (A.2.11) (see Appendix) of an atom

or molecule, $h_{X\sigma}(\mathbf{r}_1, \mathbf{r}_2)$. When a σ-spin electron is at position $\mathbf{r}_1$ the hole measures the depletion of probability, with respect to the total electron density, of finding another σ-spin electron at position $\mathbf{r}_2$. The electron plus its hole has a zero net charge. However, the hole is generally not spherically symmetric around $\mathbf{r}_1$ and the electron plus its exchange hole therefore has a non-zero dipole moment. Becke and Johnson proposed that the exchange hole instantaneous dipole moment is the source of dispersion interaction between non-overlapping systems.

For the isotropic C_6 coefficients between the two systems A and B, they obtained

$$C_6 = \frac{< d_X^2 >_A < d_X^2 >_B \alpha_A \alpha_B}{< d_X^2 >_A \alpha_B + < d_X^2 >_B \alpha_A} \tag{3.37}$$

where $< d_X^2 >$ is the expectation value of the squared exchange-hole dipole moment and α is the isotropic molecular polarizability. $< d_X^2 >$ are easily computed by numerical integration, over HF or DFT orbitals

$$< d_X^2 > = < d_{X\alpha} >^2 + < d_{X\beta} >^2 \tag{3.38}$$

$$< d_{X\sigma} >^2 = \int \rho_\sigma(\mathbf{r}_1) d_{X\sigma}^2(\mathbf{r}_1) d\mathbf{r}_1 \tag{3.39}$$

where ρ_σ is the σ-spin density and the dipole moment of the σ-spin exchange hole at reference point $\mathbf{r}_1$ is given by

$$d_{X\sigma}(\mathbf{r}_1) = \left[\frac{1}{\rho_\sigma(\mathbf{r}_1)} \sum_{ij} \psi_{i\sigma}(\mathbf{r}_1) \psi_{j\sigma}(\mathbf{r}_1) \int \mathbf{r} \psi_{i\sigma}(\mathbf{r}) \psi_{j\sigma}(\mathbf{r}) d\mathbf{r} \right] - \mathbf{r}_1 \tag{3.40}$$

3.2.4 Grimme approach to dispersion interaction.

In the present work, we are using dispersion C_6 coefficients obtained by Grimme [15, 16]. As we mentioned above, these values were obtained by fitting to a large molecular data set. In particular, the total energy is split into two parts

$$E_{DFT-D} = E_{KS-DFT} + E_{disp} \tag{3.41}$$

where E_{KS-DFT} is the usual self-consistent Kohn-Sham energy as obtained from the chosen Density Functional and E_{disp} is an empirical dispersion correction given by

$$E_{disp} = -s_6 \sum_{i=1}^{N_{at}-1} \sum_{j=i+1}^{N_{at}} \frac{C_6^{ij}}{R_{ij}^6} f_{dmp}(R_{ij}) \tag{3.42}$$

Here, N_{at} is the number of atoms in the system, C_6^{ij} denotes the dispersion coefficient for atom pair ij, s_6 is a global scaling factor that only depends on the Density Functional used, and R_{ij} is an interatomic distance. In order to avoid near-singularities for small R, a damping function f_{dmp} must be used, which is given by

$$f_{dmp}(R_{ij}) = \frac{1}{1 + e^{-d(R_{ij}/R_r - 1)}} \tag{3.43}$$

where R_r is the sum of atomic van der Waals radii.

Because higher-order dispersion terms such as C_8 and C_{10} that have been used in a similar method are more short-ranged and strongly interfere with the damping function, Grimme did not include them in his method. Grimme also proposed to use a geometric mean for the composed coefficients

$$C_6^{ij} = \sqrt{C_6^i C_6^j} \tag{3.44}$$

3.3 THE TIGHT BINDING METHODS.

The tight binding (TB) approach to electronic structure of solids is complementary to the nearly free electron picture. While the latter is a reasonably good representation of the electronic structure of simple metals, TB provides a plausible representation of systems where the electrons are localized in chemical bonds of different degrees of covalency.The starting point of this model is to consider that, in a first approximation, electrons are localized in a single atom, but they have the possibility to jump to neighboring atoms. TB methods range from very basic empirical models to the most sophisticated *ab initio* schemes, where the Hamiltonian matrix is derived from DFT.

3.3.1 Empirical tight-binding.

In a sense, this is the simplest possible form of the TB model. The state of an electron in atom i is indicated by the ket $|i>$. In a real-space representation, this would be the atomic eigenstate $\phi_i(\mathbf{r}-\mathbf{R}_i)$. If the atoms are sufficiently far apart, then, in the crudest approximation, electrons in every atom will have the same on-site energy value ϵ_0. This situation can be represented by a model Hamiltonian of the form $\hat{H}_0 = \epsilon_0 \sum_i |i><i|$.

When the atoms are brought together, we need to take into account the possibility of the electrons jumping from one atom to any other of its neighbors. This is achieved by including off-diagonal elements in the Hamiltonian, which now looks like

$$\hat{H}_{TB} = \epsilon_0 \sum_i |i><i| + t \sum_i \sum_j |i><j| \tag{3.45}$$

In a model situation, the sum on j runs only over the nearest neighbors of i. The parameter t is usually called hopping integral.

In order to understand the meaning of the hopping integrals, we need to consider the real-space version of TB Hamiltonian for a single electron in a molecule,

$$\hat{H}_{TB} = -\frac{\hbar^2}{2m}\nabla^2 + \sum_K v_K(\mathbf{r}-\mathbf{R}_K) \tag{3.46}$$

and propose a linear combination of atomic orbitals for the TB wave function, $\phi(\mathbf{r}) = \sum_j C_j \phi_j(\mathbf{r} - \mathbf{R}_j)$, where

$$\left[\frac{\hbar^2}{2m}\nabla^2 + v_j(\mathbf{r})\right]\phi_j(\mathbf{r}) = \epsilon_0\phi_j(\mathbf{r}) \tag{3.47}$$

By replacing the expression into Schrodinger's equation, we obtain

$$\hat{H}_{TB}\phi(\mathbf{r}) = \sum C_j \left[-\frac{\hbar^2}{2m}\nabla^2 + v_j(\mathbf{r} - \mathbf{R}_j) + \sum_{K\neq j} v_K(\mathbf{r} - \mathbf{R}_K)\right]\phi_j(\mathbf{r}) = \epsilon_0 \sum_j C_j\phi_j(\mathbf{r}) \tag{3.48}$$

which, after multiplication by $\phi_K^*(r)$ and integration gives

$$H_{TB}^{jj} = \int \phi_j^*(\mathbf{r})\left[-\frac{\hbar^2}{2m}\nabla^2 + v_j(\mathbf{r} - \mathbf{R}_j)\right]\phi_j(\mathbf{r})d\mathbf{r} = \epsilon_0 \tag{3.49}$$

and

$$H_{TB}^{ij} = -\frac{\hbar^2}{2m}\int \phi_j^*(\mathbf{r})\nabla^2\phi_j(\mathbf{r})d\mathbf{r} + \sum_K \int \phi_j^*(\mathbf{r})v_K(\mathbf{r} - \mathbf{R}_K)\phi_j(\mathbf{r})d\mathbf{r} \tag{3.50}$$

These are precisely hopping terms, which involve two- and three-center integrals. The two-center integrals are the most important contribution. If three-center integrals are neglected, what remains is

$$H_{TB}^{ij} \approx \int \phi_i^*(\mathbf{r})\left[-\frac{\hbar^2}{2m}\nabla^2 + v_i(\mathbf{r} - \mathbf{R}_i) + v_j(\mathbf{r} - \mathbf{R}_j)\right]\phi_j(\mathbf{r})d\mathbf{r} \tag{3.51}$$

which correspond to electron being shared between two atoms i and j. There is a kinetic energy contribution to hopping, but more important are the potentials attracting the electron to the two atoms. The most substantial part of the kinetic energy is actually associated with the on-site terms.

The electronic energy in the empirical tight-binding model, as in *ab initio* tight-binding methods, is given simply by the sum of the eigenvalues of the Hamiltonian up to the highest occupied energy level.

The above empirical tight-binding model can be made more realistic by taking into account the identity of the atoms. This is achieved by considering that the electrons in the isolated atom occupy their corresponding atomic orbitals, and these orbitals hybridize to give rise to energy bands. Therefore, for each atomic species we consider a minimal set of atomic valence orbitals,i.e. one orbital for each valence state occupied in the isolated atom.

Within this picture first- and second-row elements are represented by one s and three p orbitals, while transition metal atoms require one s and five d orbitals. This more elaborate tight-binding Hamiltonian can be expressed as:

$$\hat{H}_{TB} = \sum_i \sum_\alpha \epsilon_\alpha^i |\phi_\alpha^i >< \phi_\alpha^i| + \sum_{\alpha\beta} \sum_{i \neq j} t_{\alpha\beta}^{ij} |\phi_\alpha^i >< \phi_\beta^j| \tag{3.52}$$

where $|\phi_\alpha^i >$ represents an atomic orbital of symmetry α. ϵ_α^i are the on-site atomic energies associated with these orbitals, and depend on the type of orbitals.

3.3.2 ab initio tight-binding.

As suggested by similarity in the form of the Hamiltonian, it is possible to make a connection between the Kohn-Sham scheme for an atom-centered basis set and tight-binding approach. The Kohn-Sham Hamiltonian matrix element is

$$H_{KS\mu\nu}^{ij} =< \phi_\mu^i | \left[-\frac{\hbar^2}{2m} \nabla^2 + \hat{v}_{eff}[\rho] \right] | \phi_\nu^j > \tag{3.53}$$

The matrix element of the effective potential can be written as

$$v_{eff,\mu\nu}^{ij}[\rho] = U_{\mu\nu}^{ij} + \mu_{XC\mu\nu}^{ij}[\rho] + \sum_{\gamma\delta=1}^{M} \sum_{KL=1}^{P} \rho_{\gamma\delta}^{LK} < \mu^i \gamma^K | \nu^i \delta^L > \tag{3.54}$$

with $U_{\mu\nu}^{ij}$ the matrix element of the nuclear attraction and $\mu_{XC\mu\nu}^{ij}$ the matrix element of the exchange-correlation potential. The last term is the Hartree potential written in terms of the density matrix and the Coulomb two-electron matrix elements.

The calculation of the Kohn-Sham matrix elements formally involves the calculation of integrals ranging from one up to four centers, although the latter are reduced to two-center integrals if the density is expressed in an atom-centered basis set, or if the Hartree potential is expressed as a sum of atom-centered potentials. To avoid having to calculate the integrals every time they are required, a standard procedure in TB schemes is to parameterize them as a function of the type of basis function and distance between the centers. Two-center integrals involve one orbital at each center, and are given by

$$t_{\mu\nu}^{ij} = \int \phi_\mu^*(\mathbf{r} - \mathbf{R}_i) \hat{H}_{KS}^{TB} \phi_\nu(\mathbf{r} - \mathbf{R}_j) d\mathbf{r} \tag{3.55}$$

with

$$\hat{H}_{KS}^{TB} = -\frac{\hbar^2}{2m}\nabla^2 + v_{eff}^{(i)}(\mathbf{r} - \mathbf{R}_i) + v_{eff}^{(j)}(\mathbf{r} - \mathbf{R}_j) \tag{3.56}$$

where the approximation was used that the effective potential can be written as $\hat{v}_{eff} = \sum_K \hat{v}_{eff}^{(K)}$.

3.3.3 Harris functional.

In principle, the solution of Kohn-Sham equations requires the self-consistent determination of the density matrix. It is possible, however, to derive a non-self-consistent approach akin to the empirical tight-binding approach. This is based on the approximation to the energy proposed by Harris and Foulkes [28]. This energy functional was originally devised as an approximation to the energy of weakly interacting fragments of the system. The approximation is almost equivalent to expanding the Kohn-Sham energy expression about the density that is the sum of overlapped but frozen fragment densities and neglecting corrections which are quadratic in the difference density.

The Kohn-Sham scheme for calculating energies is based on the expression

$$E = T_0 + \int d\mathbf{r}\rho(\mathbf{r})[\frac{1}{2}\phi(\mathbf{r}) + V_{ext}(\mathbf{r})] + E_{xc} + E_N \tag{3.57}$$

where E is the total energy for N electrons in nuclear field $V_{ext}(\mathbf{r})$, $\rho(\mathbf{r})$, and $\phi(\mathbf{r})$ are the electron density and associated Poisson potential, E_N is the internuclear repulsion, T_0 is the kinetic energy of a system of independent electrons, and E_{xc} is the so-called exchange-correction energy. The solution is done by iteratively solving

$$[-\frac{1}{2}\nabla^2 + V(\mathbf{r})]\psi_i(\mathbf{r}) = \epsilon_i\psi_i(\mathbf{r}) \tag{3.58}$$

where

$$T_0 = \int \sum_i n_i\psi_i^*(\mathbf{r}) \left(-\frac{1}{2}\nabla^2\right) \psi_i(\mathbf{r})d\mathbf{r} \tag{3.59}$$

$$\rho(\mathbf{r}) = \sum_i n_i\psi_i^*(\mathbf{r})\psi_i(\mathbf{r}) \tag{3.60}$$

where n_i are occupation numbers. The potential $V(\mathbf{r})$ is then varied and the process repeated until the minimum of E is found. The potential and density at minimum, $V_0(\mathbf{r})$, $\rho_0(\mathbf{r})$ will then be found to satisfy the self-consistency condition

$$V_0(\mathbf{r}) = \phi(\mathbf{r}) + V_{ext}(\mathbf{r}) + \frac{\delta E_{xc}}{\delta \rho(\mathbf{r})}|_{\rho=\rho_0} \tag{3.61}$$

and this is used to guide the potential towards its minimum value by iteration.

Subject to some representability requirements, the Kohn-Sham scheme is formally exact. Furthermore, an exact expression for E_{xc} can be written down if one invokes the adiabatic connection formulation of the theory, where (3.57) is viewed as a functional of the one-electron potential, $V(r)$. The link between E_{xc} and the density is then implicit and the functional derivative in (3.61) is given by

$$\mu_{xc}(\mathbf{r}) = \frac{\delta E_{xc}}{\delta \rho(\mathbf{r})} = \int d\mathbf{r}' \frac{\delta E_{xc}}{\delta V(\mathbf{r}')} \frac{\delta V(\mathbf{r}')}{\delta \rho(\mathbf{r})} \tag{3.62}$$

If a local-density approximation (LDA) is used,

$$E_{xc} = \int \rho(\mathbf{r}) \epsilon_{xc}(\rho(\mathbf{r})) d\mathbf{r} \tag{3.63}$$

the above expression reduces to

$$\mu_{xc}(\mathbf{r}) = \epsilon_{xc}(\rho(\mathbf{r})) + \rho(\mathbf{r}) \frac{d\epsilon_{xc}(\rho(\mathbf{r}))}{d\rho(\mathbf{r})} \tag{3.64}$$

and the dependence of E_{xc} and μ_{xc} on the orbital structure is lost.

Consider two fragments, F_1 and F_2, whose coupling energy is of interest. Let $\rho_1(\mathbf{r})$ and $\rho_2(\mathbf{r})$ be the densities that correspond to the exact minimum of (3.57) for isolated fragments. These are given by (3.60) for some orbitals and occupations such that (3.58) and (3.61) are satisfied simultaneously. The energy of fragment F_1 can then be written in terms of the self-consistent eigenvalues ϵ_i^1 and density $\rho_1(\mathbf{r})$

$$E_1 = \sum_i n_i^1 \epsilon_i^i - \int \rho_1(\mathbf{r}) [\frac{1}{2}\phi_1(\mathbf{r}) + \mu_{xc}^{\rho_1}(\mathrm{R})] d\mathbf{r} + E_{xc}[\rho_1] + E_N^1 \tag{3.65}$$

with a similar expression for fragment F_2. Suppose we place the fragments some finite distance R apart and solve the seif-consistency equations exactly, obtaining density $\rho(\mathbf{r})$, occupations and eigenvalues n_i, ϵ_i, and energy

$$E_R = \sum_i n_i \epsilon_i - \int \rho(\mathbf{r})[\frac{1}{2}\phi(r) + \mu_{xc}^{\rho}(\mathbf{r})]d\mathbf{r} + E_{xc}[\rho] + E_N^R \tag{3.66}$$

Write now

$$\rho(\mathbf{r}) = \rho_f(\mathbf{r}) + \delta\rho(\mathbf{r}) \tag{3.67}$$

with $\rho_f(\mathbf{r}) = \rho_1(\mathbf{r}) + \rho_2(\mathbf{r})$ with the two frozen fragment densities placed at separation R, and assume $\delta\rho(\mathbf{r})$ to be sufficiently small that quadratic corrections can be ignored. Introduce a potential

$$\tilde{V}(\mathbf{r}) = \phi_f(\mathbf{r}) + \mu_{xc}^{\rho_f}(\mathbf{r}) + V_{ext}(\mathbf{r}) \tag{3.68}$$

where $V_{ext}(\mathbf{r})$ is the true external field for separation R and ϕ_f and $\mu_{xc}^{\rho_f}$ are computed from the overlapped fragment densities. Then the difference between the true self-consistent potential and $\tilde{V}(\mathbf{r})$ is

$$\Delta V(\mathbf{r}) = \phi(\mathbf{r}) - \phi_f(\mathbf{r}) + \mu_{xc}^{\rho}(\mathbf{r}) - \mu_{xc}^{\rho_f}(\mathbf{r}) \tag{3.69}$$

which will be assumed small. If $\tilde{\epsilon}_i$ are eigenvalues corresponding to $\tilde{V}(\mathbf{r})$, we have

$$\sum_i n_i \epsilon_i = \sum n_i \tilde{\epsilon}_i - \int \rho(r)\Delta V(\mathbf{r})d\mathbf{r} + O(\Delta V^2) \tag{3.70}$$

which, on substituting in (3.66) gives

$$E_R = \sum n_i \tilde{\epsilon}_i + \int \rho(\mathbf{r})[\frac{1}{2}\phi(\mathbf{r}) - \phi_f(\mathbf{r}) - \mu_{xc}^{\rho_f}(\mathbf{r})]d\mathbf{r} + E_{xc}[\rho] + E_N^R \tag{3.71}$$

Using

$$E_{xc}[\rho] = E_{xc}[\rho_f] + \int \mu_{xc}^{\rho_f}(\mathbf{r})\delta\rho(\mathbf{r})d\mathbf{r} + O(\delta\rho^2) \tag{3.72}$$

and noting that ϕ is linear in ρ, we then find

$$E_R = \sum n_i \tilde{\epsilon}_i - \int \rho_f(\mathbf{r})[\frac{1}{2}\phi_f(\mathbf{r}) + \mu_{xc}^{\rho_f}(\mathbf{r})]d\mathbf{r} + E_{xc}[\rho_f] + E_N^R \tag{3.73}$$

3.4 PRACTICAL IMPLEMENTATION.

In the following subsections, we will give a brief overview of how the solution of Kohn-Sham equations is implemented in practice. We will closely follow the paper by Pople, Gill and Johnson [110].

3.4.1 Solution of Kohn-Sham equations within a finite basis set.

In the current treatment we adopt a spin-unrestricted format, α and β electrons are assigned to sets of orthogonal orbitals $\psi_i^\alpha (i = 1, ..., n_\alpha)$ with $\psi_i^\beta (i = 1, ..., n_\beta)$ respectively. Corresponding α, β and total densities are

$$\rho_\alpha = \sum_{i=1}^{n_\alpha} |\psi_i^\alpha|^2 \tag{3.74}$$

$$\rho_\beta = \sum_{i=1}^{n_\beta} |\psi_i^\beta|^2 \tag{3.75}$$

$$\rho = \rho^\alpha + \rho^\beta \tag{3.76}$$

Kinetic energy T_s, electron-ion interaction E_V, and electron-electron interaction J energies are given by

$$T_s = \sum_i^{n_\alpha} (\psi_i^\alpha| - \frac{1}{2}\nabla^2|\psi_i^\alpha) + \sum_i^{n_\beta} (\psi_i^\beta| - \frac{1}{2}\nabla^2|\psi_i^\beta) \tag{3.77}$$

$$E_V = -\sum_A^{nucl} Z_A \int \frac{\rho(\mathbf{r})}{|\mathbf{r} - \mathbf{r}_A|} d\mathbf{r} \tag{3.78}$$

$$J = \frac{1}{2} \int \int \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 \tag{3.79}$$

For the GGA exchange-correlation functional we have

$$E_{xc} = \int f(\rho_\alpha, \rho_\beta, \gamma_{\alpha\alpha}, \gamma_{\alpha\beta}, \gamma_{\beta\beta}) d\mathbf{r} \tag{3.80}$$

$$\gamma_{\alpha\alpha} = |\nabla\rho_\alpha|^2 \quad \gamma_{\alpha\beta} = \nabla\rho_\alpha \cdot \nabla\rho_\beta \quad \gamma_{\beta\beta} = |\nabla\rho_{\beta\beta}|^2$$

where f is a functional only of the local density and its gradients. The one-electron potentials corresponding to (3.80) can be obtained by calculus of variation and are

$$V_\alpha^{xc} = \frac{\partial f}{\rho_\alpha} - 2\nabla \cdot \left(\frac{\partial f}{\partial \gamma_{\alpha\alpha}} \nabla \rho_\alpha \right) - \nabla \cdot \left(\frac{\partial f}{\partial \gamma_{\alpha\beta}} \nabla \rho_\beta \right) \tag{3.81}$$

the expression for V_β^{xc} is similar.

In practical computations, it is convenient to write the orbitals as finite expansions is a basis set ϕ_μ

$$\psi_i^\alpha = \sum_\mu c_{\mu i}^\alpha \phi_\mu \qquad \psi_i^\beta = \sum_i c_{\mu i}^\beta \phi_\mu$$

so that

$$\rho_\alpha = \sum_\mu^N \sum_\nu^N \sum_i^{n_\alpha} (c_{\mu i}^\alpha)^* c_{\nu i}^\alpha \phi_\mu \phi_\nu = \sum_{\mu\nu} P_{\mu\nu}^\alpha \phi_\mu \phi_\nu \tag{3.82}$$

$$\nabla \rho_\alpha = \sum_{\mu\nu} P_{\mu\nu}^\alpha \nabla(\phi_\mu \phi_\nu) \tag{3.83}$$

and similarly for ρ_β.By substituting these expression into energy and then minimizing with respect to unknown coefficients $c_{\mu i}^\alpha$, $c_{\mu i}^\beta$ (subject to orthogonality of ψ_i) we obtain a finite set of algebraic equations for canonical orbitals,

$$\sum_\nu^N (F_{\mu\nu}^\alpha - \epsilon_i^\alpha S_{\mu\nu}) c_{\nu i}^\alpha = 0 \tag{3.84}$$

and similarly for $c_{\nu i}^\beta$. The quantities ϵ_i^α, ϵ_i^β are one-electron eigenvalues for occupied orbitals. The Fock-type matrices $F_{\mu\nu}$ are given by

$$F_{\mu\nu}^\alpha = H_{\mu\nu}^{core} + J_{\mu\nu} + F_{\mu\nu}^{xc_\alpha} \tag{3.85}$$

$$F_{\mu\nu}^\beta = H_{\mu\nu}^{core} + J_{\mu\nu} + F_{\mu\nu}^{xc_\beta} \tag{3.86}$$

Here $S_{\mu\nu}$ and $H_{\mu\nu}^{core}$ are the overlap and bare-nucleus Hamiltonian matrices, respectively, $J_{\mu\nu}$ is the Coulomb matrix

$$J_{\mu\nu} = \sum_{\lambda\sigma}^N P_{\lambda\sigma} (\mu\nu|\lambda\sigma) \tag{3.87}$$

where $P_{\lambda\sigma}$ is the total density matrix $(P^{\alpha}_{\lambda\sigma} + P^{\beta}_{\lambda\sigma})$ and the conventional notation is used for two-electron repulsion integrals. The exchange-correlation parts of the Fock matrices are given by

$$F^{xc_\alpha}_{\mu\nu} = \int \left[\frac{\partial f}{\partial \rho_\alpha} \phi_\mu \phi_\nu + \left(2 \frac{\partial f}{\partial \gamma_{\alpha\alpha}} \nabla \rho_\alpha + \frac{\partial f}{\partial \gamma_{\alpha\beta}} \nabla \rho_\beta \right) \cdot \nabla (\phi_\mu \phi_\nu) \right] d\mathbf{r} \tag{3.88}$$

and similarly for $F^{xc_\beta}_{\mu\nu}$.

3.5 APPLICATION OF HARRIS-FUNCTIONAL APPROXIMATION.

3.5.1 Harris approximation with LDA.

In the Kohn-Sham DFT, the total energy of an N- electron system together with ν nuclei is

$$E = \sum_i n_i \int d\mathbf{r} \psi_i(\mathbf{r}) \left(-\frac{1}{2}\nabla^2 - \sum_{a=1}^{\nu} \frac{Z_a}{|\mathbf{r} - \mathbf{R}_a|} \right) \psi_i(\mathbf{r}) + E_{xc}(\rho) + E_C(\rho) + E_{NN} \tag{3.89}$$

$$\rho(\mathbf{r}) = \sum_i n_i \psi_i(\mathbf{r}) \psi_i(\mathbf{r}) \tag{3.90}$$

$$E_C = \frac{1}{2} \int d\mathbf{r} \int d\mathbf{x} \frac{\rho(\mathbf{r})\rho(\mathbf{x})}{|\mathbf{r} - \mathbf{x}|} \tag{3.91}$$

where Z_a is the positive charge of nucleus a at site $\mathbf{R}_a$, ψ_i is the ith occupied molecular orbital, and its electron occupation number is defined as n_i. E_C is the Coulomb energy, and a direct calculation of it in the basis-set approach requires the evaluation of K^4 two-electron integrals, where K is the number of basis functions. E_{xc} is the exchange-correlation energy which depends in the electron density ρ, and E_{NN} is the nuclear repulsion energy.

Harris approximated [28] E_C and E_{xc} by equations

$$E^H_C = \int d\mathbf{r} \int d\mathbf{x} \frac{\rho(\mathbf{r})\rho^H(\mathbf{x})}{|\mathbf{r} - \mathbf{x}|} - \frac{1}{2} \int d\mathbf{r} \int d\mathbf{x} \frac{\rho^H(\mathbf{r})\rho^H(\mathbf{x})}{|\mathbf{r} - \mathbf{x}|} \tag{3.92}$$

$$E^H_{xc} = E_{xc}(\rho^H) + \int d\mathbf{r} [\rho(\mathbf{r}) - \rho^H(\mathbf{r})] \frac{\delta E_{xc}(\rho^H)}{\delta \rho^H(\mathbf{r})} \tag{3.93}$$

where ρ^H is a superposition of electron densities of atomic fragments. The first-order energy correction of the density error $[\rho(\mathbf{r}) - \rho^H(\mathbf{r})]$ is exactly included in the above equations, while other higher-order corrections are partially canceled and ignored.

The molecular orbital ψ_i then satisfies the equations

$$\left(-\frac{1}{2}\nabla^2 - \sum_a \frac{Z_a}{|\mathbf{r} - \mathbf{R}_a|} + \int d\mathbf{x} \frac{\rho^H(\mathbf{x})}{|\mathbf{r} - \mathbf{x}|} + V[\rho^H(\mathbf{r})]\right) \psi_i(\mathbf{r}) = \epsilon_i \psi_i(\mathbf{r}) \tag{3.94}$$

$$V[\rho^H(\mathbf{r})] = \frac{\delta E_{xc}(\rho^H)}{\delta \rho^H(\mathbf{r})} \tag{3.95}$$

In these equations ρ^H does not depend on ψ_i, so that self-consistent iterations are not needed to obtain ψ_i. Consequently, the Harris method is often used for an initial guess to SCF methods like Hartree-Fock and DFT. The above approach uses the local density approximation (LDA) but others have extended the Harris approach to the GGA.

3.5.2 Generalized-gradient approximation.

GGA functionals depend not only on the values of electron density but also on its gradient. This approximation has been shown to give more accurate results.[19, 20] Generally speaking, the total exchange-correlation energy of a spin-unpolarized state is obtained as:

$$E_{xc} = E_x(\rho, |\nabla\rho|) + E_c(\rho, |\nabla\rho|) = \int d\mathbf{r} \Omega(\rho, |\nabla\rho|) \tag{3.96}$$

In the total exchange-correlation-energy formula (3.96), if the terms of order of $(\rho - \rho^H)^2$, $(\nabla\rho - \nabla\rho^H)^2$, $(\rho - \rho^H)(\nabla\rho - \nabla\rho^H)$, and the higher-order terms are neglected, we obtain the following exchange-correlation form in the Harris-GGA approximation [29]:

$$E_{xc}^H = E_{xc}(\rho^H, |\nabla\rho^H|) + \int d\mathbf{r}[\rho(\mathbf{r}) - \rho^H(\mathbf{r})] \frac{\partial\Omega(\rho^H, |\nabla\rho^H|)}{\partial\rho^H(\mathbf{r})} + \tag{3.97}$$

$$\int + d\mathbf{r}[\nabla\rho(\mathbf{r}) - \nabla\rho^H(\mathbf{r})] \frac{\partial\Omega(\rho^H, |\nabla\rho^H|)}{\partial(\nabla\rho^H(\mathbf{r}))}$$

The above equation can be modified to the following equation by partial integration of the third term:

$$E_{xc}^H = E_{xc}(\rho^H, |\nabla\rho^H|) + \int d\mathbf{r}[\rho(\mathbf{r}) - \rho^H(\mathbf{r})] V(\rho^H) \tag{3.98}$$

where

$$V(\rho^H) = \frac{\partial \Omega(\rho^H, |\nabla \rho^H|)}{\partial \rho^H(\mathbf{r})} - \nabla \left(\frac{\partial \Omega(\rho^H, |\nabla \rho^H|)}{\partial (\nabla \rho^H(\mathbf{r}))} \right) \tag{3.99}$$

3.5.3 Using the Harris approximation in new ways.

Originally, Harris applied his method in a context of atoms. In other words, one would consider ρ^H as the electron density of isolated atoms assembled into a molecule. Instead, we propose to use Harris approximation within a larger framework. Imagine that someone needs to calculate the binding curve between two fragments. Those can be large molecules. We propose to use the Harris approximation by creating $\rho^H = \rho_{M1} + \rho_{M2}$ where ρ_{M1} and ρ_{M2} are electron densities of molecules computed using full SCF DFT. The reuse of information in such a situation is enormous. Instead of running the full DFT for a set of fragments, one can easily reuse the electron density of those pieces and not the atomic densities. This should significantly speed up the calculation of the energies. Similar ideas are in use in other techniques such as SAPT.

We also propose to correct the original Harris energies by adding an empirical dispersion correction [15]. The dispersion correction plays a very important role when looking at π-stacking. In chemistry, π- stacking refers to attractive, non-covalent interactions between aromatic rings.

We use the dispersion correction of the form proposed by Grimme [15]:

$$E_D = -\sum_i \sum_j \frac{C_{ij}}{r^6} f(r_{ij}) \tag{3.100}$$

$$f(r_{ij}) = \frac{1}{1 + e^{-d(r_{ij}/R_{ij} - 1)}} \tag{3.101}$$

where $d = 23$, $f(r_{ij})$ is the damping function and R_{ij} is the sum of the van-der-Waals radii of the corresponding atoms.

3.6 RESULTS AND DISCUSSIONS.

3.6.1 Results.

In this work, we do not attempt to present a full picture of all types of interactions involved. Instead, we compare the results obtained by our method with a set of well-known functionals. More specifically, we compare Harris+LDA+D and Harris+GGA+D with PW91 [38, 39, 40, 41, 42], PBE [19, 20], M062X [21, 22, 23, 24], B3LYP [25] and MP2 [26, 27]. We consider M062X and MP2 as a benchmark. To generate binding curves from the latter methods we use the Gaussian 09 [30] program. Harris based methods were implemented by modifying the Erkale package. [31, 32] All calculations, including Harris based ones, were done with a cc-pVDZ [33, 34, 35, 36, 37] basis set. Harris+LDA+D is computed based on Dirac exchange[17] and Perdew Wang correlation [18] functionals and Harris+GGA+D is derived from PBE [19, 20] exchange-correlation functional. The initial electron densities for each molecule were generated using B3LYP functional with cc-pVDZ basis set. We found that the final results are not sensitive to the choice of a method used to produce the initial guess for each fragment.

To study the performance of our method, we look at the shapes of binding curves and the locations of minima which are produced. We examine systems based on various orientations of two benzene and two thiophene molecules. The sandwich configuration was created by stacking of two molecules without performing any rotations. The parallel-displaced configuration was created by shifting one of molecules horizontally at separation of 3.5 Å. Rotations of thiophene molecule were performed at the distance of 5 Å from the initial position where two thiophene molecules are parallel to each other and one of them in rotated by 180 degrees around the vertical axis going through the geometric center. All curves are shifted to have zero value of energy at infinity. Results are obtained by shifting molecules by a tenth of an angstrom in case of sandwiched, T-shaped and parallel-displaced orientations and by rotating by 3 degrees in case of clockwise and counterclockwise rotations.

Additionally we compute the maximum and average total energy deviation of Harris+GGA+D method from PW91, PBE, B3LYP, M062X, and MP2 to demonstrate the

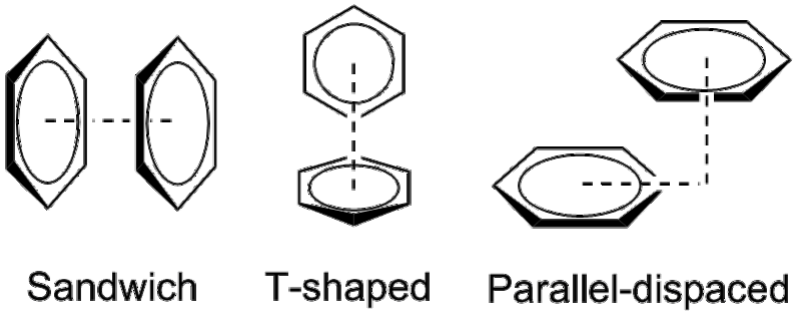


Figure 3.1: Orientations

proximity of the binding curves Tab. 3.1,3.3 as well as locations of minima across all methods presented in this work Tab.3.2,3.4. Additional digits in the minima locations were obtained by spline fitting of the binding curves and solving for a minimum.

3.6.2 Discussions.

As we mentioned above, we judge the effectiveness of our approach by examining its proximity to the benchmark curves generated with MP2 and M062X functionals across different configurations. For the case of benzene we examine the sandwich, T-shaped and parallel-displaced orientations, while in case of thiophene we look at the sandwich, parallel-displaced orientations and rotations. We perform a 360 degrees rotation of a thiophene molecules at the distance of 5 Å starting from a sandwich orientation when sulfur atoms point in opposite directions. When sulfur atom goes up, we call it a clockwise rotation. When it goes down, we call it counterclockwise rotation Fig. 3.8

We first examine closely the binding curves for the benzene dimer. The performance of the Harris approximation in case of the sandwiched configuration is excellent. The Harris+GGA+D curve almost exactly follows the MP2 curve giving an average deviation of 0.18 kcal/mol. The M06-2X curves is slightly shifted upward which results into a higher average deviation. The minimum of the Harris+GGA+D curve almost exactly coincides with the minimum of MP2 and is slightly smaller than the minimum of M06-2X. A similar behavior

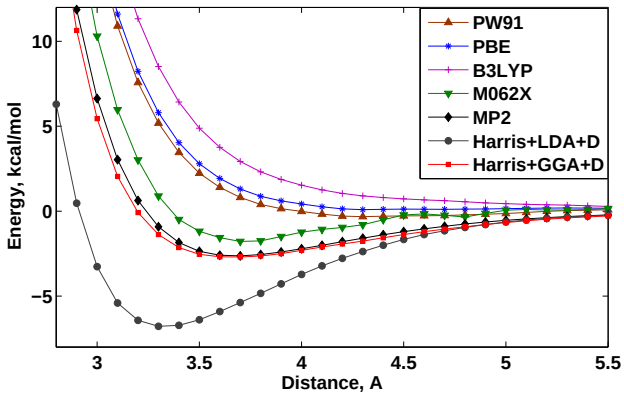


Figure 3.2: Sandwich configuration of benzene

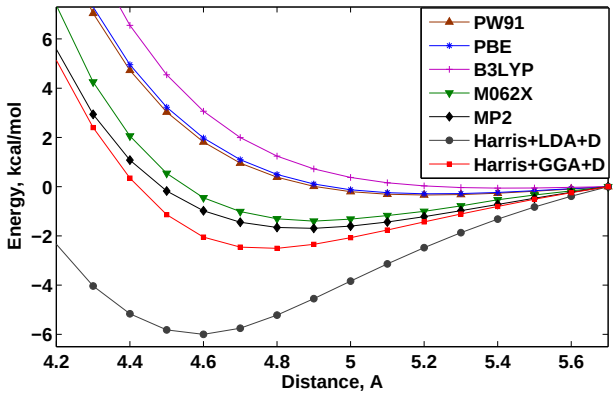


Figure 3.3: T-shaped configuration of benzene

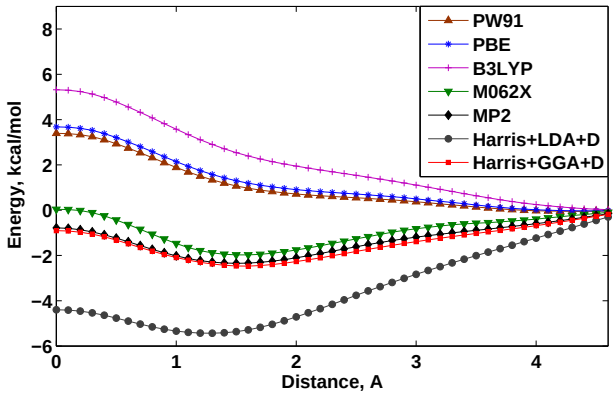


Figure 3.4: Parallel-displaced configuration of benzene at separation of 3.5 Å

	Sandwich		T-shaped		Parallel-displaced	
Method	Max	Average	Max	Average	Max	Average
PW91	12.02	2.00	13.17	5.50	4.30	2.39
PBE	12.89	2.26	12.92	5.71	4.58	2.55
B3LYP	18.41	3.13	42.49	11.81	6.22	3.41
M062X	6.78	1.01	32.02	6.36	0.95	0.51
MP2	1.22	0.18	13.69	1.95	0.21	0.12

Table 3.1: Maximum and average total energy deviation (kcal/mol) of Harris+GGA+D vs other methods applied to benzene.

Table 3.2: Locations of minima of binding curves of benzene.Å

Orientation/Method	Sandwich	T-shaped	Parallel-displaced
PW91	4.30	5.20	-[a]
PBE	4.31	5.22	-[a]
B3LYP	-[a]	-[a]	-[a]
M062X	3.73	4.89	1.54
MP2	3.67	4.87	1.50
Harris+LDA+D	3.32	4.58	1.27
Harris+GGA+D	3.66	4.76	1.56

[a] No minimum;

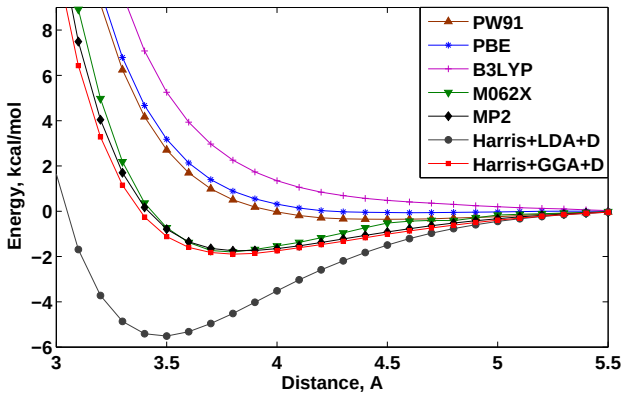


Figure 3.5: Sandwich configuration of thiophene

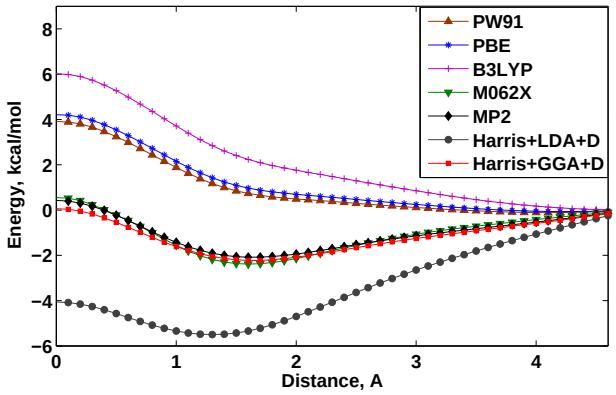
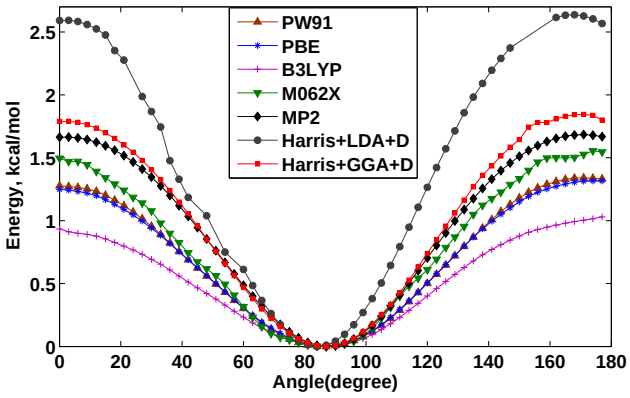


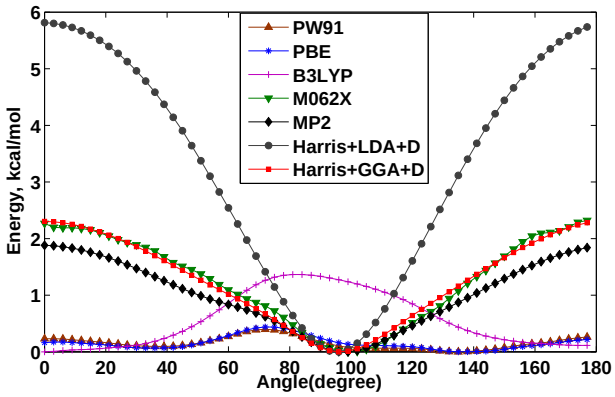
Figure 3.6: Parallel-displaced configuration of thiophene at separation of 3.5 Å

is observed in case of the T-shaped and the parallel-displaced configurations. The former having higher average deviations. The location of the minimum of Harris+GGA+D is about 0.1Å off compared to M06-2X and MP2. Finally, the parallel-displaced configuration exhibits similar properties by having even very small average deviation and the minimum being closer to M06-2X result. Overall, we observe very good agreement of the Harris+GGA+D method compared with the benchmark results. Needless to say, that other functional that we have used give a very different performance. In case of the sandwiched configuration PW91 and PBE give minima at very long distances, while B3LYP does not give any minimum at all. The above statement is also true for the T-shaped configuration. For the parallel-displaced arrangement neither of these functionals give any binding.

A similar behavior is observed for the thiophene dimer. The sandwiched and the parallel-displaced configurations exhibit similar trends, with the minima being very close to the benchmark results. Again, functionals like PW91,PBE, and B3LYP either give minima at very long distances (in case of the sandwiched configuration) or do not show any binding (in case of the parallel-displaced configuration). In addition to translations we perform the rotation of one of the thiophene molecules. In case of the clockwise rotation Harris+GGA+D follows closely the curve generated by MP2. One should notice a similar behavior of all functional in this situation. On the other hand, counterclockwise rotation is describes differently



(a) Clockwise rotation of thiophene



(b) Counterclockwise rotation of thiophene

Figure 3.7: Rotation performed at separation of 5.0 Å

Table 3.3: Maximum and average total energy deviations (kcal/mol) of Harris+GGA+D vs other methods applied to thiophene, for all points on the illustrated binding curves.

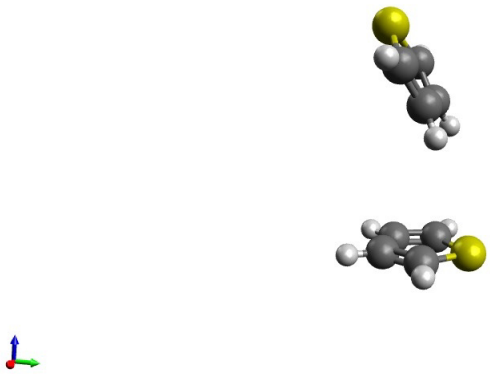
	Sandwich		Parallel-displaced		CW[a] rotation		CCW[b] rotation	
Method	Max	Average	Max	Average	Max	Average	Max	Average
PW91	9.29	2.43	3.85	2.04	0.51	0.30	2.07	1.11
PBE	10.09	2.78	4.15	2.21	0.55	0.31	2.14	1.13
B3LYP	16.26	4.26	5.95	3.12	0.88	0.48	2.30	1.22
M062X	5.80	0.84	0.49	0.16	0.36	0.20	0.15	0.07
MP2	2.17	0.35	0.36	0.14	0.18	0.06	0.44	0.25

[a] Clockwise; [b] Counterclockwise;

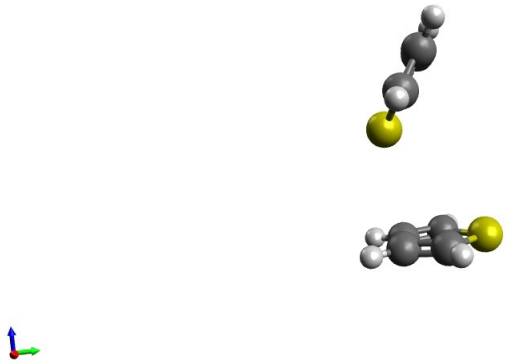
Table 3.4: Locations of minima of binding curves of thiophene.

Orient./Meth.	Sandwich(Å)	PD[b](Å)	CWR[c] (°)	CCWR[d] (°)
PW91	4.51	-[a]	86.00	134.85
PBE	4.61	-[a]	86.14	136.72
B3LYP	-[a]	-[a]	86.98	-[a]
M062X	3.78	1.60	86.11	99.73
MP2	3.83	1.63	87.43	97.92
Harris+LDA+D	3.47	1.30	84.29	95.34
Harris+GGA+D	3.80	1.64	87.09	96.40

[a] No minimum; [b] Parallel-displaced; [c] Clockwise rotation; [d] Counterclockwise rotation;



(a) Clockwise rotation



(b) Counterclockwise rotation

Figure 3.8: Rotations. Distance between molecules is 5.0 Å

by different functionals. As before, Harris+GGA+D closely follows the benchmark curves, while PW91, PBE, and B3LYP give qualitatively different curves. The locations of minima produced by Harris+GGA+D are in a very good agreement with the benchmark results.

By taking a quick look at the binding curves one can immediately notice the bad performance of the Harris+LDA+D method. It consistently overbinds and produces very deep potential wells. We are not going to discuss this combination further because it obviously stands on its own. This is one of the reasons why we do not present any comparisons of Harris+LDA+D with other methods.

On the other hand, Harris+GGA+D gives good results for all instances. A closer look at the Tab.3.1,3.3 and Tab.3.2,3.4 can clearly demonstrate this. Firstly, we examined the maximum and average deviation of the Harris+GGA+D method from the other functionals. We present maximum deviation for completeness here. Some of them are not visible in the plots of binding curves and occur at 2.8Å in case of sandwich and T-shaped orientations. A more interesting measure of proximity is the average deviation of the curves. In case of Harris+GGA+D vs M062X and MP2 we observe an excellent proximity across all orientations. Secondly, we obtained the minima by spline smoothing of the original data and the results are given in Tab. 3.2,3.4. Again, we observe a good agreement of Harris+GGA+D with our benchmark methods.

An interesting feature is observes when counterclockwise rotations are performed with thiophene molecules. We observe a drastic disagreement of B3LYP, PW91 and PBE vs M062X, MP2. On the other hand, Harris+GGA+D is in excellent agreement with our benchmark functionals as one can see by examining the average deviation and the locations of minima. This proves the value of our approach.

3.6.3 Timing results.

In this section we present parallel speed-up 9(a),9(b) and relative timing 3.10 results of our code. We run GGA calculation with PBE exchange-correlation functional and LDA calculation with Dirac exchange and PW91 correlation. We scale up the calculation up to 24 processors via OpenMP. The system under study here is a benzene dimer. This

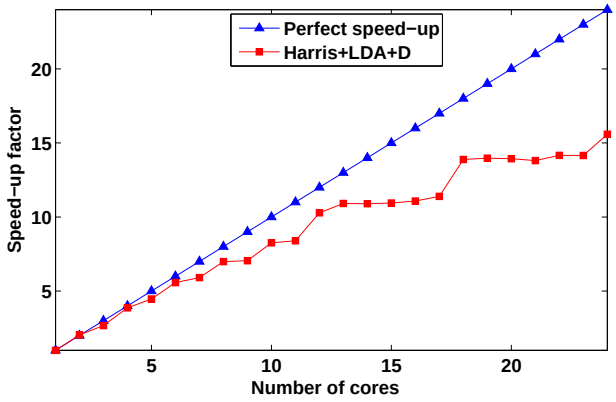
system is perfect for this sort of studies because it contains 24 atoms and 24 is divisible by 2,3,4,6,8,12,24. This can be explained by the fact the the code uses grid construction technique where the latter is separately generated for each atom. Likewise, the routines which compute the total energy also do it in atom by atom manner. This way each thread created by openMP is able to process atoms individually. The most computationally expensive part of the calculation here is the adaptive grid generation. The results show that the speed-up factor is close to perfect when up to 12 processors are used. When the number of CPUs exceeds 12 we observe a steady deterioration of the performance which may be explained by the fact that not all atoms are processed by threads simultaneously. For example, in case of 11 CPUs we have 2 atoms left for the third round of grid generation. In particular, this explains why the speed-up curve is essentially flat in the regions between 12 and 17 CPUs and 18 and 23 CPUs. Obviously, one needs more time to generate grid for a carbon atom that for a hydrogen atom. This results in a slight imbalance.

We also demonstrate timing results of Harris+GGA+D vs full SCF GGA calculation (as described above). We run the code with 3 different basis sets: cc-pVDZ, aug-cc-pVDZ, and cc-pVTZ respectively. We also run the Harris approximated calculation on a fine and a coarse grids. The bar plots Fig. 3.10 presented below shows the results of this comparison. It is very clear that this approximation does indeed bring a substantial acceleration when compared to the full SCF and clearly demonstrate the utility of the Harris approximation. When run on a fine grid we get 6.4,11.1, and 6.4 acceleration respectively when compared to the full SCF method. Higher speed-up is attained when we run our code on a coarse grid. This brings 16.4,27.6 and 15.7 acceleration respectively.

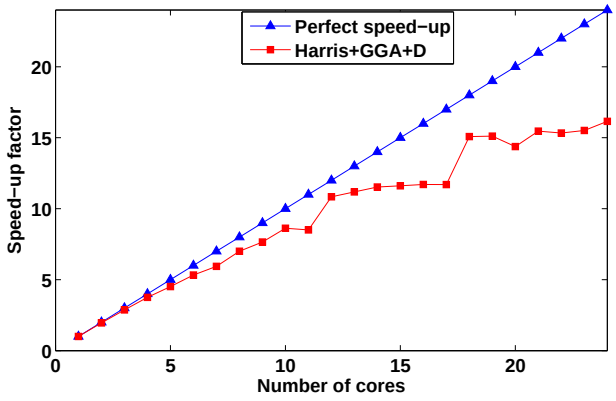
One major drawback of the Harris based methodology is that it lacks an analytical expression for the gradient of energy with respect to molecular coordinates. Besides numerical gradients the only possible way of minimizing the energy of any system in such a setup can be a stochastic algorithm such as simulated annealing, parallel-tempering etc. Lastly, one needs to be able to rotate the density matrix of a molecule when requested. In case of a stochastic optimizer one needs to compute the energy of a proposed state and then either accept it or reject. That requires knowledge of the density matrix at all possible orientations.

Although, Harris+GGA+D methods produce accurate binding curves together with dis-

persion correction, they will most likely fail to account for any sort of polarization effects. Molecules with high dipole moment will require additional energy corrections.



(a) Speed-up curve for Harris+LDA+D



(b) Speed-up curve for Harris+GGA+D

Figure 3.9: Speed-up curves

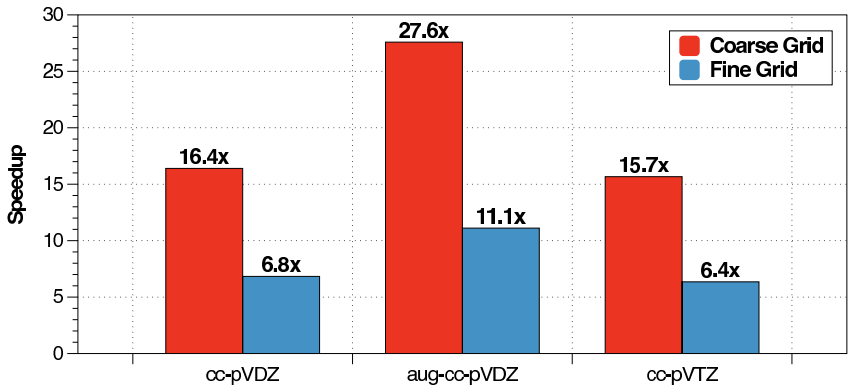


Figure 3.10: Relative acceleration with different basis sets and discretizations.

4.0 MOLECULAR PACKING OPTIMIZATION.

As we discussed in the previous section, Harris approximation happens to work very well in a number of systems. The accuracy of the method suggests that it can be used for molecular packing optimization. Ideally, one can use the Harris approximation in order to quickly find an approximate minimum, and then adjust the geometry by means of more accurate approaches. In this section, we discuss how one can run the geometry optimization with the Harris approximation, and give an example of optimizing the benzene dimer. We will apply an unconventional minimization technique called Simulated Perturbation Stochastic Approximation (SPSA) [112]. We are using a stochastic algorithms because of the lack of an analytic expression for the gradient of the total energy with respect to the positions of the nuclei. This prevents us from using deterministic local gradient based optimizers such as the gradient descent, etc. Instead, we choose to use a stochastic algorithm which also takes a step using a similar update as the deterministic gradient descent, but the direction of this step is now determined stochastically.

4.1 SIMULATED PERTURBATION STOCHASTIC APPROXIMATION ALGORITHM.

4.1.1 Basic algorithm.

We assume that no direct measurements of the gradient $g(\theta)$ are available. The basic unconstrained SPSA algorithm [112] in the general recursive form looks like:

$$\hat{\theta}_{k+1} = \hat{\theta}_k - a_k \hat{g}_k(\hat{\theta}_k) \tag{4.1}$$

where $\hat{g}_k(\theta_k)$ is the simultaneous perturbation estimate of the gradient at the iterate $\hat{\theta}_k$ based on the measurements of the loss function and a_k is a nonnegative scalar gain coefficient. The loss function in the current context is equivalent to the energy the system.

The essential part of of (4.1) is the gradient approximation $\hat{g}_k(\hat{\theta}_k)$. With simultaneous perturbation, all elements of $\hat{\theta}_k$ are randomly perturbed together to obtain two loss measurements $y(\cdot)$. For the two-sided gradient approximation, this leads to

$$\hat{g}_k(\hat{\theta}_k) = \begin{bmatrix} \frac{y(\hat{\theta}_k+c_k\Delta_k)-y(\hat{\theta}_k-c_k\Delta_k)}{2c_k\Delta_{k1}} \\ \vdots \\ \frac{y(\hat{\theta}_k+c_k\Delta_k)-y(\hat{\theta}_k-c_k\Delta_k)}{2c_k\Delta_{kp}} \end{bmatrix} = \tag{4.2}$$

$$= \frac{y(\hat{\theta}_k + c_k\Delta_k) - y(\hat{\theta}_k - c_k\Delta_k)}{2c_k}[\Delta_{k1}^{-1}, \Delta_{k2}^{-1}, ..., \Delta_{kp}^{-1}]^T$$

where the mean-zero p-dimensional random perturbation vector, $\Delta_k = [\Delta_{k1}, \Delta_{k2}, ..., \Delta_{kp}]^T$, has a user-specified distribution with finite inverse moments. Because the numerator is the same in all p components of $\hat{g}_k(\hat{\theta}_k)$, the number of loss measurements needed to estimate the gradient is SPSA is two, regardless of the dimension p.

This can be contrasted with the conventional finite difference approximation of the gradient where one needs $2p$ loss function evaluations: two for each degree of freedom.

$$\hat{g}_k^{FD}(\hat{\theta}_k) = \begin{bmatrix} \frac{y(\hat{\theta}_{k1}+c_k\Delta_{k1})-y(\hat{\theta}_{k1}-c_k\Delta_{k1})}{2c_k\Delta_{k1}} \\ \vdots \\ \frac{y(\hat{\theta}_{kp}+c_k\Delta_{kp})-y(\hat{\theta}_{kp}-c_k\Delta_{kp})}{2c_k\Delta_{kp}} \end{bmatrix} \tag{4.3}$$

The step-by-step implementation summary below shows how SPSA iteratively produces a sequence of estimates that progressively minimizes the function.

4.1.2 Basic SPSA algorithm.

- **Initialization and coefficient selection**. Set iteration counter $k = 0$. Pick the initial guess $\hat{\theta}_0$ and nonnegative coefficients a, c, A,α, and γ in the SPSA gain sequences $a_k = a/(k+1+A)^\alpha$ and $c_k = c/(k+1)^\gamma$. Practically effective values for α and γ are 0.602 and 0.101, respectively [112]; other parameters are discussed below.
- Generation of the simultaneous perturbation vector. Generate by Monte-Carlo a p-dimensional random perturbation vector Δ_k, where each of the p components are independently generated from a zero-mean probability distribution with finite inverse moments. An effective choice for each component of Δ_k is to use a Bernoulli ± 1 distribution with probability of 1/2 for each outcome, although other choices are valid and be desirable in some applications.
- Loss function evaluation. Obtain two measurements of the loss function based on the simultaneous perturbation around the current $\hat{\theta}_k$: $y(\hat{\theta}_k + c_k\Delta_k)$ and $y(\hat{\theta}_k - c_k\Delta_k)$.
- Gradient approximation. Generate the simultaneous perturbation approximation to the unknown gradient $g_k(\hat{\theta}_k)$ according to (4.2). It is sometimes useful to average several gradient approximations at $\hat{\theta}_k$, each formed from an independent generation of Δ_k. The benefits are especially apparent if the noise effects are relatively large.
- Update θ estimate. Use the standard SA form (4.1) to update $\hat{\theta}_k$ to a new value $\hat{\theta}_{k+1}$.
- Iteration or termination. Return to step 1 with $k+1$ replacing k. Terminate the algorithm if there is little change in several successive iterates or if the maximum allowable number of iterations has been reached. The algorithm can also be terminated if the gradient is vanishingly small, or by meeting multiple conditions.

4.1.3 Choice of gain sequence.

The choice of the gain sequence is critical to the performance. With α and γ as specified in the description of the algorithm above, one typically finds that in a high-noise setting it is necessary to pick a smaller a and larger c than in a low-noise setting. The asymptotically optimal values of α and γ with noisy loss measurements are 1 and 1/6, respectively [112]. In practice, however, it is usually the case that $\alpha < 1$ yields better finite-sample performance

through maintaining a larger step size. Hence the recommendation is to use values (0.602 and 0.101) that are effectively the lowest allowable subject to satisfying certain theoretical conditions. When the algorithm is being run with a larger number of iterations, it may be beneficial to convert to $\alpha = 1$ and $\gamma = 1/6$ at some point in the iteration process to take advantage of the asymptotic optimality.

With the Bernoulli ± 1 distribution (this distribution has finite inverse moments) for the elements of Δ_k and the α and γ specified, a rule of thumb is to set c at a level approximately equal to the standard deviation of the measurement noise in $y(\theta)$. This helps keep the p elements of $\hat{g}_k(\hat{\theta}_k)$ from getting excessively large in magnitude. The standard deviation can be estimated by collecting several $y(\theta)$ values at the initial guess $\hat{\theta}_0$. When perfect measurements are available, then c should be chosen as some small positive number.

The values of a,A can be chosen together to ensure effective practical performance of the algorithm. A useful rule of thumb is to choose $A > 0$ such that it is 10 percent or less of the maximum number of expected/allowed iterations. After choosing A, one can choose a such that $a_0 = a/(1+A)^{0.602}$ times the magnitude of the elements in $\hat{g}_0(\hat{\theta}_0)$ is approximately equal to the smallest of the desired change magnitudes among the elements of θ in the early iterations.

4.2 RESULTS AND DISCUSSIONS.

In this section, we consider some results obtained for the molecular packing optimization of the benzene dimer that we have studied in detail in the previous chapter. In order to find a configuration of these molecules with the minimum of energy, we choose to move just one of the monomers relative to the other. This avoids the recalculation of the density matrix for the other monomer, which in turn increases the performance of our code. One might think that we have introduced an unnecessary constraint into the system, but as we will see later, it does not introduce any biases into the system, and the algorithm finds the correct minima.

Additionally, one should be aware that in the present situation we perform the optimization with only 6 degrees of freedom. This comes from the fact that we keep our molecules rigid

and do not relax the actual nuclei positions. The degrees of freedom include 3 translational variables (x,y,z) which correspond to the coordinates of the geometric center (the molecule is highly symmetric) and 3 rotational variables which correspond to rotations around the geometric center. This setup is different from the conventional analytic gradient bases optimizers where the gradient of the energy with respect to each nucleus position is readily available. We present some optimization results below. We consider 3 different situations. First we constraint the system to translations only in the z direction and try to assess its convergence speed compared to the ADF package [114] simulation of the same system. Next, we relax the above constraint to include x and y translations (rotations are forbidden). Finally, we let the system to optimize without any constraints. Each of the above cases will be discussed separately.

4.2.1 The gain sequence choice.

The above description of the algorithm is deceptively simple. Indeed, we propose two possible orientations of the the monomer, compute the gradient based on the values of the energies obtained and update the coordinates. As discussed above,the most important ingredient of this algorithm is the gain sequence. One may falsely assume that since the algorithm is base on the gradient descent idea, then it should be local. In other words, being able to converge to the local minimum is the only possibility. It has been proven that the SPSA procedure can also be used as a global optimizer [113]. This is a very important feature because being able to locate the global minimum is a key goal, avoiding local barriers if desired.

In the last section we will show that the careful choice of the gain sequence is critical for the performance of the algorithm. We used different gain sequences for the different cases that we have studied. The convergence is assessed by examining the energy change from iteration to iteration. The maximum number of iterations is set to be intentionally large to see the evolution of the system. Our experiments show that running the simulation for $500 - 600$ iteration is more than enough to find a minimum.

One last important point which is worth discussing here is the initial orientation - an initial guess. This factor plays a very important role here and has a direct influence on

the choice of the gain sequence. In case of the sandwiched and T-shaped orientations fig. 3.2,3.3, the binding curve becomes very steep at short distances. This results in very high values of the gradient which in turn may cause substantial coordinate updates during the first few iterations. On the other hand, the valleys of those curves are much shallower. The problem is that all of the interesting things happens exactly there. If we attempt to place the monomers at the distance of 2.8Å and pick a relatively large value for a, then the system will jump too far from the the minimum and may not be able to converge at all. On the other hand, assigning small values to a will results in a very,very slow convergence in the regions close to the minimum. Although, any optimization code should be able to handle different situations, we suggest having reasonable initial orientations where the energy change from one configuration to another does not exceed 0.1 kcal/mol. We set $A = 200$,$c = 0.2$ for all our simulations and use the suggested values for α and γ. The value of a will differ across our simulations.

4.2.2 Vertical translations.

In this section we consider only vertical translations figures 4.1 and try to test whether the algorithm is able to find the minimum. Indeed, we know the answer beforehand from Table 3.2 which is 3.66Å. We consider two initial orientations: the sandwiched one at separation of 2.8Å and the t-shaped one at separation of 3.0Å. We set the value of $a = 50$. We wish to follow the general convergence and compare the number of iterations requested to reach a certain convergence criterion. We compare the performance of our approach to the conventional gradient optimization run with ADF software package [114]. The convergence criterion considered here is the energy change from iteration to iteration. The tolerance of 10^{-4} Hartree is achieved within 32 iteration using SPSA and within 16 iterations by the conventional gradient approach. In case of the t-shaped optimization SPSA took 8 iterations and the conventional method took 46 iterations. Indeed, one may be able to achieve convergence even faster if a different gain sequence is used.

This tells us that the SPSA algorithm is suitable for this kind of problems and can be a very promising alternative. As we expected, the minima are attained at 3.66Å for

the sandwiched configuration and at 4.76Å for the t-shaped configuration. SPSA would compare to numerical estimates of the gradient,e.g. by finite differences,requiring two energy evaluations per degree of freedom, versus only two for SPSA. Thus SPSA compares well to gradient methods.

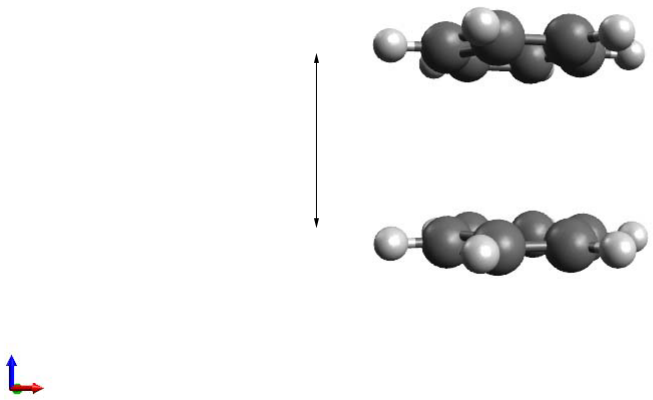
4.2.3 Simulation with translations only.

In this section we try to run the SPSA algorithm with $a = 50$ as in the previous section. Now we let the second monomer not only move along the z axis but also along the x and y axes. This can be viewed as a trial run before we remove any constraints and optimize along 3 axes and 3 rotations. We run the simulation for the sandwiched and the parallel-displaced initial orientations. The sandwiched configuration is set up, as in the previous section where the monomers were located on top of each other at distance of 2.8Å. The parallel-displaced configuration was obtained from the sandwiched one by shifting the second monomer by 3.0Å horizontally and by 0.7Å vertically placing it at 3.5Å from the first monomer. We will not discuss the t-shaped orientation because it happens to be very stable and the simulation simply shifts it upward.

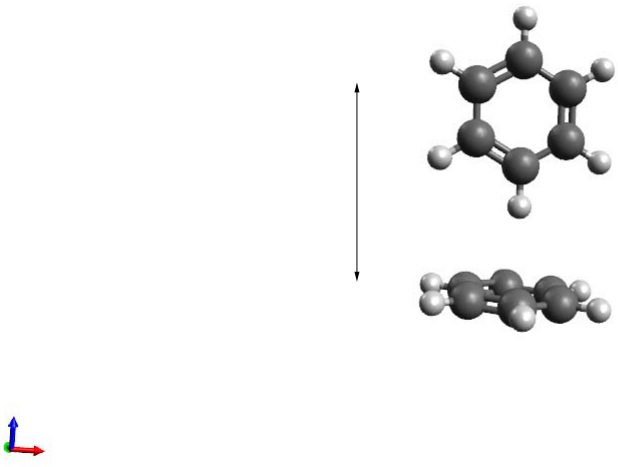
We run the simulation with 250 iterations and observe that both initial orientations lead to the final configuration which is similar to the initial parallel-displaced orientation. The initial coordinate of the geometric center of the sandwiched configuration was $(0.0, 0.0, 2.8)$ Å and the final configuration is located at $(1.0470, -1.2482, 3.25)$ Å. The initial coordinate of the parallel-displaced monomer is configuration was at $(3.0, 0.0, 3.5)$ Å and the final coordinate is $(0.9735, 0.0253, 3.23)$ Å. Longer runs should ultimately produce the same parallel-displaced configuration. The results are shown in figures 4.2,4.3,4.4,4.5

4.2.4 Optimization with translations and rotations.

Finally we attempt to run the optimization algorithm with 6 degrees of freedom. As we mentioned above, one should focus on the choice of the gain sequence and to the initial orientations of the molecules. In this simulation we create the sandwiched configuration by placing two benzene monomers at a distance of 3.5Å. The initial parallel-displaced

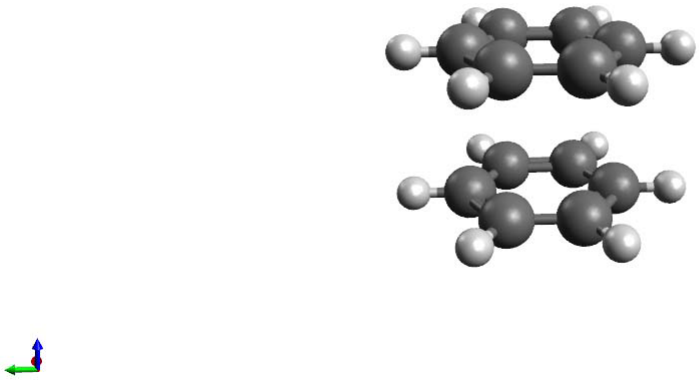


(a) Sandwiched configuration

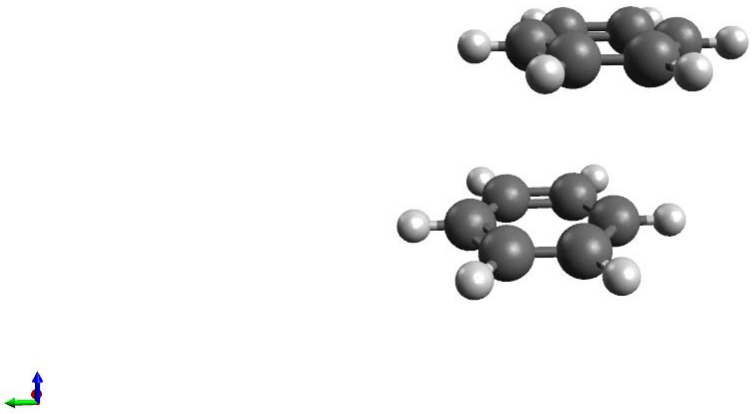


(b) T-shaped configuration

Figure 4.1: Vertical translations

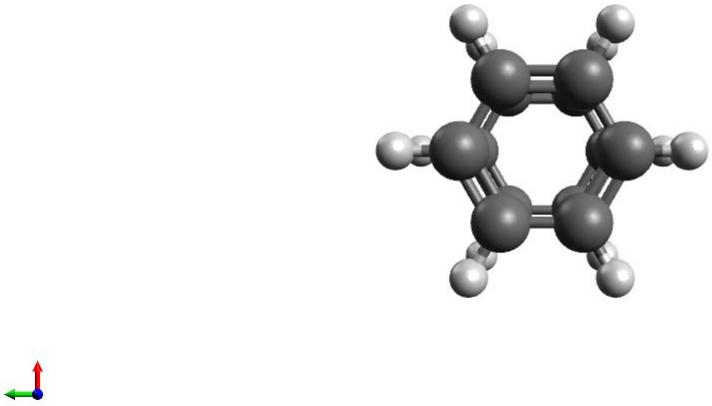


(a) Initial sandwiched configuration

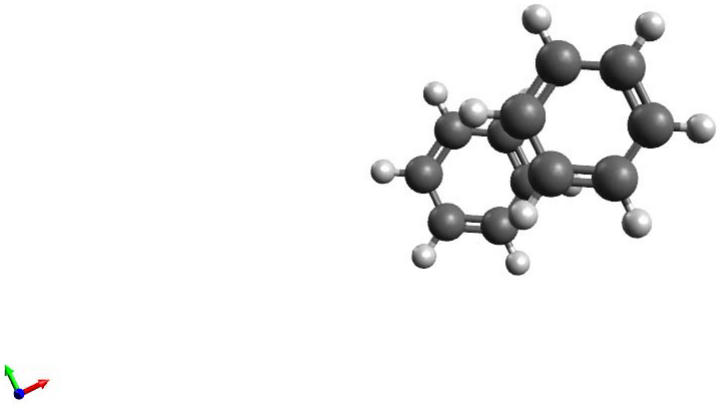


(b) Final sandwiched configuration

Figure 4.2: Sandwiched configuration optimization

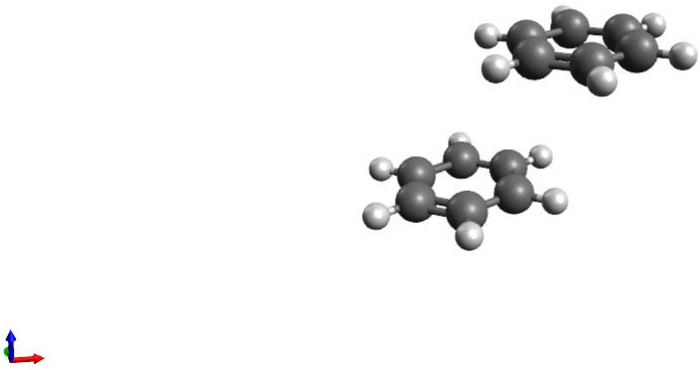


(a) Initial sandwiched configuration

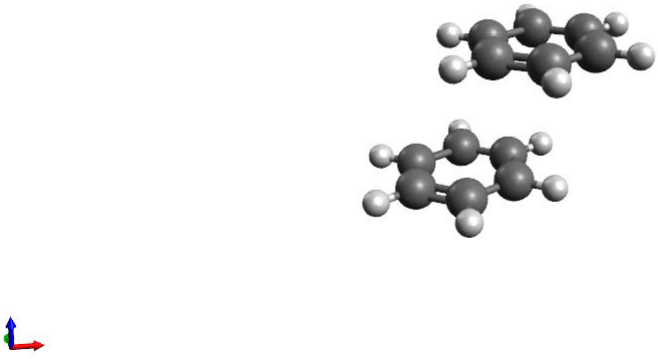


(b) Final sandwiched configuration

Figure 4.3: Sandwiched configuration optimization(view from the top)

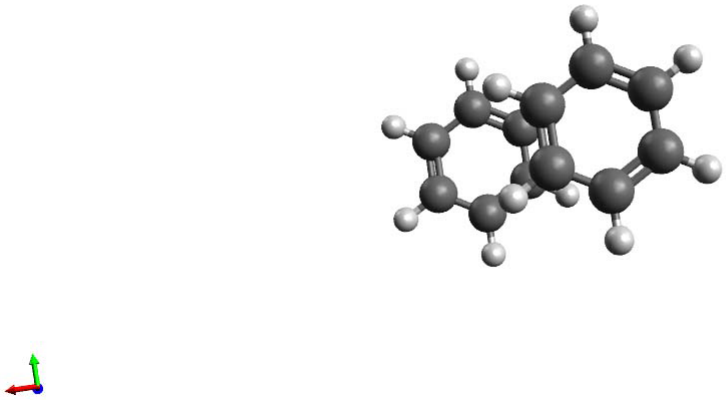


(a) Initial parallel-displaced configuration

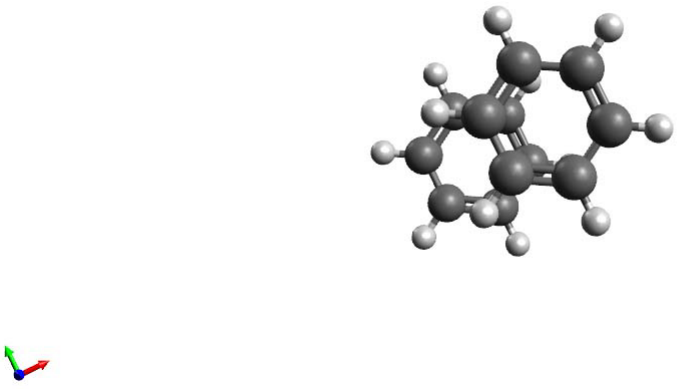


(b) Final parallel-displaced configuration

Figure 4.4: Parallel-displaced configuration optimization



(a) Initial parallel-displaced configuration



(b) Final parallel-displaced configuration

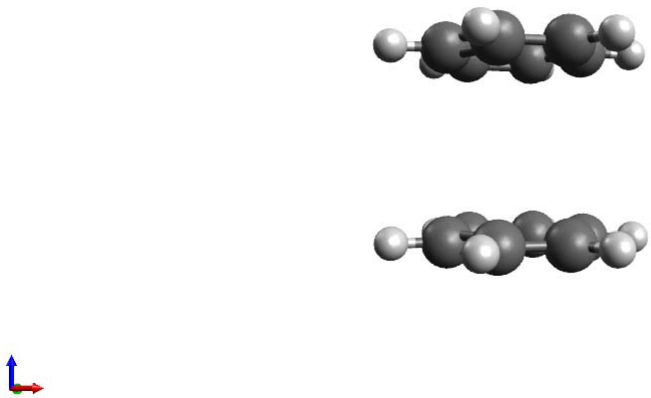
Figure 4.5: Parallel-displaced configuration optimization (view from the top)

configuration is the same as in the previous section. The t-shaped configuration is now such that the geometric center of the second monomer is 5.0Å over the first monomer. We choose the value for the parameter to be $a = 400$.

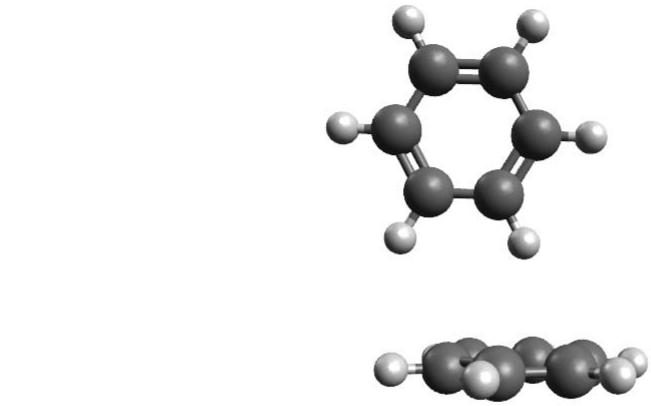
When we run our program with the sandwiched initial configuration, we observe that the second monomer rotates and ends up in a t-shaped -configuration. The parallel-displaced initial configuration results in another parallel-displaced configuration which is located closer to the first monomer. The initial t-shaped configuration does not change almost at all. After 462 iterations we find that the t-shaped orientation generated from the initial sandwiched configuration has centroid 0.1Å higher over the first monomer than the t-shaped configuration generated from the initial t-shaped configuration. As we approach the minimum, the energy change from iteration to iteration does not exceed 10^{-6} Hartree in all of the above cases. It is clear that both should converge to same minimum after running for a longer time.

We have mentioned above, that unlike the standard gradient decent algorithm, the SPSA algorithm is a global optimizer. We have also mentioned that the choice of the gain sequence may directly influence that property. The benzene dimer has two possible metastable configurations with very close values of the energy. Those are the t-shaped and the parallel-displaced orientations. The energy difference between the two configurations is 0.0014 Hartree or 0.8 kcal/mol.In this section we have demonstrated that the SPSA algorithm was able to find both of them. In contrast, when we set $c = 0.1$,$a = 400$ and start from the same initial positions, the algorithm is only able find the parallel-displaced configuration. This fact clearly demonstrates our claim of the sensitivity of the algorithm to the choice of the gain sequence as well as its ability to locate the global minimum. The results are given in Fig. 4.6,4.7,4.8

Furthermore, by setting $c = 0.3$ the algorithm only finds the t-shaped configuration, which corresponds to the overall global minimum. This demonstrates that the proper choice of the gain sequence may result into different final configurations and makes the algorithm very flexible and capable of finding not only the global minimum but also the local minima.

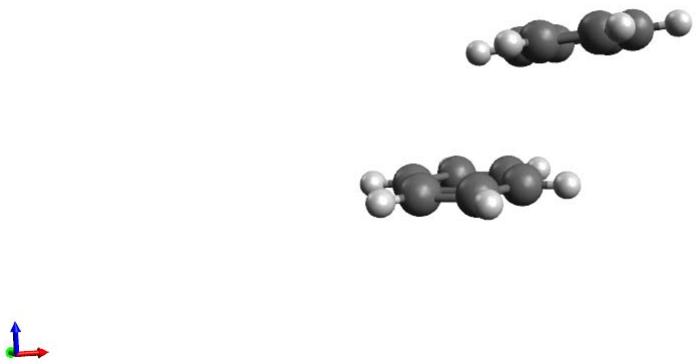


(a) Initial sandwiched configuration

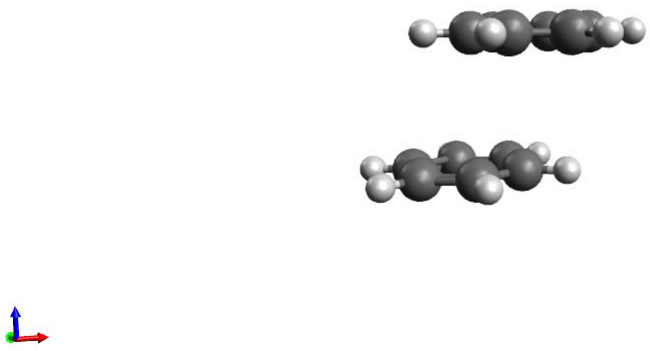


(b) Final sandwiched configuration

Figure 4.6: Sandwiched configuration optimization with 6 degrees of freedom

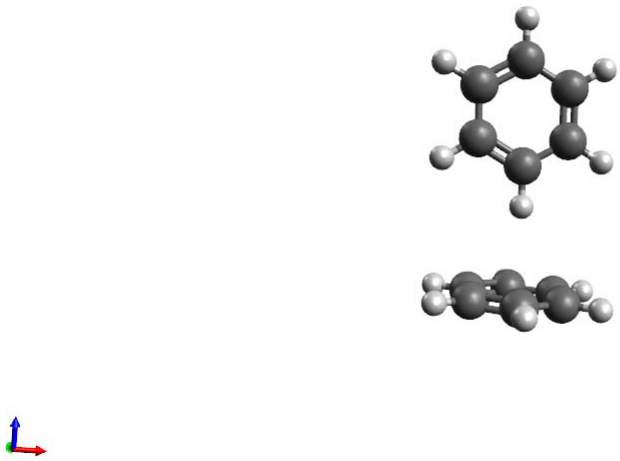


(a) Initial parallel-displaced configuration

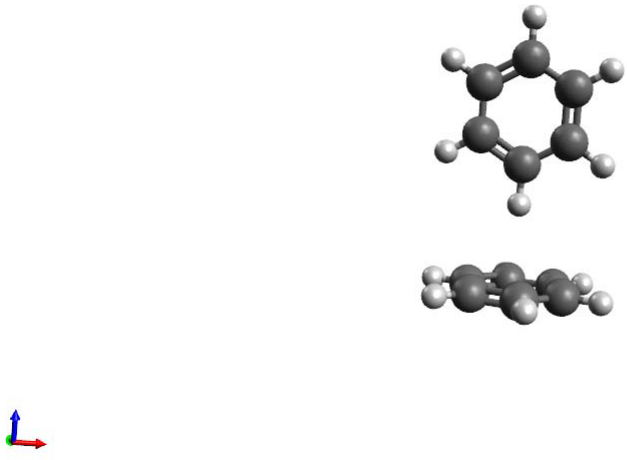


(b) Final parallel-displaced configuration

Figure 4.7: Parallel-displaced configuration optimization



(a) Initial t-shaped configuration



(b) Final t-shaped configuration

Figure 4.8: T-shaped configuration optimization

5.0 CONCLUSION.

In this work we have considered two unconventional methods of electronic structure calculations: OF-DFT and Harris functional approximation. In the OF-DFT approach one models the molecular electron density by means of the Thomas-Fermi model and its extensions. Although OF-DFT is a very appealing method from the point of view of computational complexity because of its linear scaling with the system size. It is still not mature enough to be applied to molecules. This can be explained by the lack of a meaningful kinetic energy functional. Indeed, in this work we used a very crude approximation for the kinetic energy by means of the Thomas-Fermi-Weizsacker functional. This expression is only valid for the uniform electron gas and systems with no more than two fermions. The more advanced forms of the kinetic energy which involve non-local contributions may result in prohibitively long calculations. The famous expansions [62] around the average electron density are not applicable to molecular systems. Although, the binding curves that we have obtained for the benzene and thiophene dimers 2.2,2.3 look close to our benchmark results, the OF-DFT methodology is not suitable for molecular packing optimization due to the lack of the correct kinetic energy functional.

On the other hand, Harris functional approximation gave very promising results. First of all, we have obtained the binding curves for various benzene and thiophene monomer orientations. These curves were in a very good agreement with our benchmark results. Moreover these results suggest that the Harris functional approximation augmented by the empirical dispersion correction can be used for the geometry optimization.

In the final chapter we applied the SPSA algorithm to simulate the molecular packing of two benzene monomers and found that armed with the Harris+GGA+D method and the SPSA procedure one can find all possible minima of the proposed system.

APPENDIX A

DENSITY MATRIX FORMALISM.

A.1 DENSITY OPERATORS AND REDUCED DENSITY OPERATORS.

Consider a general description of a quantum state, where $\mathbf{x} = (\mathbf{r}, s)$. Here $\mathbf{r}$ represents the spacial coordinate and s is the spin coordinate. The quantity,

$$\Psi_N(\mathbf{x}_1\mathbf{x}_2\cdots\mathbf{x}_N)\Psi_N^*(\mathbf{x}_1\mathbf{x}_2\cdots\mathbf{x}_N) \tag{A.1.1}$$

is the probability distribution associated with a solution of the Schrodinger equation, with the Hamiltonian operator $\hat{H}_N$. The main result we are looking to establish here is the usefulness of quantities like

$$\gamma_N(\mathbf{x}_1'\mathbf{x}_2'\cdots\mathbf{x}_N', \mathbf{x}_1\mathbf{x}_2\cdots\mathbf{x}_N) = \Psi_N(\mathbf{x}_1'\mathbf{x}_2'\cdots\mathbf{x}_N')\Psi_N^*(\mathbf{x}_1\mathbf{x}_2\cdots\mathbf{x}_N) \tag{A.1.2}$$

which is more general than (A.1.1). The two sets of independent quantities $\mathbf{x}_1'\mathbf{x}_2'\cdots\mathbf{x}_N'$ and $\mathbf{x}_1\mathbf{x}_2\cdots\mathbf{x}_N$ can be thought of as two sets of indices that give (A.1.2) a numerical value. We can think of (A.1.2) as an element of a matrix, which we shall call a density matrix. It can also be viewed as the coordinate representation of the density operator

$$|\Psi_N><\Psi_N| = \hat{\gamma}_N \tag{A.1.3}$$

The basic Hamiltonian operator is usually the sum of two symmetric one-electron operators and a symmetric two-electron operator. It also does not depend on spin. Similarly, operators corresponding to other physical observables are of one-electron or two-electron type and

usually are spin free. Wave functions Ψ_N are antisymmetric. All of the above means that the expectation formulas like

$$< \hat{A} >= tr(\hat{\gamma}_N \hat{A}) \tag{A.1.4}$$

can be simplified by integrating $\Psi_N \Psi_N^*$ over $N-2$ of its variables. This gives rise to the concept of reduced density matrix and spinless density matrix.

One calls (A.1.1) the Nth order density matrix for a pure state of an N-electron system. One can then define the reduced density matrix of order p by the formula

$$\gamma_p(\mathbf{x}_1^{'}\mathbf{x}_2^{'}\cdots\mathbf{x}_p^{'}, \mathbf{x}_1\mathbf{x}_2\cdots\mathbf{x}_p) =$$

$$\binom{N}{p} \int\cdots\int \gamma_N(\mathbf{x}_1^{'}\mathbf{x}_2^{'}\cdots\mathbf{x}_p^{'}\mathbf{x}_{p+1}\cdots\mathbf{x}_N, \mathbf{x}_1\mathbf{x}_2\cdots\mathbf{x}_p\cdots\mathbf{x}_N) d\mathbf{x}_{p+1}\cdots d\mathbf{x}_N$$

where $\binom{N}{p}$ is a binomial coefficient. For example, the density matrix of order 2 looks like

$$\gamma_2(\mathbf{x}_1^{'}\mathbf{x}_2^{'}, \mathbf{x}_1\mathbf{x}_2) = \frac{N(N-1)}{2} \int\cdots\int \Psi_N(\mathbf{x}_1^{'}\mathbf{x}_2^{'}\mathbf{x}_3\cdots\mathbf{x}_N)\Psi_N^*(\mathbf{x}_1\mathbf{x}_2\mathbf{x}_3\cdots\mathbf{x}_N) d\mathbf{x}_3\cdots d\mathbf{x}_N \tag{A.1.5}$$

and the density matrix of order 1 is

$$\gamma_1(\mathbf{x}_1^{'}, \mathbf{x}_1) = N \int\cdots\int \Psi_N(\mathbf{x}_1^{'}\mathbf{x}_2\cdots\mathbf{x}_N)\Psi_N^*(\mathbf{x}_1\mathbf{x}_2\cdots\mathbf{x}_N) d\mathbf{x}_2\cdots d\mathbf{x}_N \tag{A.1.6}$$

One can also obtain γ_1 from γ_2 by performing a simple integration

$$\gamma_1(\mathbf{x}_1^{'}, \mathbf{x}_1) = \frac{2}{N-1} \int \gamma_2(\mathbf{x}_1^{'}\mathbf{x}_2, \mathbf{x}_1\mathbf{x}_2) d\mathbf{x}_2 \tag{A.1.7}$$

Like $\hat{\gamma}_N$, all reduced density matrix operators are positive semi-definite and Hermitian. The Hermitian reduced density operators $\hat{\gamma}_1$ and $\hat{\gamma}_2$ admit eigenfunctions and associated eigenvalues,

$$\int \gamma_1(\mathbf{x}_1^{'}, \mathbf{x}_1)\psi_i(\mathbf{x}_1) d\mathbf{x}_1 = n_i \psi_i(\mathbf{x}_1^{'}) \tag{A.1.8}$$

and

$$\int \gamma_2(\mathbf{x}_1^{'}\mathbf{x}_2^{'}, \mathbf{x}_1\mathbf{x}_2)\theta_i(\mathbf{x}_1\mathbf{x}_2) d\mathbf{x}_1 d\mathbf{x}_2 = g_i \theta_i(\mathbf{x}_1^{'}\mathbf{x}_2^{'}) \tag{A.1.9}$$

For $\hat{\gamma}_1$, the eigenfunctions $\psi_i(\mathbf{x})$ are called natural spin orbitals, and the eigenvalues n_i the occupation numbers. We can express an operator in terms of its eigenvectors and obtain

$$\hat{\gamma}_1 = \sum_i n_i |\psi_i >< \psi_i| \tag{A.1.10}$$

or

$$\gamma_1(\mathbf{x}_1^{'}, \mathbf{x}_1) = \sum_i n_i \psi_i(\mathbf{x}_1^{'}) \psi_i^*(\mathbf{x}_1) \tag{A.1.11}$$

Similarly,

$$\hat{\gamma}_2 = \sum_i g_i |\theta_i >< \theta_i| \tag{A.1.12}$$

where g_i is called an occupation number again and $|\theta_i >$ is called a natural geminal.

The interpretation of $n_i$and g_i is very straightforward. n_i is proportional to the probability of the one-electron state $|\psi_i >$ being occupied and g_i is proportional to the probability of the two-electron state $|\theta_i >$ being occupied.

Consider the expectation value for an antisymmetric N-body wave function Ψ of a one-electron operator

$$\hat{P}_1 = \sum_{i=1}^{N} P_1(\mathbf{x}_i, \mathbf{x}_i^{'}) \tag{A.1.13}$$

We have

$$< \hat{P}_1 >= tr(\hat{P}_1 \gamma_N) = \int P_1(\mathbf{x}_1^{'} \mathbf{x}_1) \gamma_1(\mathbf{x}_1^{'}, \mathbf{x}_1) d\mathbf{x}_1 d\mathbf{x}_1^{'} \tag{A.1.14}$$

For local operators we will keep a simpler notation

$$\hat{P}_1 = \sum_{i=1}^{N} P_1(\mathbf{x}_i) \tag{A.1.15}$$

and the corresponding expectation value is

$$< \hat{P}_1 >= \int [P_1(\mathbf{x}_1) \gamma_1(\mathbf{x}_1^{'}, \mathbf{x}_1)]_{\mathbf{x}_1^{'} = \mathbf{x}_1} d\mathbf{x}_1 \tag{A.1.16}$$

All local two-electron operators we can denote by their diagonal part and write

$$\hat{P}_2 = \sum_{i<j} P_2(x_i, x_j) \tag{A.1.17}$$

and obtain for the corresponding expectation value

$$< \hat{P}_2 >= tr(\hat{P}_2\gamma_N) = \int\int [P_2(\mathbf{x}_1,\mathbf{x}_2)\gamma_2(\mathbf{x}_1^{'}\mathbf{x}_2^{'},\mathbf{x}_1\mathbf{x}_2)]_{\mathbf{x}_1^{'}=\mathbf{x}_1,\mathbf{x}_2^{'}=\mathbf{x}_2} d\mathbf{x}_1 d\mathbf{x}_2 \tag{A.1.18}$$

For the expectation value of the Hamiltonian we obtain

$$E = tr(\hat{H}\gamma_N) = \int [(-\frac{1}{2}\nabla_1^2 + v(r_1))\gamma_1(\mathbf{x}_1^{'},\mathbf{x}_1)]]_{\mathbf{x}_1^{'}=\mathbf{x}_1} d\mathbf{x}_1 + \int\int \frac{1}{r_{12}}\gamma_2(\mathbf{x}_1\mathbf{x}_2,\mathbf{x}_1\mathbf{x}_2) d\mathbf{x}_1 d\mathbf{x}_2 \tag{A.1.19}$$

A.2 SPINLESS DENSITY MATRICES.

Many operators of interest do not have spin dependence. This makes it possible to reduce the reduced density matrices even further by performing summation over spin degrees of freedom.

We define the first-order and second order spinless density matrices by

$$\rho_1(\mathbf{r}_1^{'},\mathbf{r}_1) = \int \gamma(\mathbf{r}_1^{'}s_1,\mathbf{r}_1 s_1) ds_1 = N \int\cdots\int \Psi(\mathbf{r}_1^{'}s_1\mathbf{x}_2\cdots\mathbf{x}_N)\Psi(\mathbf{r}_1 s_1\mathbf{x}_2\cdots\mathbf{x}_N) ds_1 d\mathbf{x}_2\cdots d\mathbf{x}_N \tag{A.2.1}$$

and

$$\rho_2(\mathbf{r}_1^{'}\mathbf{r}_2^{'},\mathbf{r}_1\mathbf{r}_2) = \int\int \gamma_2(\mathbf{r}_1^{'}s_1\mathbf{r}_2^{'}s_2,\mathbf{r}_1 s_1\mathbf{r}_2 s_2) ds_1 ds_2 = \tag{A.2.2}$$

$$\frac{N(N-1)}{2}\int\cdots\int \Psi(\mathbf{r}_1^{'}s_1\mathbf{r}_2^{'}s_2\cdots\mathbf{x}_N)\Psi(\mathbf{r}_1 s_1\mathbf{r}_2 s_2\cdots\mathbf{x}_N) ds_1 ds_2 d\mathbf{x}_3\cdots d\mathbf{x}_N$$

We also introduce a shorthand notation for the diagonal elements of ρ_2

$$\rho_2(\mathbf{r}_1,\mathbf{r}_2) = \rho_2(\mathbf{r}_1\mathbf{r}_2,\mathbf{r}_1\mathbf{r}_2) = \frac{N(N-1)}{2}\int\cdots\int |\Psi|^2 ds_1 ds_2 d\mathbf{x}_3\cdots d\mathbf{x}_N \tag{A.2.3}$$

and note that the diagonal elements of $\rho_1(\mathbf{r}_1^{'},\mathbf{r}_1)$ is just the electron density

$$\rho(\mathbf{r}_1) = \rho_1(\mathbf{r}_1,\mathbf{r}_1) = N\int\cdots\int |\Psi|^2 ds_1 d\mathbf{x}_2\cdots d\mathbf{x}_N \tag{A.2.4}$$

We can also recall the relationship between first and second order density matrices and write

$$\rho(\mathbf{r}_1) = \frac{2}{N-1}\int \rho_2(\mathbf{r}_1, \mathbf{r}_2) d\mathbf{r}_2 \tag{A.2.5}$$

The expectation value for spin free operators $Q_1(r_1)$ and $Q_2(r_1 r_2)$

$$< \hat{Q}_1 >= \int [Q_1(\mathbf{r}_1)\rho_1(\mathbf{r}_1^{'}, \mathbf{r}_1)]_{\mathbf{r}_1^{'}=\mathbf{r}_1} d\mathbf{r}_1 \tag{A.2.6}$$

and

$$< \hat{Q}_2 >= \int\int [Q_2(\mathbf{r}_1\mathbf{r}_2)\rho_2(\mathbf{r}_1^{'}\mathbf{r}_2^{'}, \mathbf{r}_1\mathbf{r}_2)]_{\mathbf{r}_1^{'}=\mathbf{r}_1, \mathbf{r}_2^{'}=\mathbf{r}_2} d\mathbf{r}_1 d\mathbf{r}_2 \tag{A.2.7}$$

The energy is now given by

$$E = \int [-\frac{1}{2}\nabla^2_{\mathbf{r}}\rho_1(\mathbf{r}^{'}, \mathbf{r})]_{\mathbf{r}^{'}=\mathbf{r}} d\mathbf{r} + \int v(\mathbf{r})\rho(\mathbf{r}) d\mathbf{r} + \int\int \frac{1}{r_{12}}\rho_2(\mathbf{r}_1, \mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2 \tag{A.2.8}$$

Additional insight can be gained by considering the last term in (A.2.8) which represents the electron-electron repulsion energy. If this were purely classical, it would just be the self-repulsion energy of $\rho(\mathbf{r})$, which is

$$J[\rho] = \frac{1}{2}\int\int \frac{1}{r_{12}}\rho(\mathbf{r}_1)\rho(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2 \tag{A.2.9}$$

We can obtain a similar formula in (A.2.8) if we write

$$\rho_2(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{2}\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)[1 + h(\mathbf{r}_1, \mathbf{r}_2)] \tag{A.2.10}$$

where $h(\mathbf{r}_1, \mathbf{r}_2)$ is the pair correlation function which incorporates non-classical effects. This function satisfies an important integral condition

$$\int \rho(\mathbf{r}_2) h(\mathbf{r}_1, \mathbf{r}_2) d\mathbf{r}_2 = -1 \tag{A.2.11}$$

The quantity inside the integral is called the exchange-correlation hole of an electron at $\mathbf{r}_1$ and is given by $\rho_{xc}(\mathbf{r}_1, \mathbf{r}_2) = \rho(\mathbf{r}_2)h(\mathbf{r}_1, \mathbf{r}_2)$. Now the expression for the electron repulsion energy can be written as

$$V_{ee} = J[\rho] + \frac{1}{2}\int\int \frac{1}{r_{12}}\rho(\mathbf{r}_1)\rho_{xc}(\mathbf{r}_1, \mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2 \tag{A.2.12}$$

Sometimes it is convenient to have the spinless density matrices resolved into components arising from different spins. For $\rho_1(\mathbf{r}_1^{'}, \mathbf{r}_1)$ and for any values of $\mathbf{r}_1^{'}$ and $\mathbf{r}_1$, this is the sum over the spin of the diagonal parts of γ_1. So we have then

$$\rho_1(\mathbf{r}_1^{'}, \mathbf{r}_1) = \rho_1^{\alpha\alpha}(\mathbf{r}_1^{'}, \mathbf{r}_1) + \rho_1^{\beta\beta}(\mathbf{r}_1^{'}, \mathbf{r}_1) \tag{A.2.13}$$

Similarly, for the second order density matrix we have

$$\rho_2(\mathbf{r}_1^{'}\mathbf{r}_2^{'}, \mathbf{r}_1\mathbf{r}_2) = \rho_2^{\alpha\alpha,\alpha\alpha}(\mathbf{r}_1^{'}\mathbf{r}_2^{'}, \mathbf{r}_1\mathbf{r}_2) + \rho_2^{\beta\beta,\beta\beta}(\mathbf{r}_1^{'}\mathbf{r}_2^{'}, \mathbf{r}_1\mathbf{r}_2) + \rho_2^{\alpha\beta,\alpha\beta}(\mathbf{r}_1^{'}\mathbf{r}_2^{'}, \mathbf{r}_1\mathbf{r}_2) + \rho_2^{\beta\alpha,\beta\alpha}(\mathbf{r}_1^{'}\mathbf{r}_2^{'}, \mathbf{r}_1\mathbf{r}_2) \tag{A.2.14}$$

For example, the famous Hartree-Fock energy can be written in terms of these density matrices in a very compact way. We provide here just the final result. The detailed derivation can be found in any book on density functional theory [87].

$$E_{HF}[\rho_1] = \int [(-\frac{1}{2}\nabla_{\mathbf{r}}^2 \rho_1(\mathbf{r}^{'}, \mathbf{r}) + v(\mathbf{r}_1))]_{\mathbf{r}_1^{'}=\mathbf{r}_1} dr_1 + \tag{A.2.15}$$

$$\frac{1}{2}\int\int \frac{1}{r_{12}}\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)d\mathbf{r}_1 d\mathbf{r}_2 - \frac{1}{2}\int\int [\rho_1^{\alpha\alpha}(\mathbf{r}_1, \mathbf{r}_2)\rho_1^{\alpha\alpha}(\mathbf{r}_2, \mathbf{r}_1) + \rho_1^{\beta\beta}(\mathbf{r}_1, \mathbf{r}_2)\rho_1^{\beta\beta}(\mathbf{r}_2, \mathbf{r}_1)]d\mathbf{r}_1 d\mathbf{r}_2$$

APPENDIX B

ORBITAL FREE DENSITY FUNCTIONAL THEORY BACKGROUND.

B.1 THE THOMAS-FERMI MODEL AS THE ORIGINAL IDEA.

The history begins with the works of Thomas [64] and Fermi [65] in the 1920s. These authors realized that it was possible to approximate the electron density within an atom by means of statistical mechanics considerations. The model they proposed is quite crude because they suggested to compare the atomic electron density locally with a uniform electron gas. The formula that they derived seeks the value of the kinetic energy of electrons of the uniform electron gas. In fact, the modern OF-DFT is still in search of a good kinetic energy functional. Here we present a short derivation of that expression for the kinetic energy.

We divide the space into many small cubes, each of side l and volume $\Delta V = l^3$, each containing some fixed number of electrons ΔN, and assume that the electrons in each cell behave like independent fermions at the temperature of $0K$, with cells being independent.

The energy levels of a particle in a 3D infinite well are given by the formula

$$\varepsilon(n_x, n_y, n_z) = \frac{h^2}{8ml^2}(n_x^2 + n_y^2 + n_z^2) = \frac{h^2}{8ml^2}R^2 \tag{B.1.1}$$

where $n_x, n_y, n_z = 1, 2, 3, ...$ For high quantum numbers, that is, for large R, the number of distinct energy levels with energy smaller that ε can be approximated by the volume of one octant of a sphere with radius R in the space (n_x, n_y, n_z). The number is

$$\Phi(\varepsilon) = \frac{1}{8}\left(\frac{4\pi R^3}{3}\right) = \frac{\pi}{6}\left(\frac{8ml^2\varepsilon}{h^2}\right)^{3/2} \tag{B.1.2}$$

The number of energy levels between ε and $\varepsilon + \delta\varepsilon$ is

$$h(\varepsilon)\Delta\varepsilon = \Phi(\varepsilon + \delta\varepsilon) - \Phi(\varepsilon) = \frac{\pi}{4}\left(\frac{8ml^2}{h^2}\right)^{3/2}\varepsilon^{1/2}\delta\varepsilon \tag{B.1.3}$$

where $h(\varepsilon)$ is the density of states.

To compute the total energy for the cell with ΔN electrons, we need the probability for the state with energy ε, to be occupied, which in our case is the Fermi-Dirac distribution,

$$f(\varepsilon) = \frac{1}{1 + e^{\beta(\varepsilon-\mu)}} \tag{B.1.4}$$

which at $0K$ reduces to a step function:

$$f(\varepsilon) = \begin{cases} 1, & \varepsilon < \varepsilon_F \\ 0, & \varepsilon > \varepsilon_F \end{cases} \tag{B.1.5}$$

where ε_F is the Fermi energy.

Now we find the total energy of the electrons in this cell by summing the contributions from the different energy states:

$$\Delta E = 2\int \varepsilon f(\varepsilon)h(\varepsilon)d\varepsilon = \frac{8\pi}{5}\left(\frac{2m}{h^2}\right)^{3/2} l^3\varepsilon_F^{5/2} \tag{B.1.6}$$

The Fermi energy is related to the number of electrons ΔN in the cell, thought the formula

$$\Delta N = 2\int f(\varepsilon)g(\varepsilon)d\varepsilon = \frac{8\pi}{3}\left(\frac{2m}{h^2}\right)^{3/2} l^3\varepsilon_F^{3/2} \tag{B.1.7}$$

Eliminating ε_F from these equations we obtain

$$\Delta E = \frac{3}{5}\Delta N\varepsilon_F = \frac{3h^2}{10}\left(\frac{3}{8\pi}\right)^{2/3} l^3\left(\frac{\Delta N}{l^3}\right)^{5/3} \tag{B.1.8}$$

Adding the contributions from all cells, we find the total kinetic energy in atomic units to be,

$$T_{TF}[\rho] = C_F\int \rho^{5/3}(\mathbf{r})d\mathbf{r} \tag{B.1.9}$$

where $C_F = 0.3(3\pi^2)^{2/3} = 2.871$. This is the famous Thomas-Fermi kinetic energy functional. In fact, this is also the first LDA (local density approximation) functional.

Neglecting non-classical terms, we obtain the energy functional for the Thomas-Fermi theory of atoms

$$E_{TF}[\rho] = C_F \int \rho^{5/3}(\mathbf{r})d\mathbf{r} - Z \int \frac{\rho(\mathbf{r})}{\mathbf{r}} d\mathbf{r} + \frac{1}{2} \int \int \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2 \tag{B.1.10}$$

The ground state energy can be obtained by minimizing this functional with respect to the electron density subject to the constraint that the electron density integrates up to the number of electrons in the system.

Needless to say, this functional gives a poor approximation for any molecular system because it compares the electron density to the uniform electron gas. Furthermore, this approach completely disregards any exchange-correlation effects. It was also shown that the model in the present form predicts no molecular binding.

B.2 FROM THOMAS-FERMI MODEL TO THOMAS-FERMI-DIRAC MODEL.

For simplicity, consider a non-degenerate closed-shell ground state described by a single-determinant wave function, a first-order density matrix, and a spinless first-order density matrix of the form

$$\rho_1(\mathbf{r}_1, \mathbf{r}_1^{'}) = 2 \sum_{i}^{N/2} \phi_i(\mathbf{r}_1)\phi_i^*(\mathbf{r}_1^{'}) \tag{B.2.1}$$

where the ϕ_i are the doubly occupied spatial orbitals. The energy is then given by the Hartree-Fock formula (A.2.15)

$$E_{HF}[\rho_1] = \int \left[-\frac{1}{2} \nabla_1^2 \rho_1(\mathbf{r}_1, \mathbf{r}_2) \right]_{\mathbf{r}_2 = \mathbf{r}_1} d\mathbf{r}_1 + \int \rho(\mathbf{r}) v(\mathbf{r}) d\mathbf{r} + \tag{B.2.2}$$

$$+ J[\rho] - \frac{1}{4} \int \int \frac{1}{r_{12}} \rho_1(\mathbf{r}_1, \mathbf{r}_2)\rho_1(\mathbf{r}_2, \mathbf{r}_1) d\mathbf{r}_1 d\mathbf{r}_2$$

Here, the kinetic energy is

$$T[\rho] = \int \left[-\frac{1}{2} \nabla_1^2 \rho_1(\mathbf{r}_1, \mathbf{r}_2) \right]_{\mathbf{r}_2 = \mathbf{r}_1} d\mathbf{r}_1 \tag{B.2.3}$$

and

$$V_{ee}[\rho] = J[\rho] - K[\rho] \tag{B.2.4}$$

where $K[\rho]$ is the Hartree-Fock exchange energy functional

$$K[\rho] = \frac{1}{4}\int \frac{1}{r_{12}}|\rho_1(\mathbf{r}_1, \mathbf{r}_2)|^2 d\mathbf{r}_1 d\mathbf{r}_2 \tag{B.2.5}$$

and

$$J[\rho] = \frac{1}{2}\int\int \frac{1}{r_{12}}\rho(\mathbf{r}_1)\rho(\mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2 \tag{B.2.6}$$

We are facing the problem of expressing $T[\rho]$ and $K[\rho]$ in terms of the diagonal elements of ρ_1,the electron density. To solve this problem we take a second look at the uniform electron gas.

In the previous section we considered the uniform-gas description for particle-in-a-box states for which the boundary conditions had the form $\psi(x = 0) = \psi(x = l) = 0$. Equivalently, for a large number of particles, one can employ periodic boundary conditions of the type $\psi(x + l) = \psi(x)$. These lead to the orbitals

$$\psi(k_x, k_y, k_z) = \frac{1}{l^{3/2}} e^{i(k_x x + k_y y + k_z z)} = \frac{1}{V^{1/2}} e^{i\mathbf{kr}} \tag{B.2.7}$$

where $k_x = (2\pi/l)n_x$, $k_y = (2\pi/l)n_y$, $k_z = (2\pi/l)n_z$ with $n_x, n_y, n_z = 0, \pm 1, \pm 2 \cdots$. The energy levels are

$$E(n_x, n_y, n_z) = \frac{h^2}{8ml^2}((2n_x)^2 + (2n_y)^2 + (2n_z)^2) \tag{B.2.8}$$

The first order density matrix then becomes

$$\rho_1(\mathbf{r}_1, \mathbf{r}_2) = \frac{2}{V}\sum_{k_{occupied}} e^{i\mathbf{k}(\mathbf{r}_1 - \mathbf{r}_2)} \tag{B.2.9}$$

If there are many occupied states, the sum can be replaced by an integral, giving

$$\rho_1(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{4\pi^3}\int e^{i\mathbf{k}(\mathbf{r}_1 - \mathbf{r}_2)} d\mathbf{k} = \frac{1}{4\pi^3}\int_0^{k_F} k^2 dk \int\int e^{i\mathbf{kr}_{12}} \sin(\theta) d\theta d\phi \tag{B.2.10}$$

Since $\rho_1(r, r) = \rho(r)$, we obtain

$$\rho(\mathbf{r}) = \frac{k_F^3}{3\pi^2} \tag{B.2.11}$$

and then

$$k_F(\mathbf{r}) = [3\pi^2\rho(\mathbf{r})]^{1/3} \tag{B.2.12}$$

For an inhomogeneous system, the natural choice of the argument of $k_F(\mathbf{r})$ to be used in (B.2.10) is the average of $\mathbf{r}_1$ and $\mathbf{r}_2$. Let denote

$$\mathbf{r} = \frac{1}{2}(\mathbf{r}_1 + \mathbf{r}_2) \tag{B.2.13}$$

and

$$\mathbf{s} = \mathbf{r_1} - \mathbf{r}_2 \tag{B.2.14}$$

and proceed to out the integration in (B.2.10).

Choose $\mathbf{s}$ to lie along the k_z axis. Then (B.2.10) can be evaluated as follows:

$$\rho_1(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{4\pi^3}\int_0^{k_F} k^2 dk \int_0^{\pi} \sin(\theta) e^{ikr_{12}\cos(\theta)} d\theta \int_0^{2\pi} d\phi = \tag{B.2.15}$$

$$= 3\rho(\mathbf{r})\left[\frac{\sin(t) - t\cos(t)}{t^3}\right] = \rho_1(\mathbf{r}, s)$$

where $t = k_F(\mathbf{r})s$.

To evaluate the kinetic energy, we need

$$\nabla^2_{\mathbf{r}_1} = \frac{1}{4}\nabla^2_{\mathbf{r}} + \nabla^2_s + \nabla_{\mathbf{r}}\nabla_s \tag{B.2.16}$$

and

$$\nabla^2_{\mathbf{r}_2} = \frac{1}{4}\nabla^2_{\mathbf{r}} + \nabla^2_s - \nabla_{\mathbf{r}}\nabla_s \tag{B.2.17}$$

Thus

$$[\nabla^2_1 \rho_1(\mathbf{r}_1, \mathbf{r}_2)]_{r_2 = r_1} = [(\frac{1}{4}\nabla^2_{\mathbf{r}} + \nabla^2_s + \nabla_{\mathbf{r}}\nabla_s)\rho_1(\mathbf{r}, s)]_{s=0} = \frac{1}{4}\nabla^2_{\mathbf{r}}\rho(\mathbf{r}) - \frac{3}{5}(3\pi^2)^{2/3}\rho(\mathbf{r})^{5/3} \tag{B.2.18}$$

For any well-behaved $\rho(\mathbf{r})$,

$$\int \nabla^2_{\mathbf{r}}\rho(\mathbf{r}) d\mathbf{r} = 0 \tag{B.2.19}$$

so that the kinetic energy becomes

$$T_{TF}[\rho] = C_F \int \rho(\mathbf{r})^{5/3} d\mathbf{r} \tag{B.2.20}$$

The exchange energy can be obtained similarly

$$K_D[\rho] = \frac{1}{4}\int\int \frac{[\rho_1(\mathbf{r}, s)]^2}{s} d\mathbf{r} ds = \tag{B.2.21}$$

$$= 9\pi \int \rho^2(\mathbf{r}) \frac{1}{k_F^2} d\mathbf{r} \left[\int_0^\infty \frac{(\sin(t) - t\cos(t))^2}{t^5} dt \right] = C_x \int \rho^{4/3}(\mathbf{r}) d\mathbf{r}$$

where $C_x = \frac{3}{4}\left(\frac{3}{\pi}\right)^{1/3} = 0.7386$. This is the famous exchange-energy formula of Dirac [17].

The total energy now becomes

$$E_{TFD}[\rho] = C_F \int \rho(\mathbf{r})^{5/3} d\mathbf{r} + \int \rho(\mathbf{r}) v(\mathbf{r}) d\mathbf{r} + J[\rho] - C_x \int \rho(\mathbf{r})^{4/3} d\mathbf{r} \tag{B.2.22}$$

B.3 GRADIENT CORRECTION.

Consider the ground state of N noninteracting electrons. The spinless first-order reduced density matrix can be written as

$$\rho_1(\mathbf{r}, \mathbf{r}') = 2 \sum_i^\infty \phi_i(\mathbf{r}) \phi_i^*(\mathbf{r}') \eta(\varepsilon_F - \varepsilon_i) = 2 < \mathbf{r} | \eta(\varepsilon_F - \hat{H}) | \mathbf{r}' > \tag{B.3.1}$$

where $\eta(x)$ is the Heaviside step function and $\hat{H}$ is the one-particle Hamiltonian having ϕ_i and ε_i as its eigenstates and eigenvalues:

$$\hat{H}\phi_i(\mathbf{r}) = \{-\frac{1}{2}\nabla^2 + w(\mathbf{r})\}\phi_i(\mathbf{r}) = \varepsilon_i \phi_i(\mathbf{r}) \tag{B.3.2}$$

with $w(\mathbf{r})$ the local potential function and ε_F can take values between the highest occupied and lowest unoccupied eigenvalues.

Our goal is to express the kinetic energy T as a function of the electron density. The idea is the following: T is determined by $\rho_1(\mathbf{r}, \mathbf{r}')$, which in turn is determined by $w(\mathbf{r})$ though (B.3.1); electron density $\rho(\mathbf{r})$ as the diagonal of $\rho_1(\mathbf{r}, \mathbf{r}')$ is also determined by $w(\mathbf{r})$;therefore we can hope to use $w(\mathbf{r})$ as a bridge to connect T to $\rho(\mathbf{r})$.

The key problem is to find $\rho_1(\mathbf{r}, \mathbf{r}')$ in terms of $w(\mathbf{r})$. Note that as expressed in (B.3.1), the N-electron quantity $\rho_1(\mathbf{r}, \mathbf{r}')$ is the matrix representation of a one-particle operator $\hat{\rho}_1 = \eta(\varepsilon_F - \hat{H})$. A one-electron problem is much easier to handle than an N-electron one. Many

techniques have been developed to manipulate and approximate the single-electron Green's function, defined as

$$G(\mathbf{r},\mathbf{r}';\beta) =< \mathbf{r}|e^{-\beta\hat{H}}|\mathbf{r}' >= \sum_{i}^{\infty} \phi_i(\mathbf{r})\phi_i^*(\mathbf{r}')\varepsilon^{-\beta\varepsilon_i} \tag{B.3.3}$$

$G(\mathbf{r},\mathbf{r}')$ is in turn related to $\rho_1(\mathbf{r},\mathbf{r}')$ by an inverse Laplace transform

$$\rho_1(\mathbf{r},\mathbf{r}') = \frac{2}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty} \frac{d\beta}{\beta} e^{\beta\varepsilon_F} G(\mathbf{r},\mathbf{r}';\beta) \tag{B.3.4}$$

where γ is any positive constant.

We now invoke $\tilde{G}(\mathbf{r},\mathbf{p};\beta)$, the Wigner transformation of $G(\mathbf{r},\mathbf{r}';\beta)$. Define

$$\tilde{G}(\mathbf{r},\mathbf{p};\beta) = \frac{1}{(2\pi\hbar)^3}\int d\mathbf{s} < \mathbf{r}-\frac{\mathbf{s}}{2}|e^{-\beta\hat{H}}|\mathbf{r}+\frac{\mathbf{s}}{2} > e^{i\mathbf{ps}/\hbar} = \frac{1}{(2\pi\hbar)^3}\int d\mathbf{s} G(\mathbf{r}-\frac{\mathbf{s}}{2},\mathbf{r}+\frac{\mathbf{s}}{2};\beta)e^{i\mathbf{ps}/\hbar} \tag{B.3.5}$$

Then the inverse is

$$G(\mathbf{r},\mathbf{r}';\beta) = \int d\mathbf{p} e^{-i\mathbf{p}(\mathbf{r}-\mathbf{r}')/\hbar}\tilde{G}(\frac{1}{2}(\mathbf{r}+\mathbf{r}'),\mathbf{p};\beta) \tag{B.3.6}$$

Inserting (B.3.6) in (B.3.4), we obtain

$$\rho_1(\mathbf{r},\mathbf{r}') = \int d\mathbf{p} e^{-i\mathbf{p}(\mathbf{r}-\mathbf{r}')/\hbar}\tilde{\rho}_1(\frac{1}{2}(\mathbf{r}+\mathbf{r}'),\mathbf{p}) \tag{B.3.7}$$

where

$$\tilde{\rho}_1(\mathbf{r},\mathbf{p}) = \frac{2}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty} \frac{d\beta}{\beta} e^{\beta\varepsilon_F}\tilde{G}(\mathbf{r},\mathbf{p};\beta) \tag{B.3.8}$$

Using Wigner expansion of $\tilde{G}$ we can find an explicit approximation of $\rho_1(\mathbf{r},\mathbf{r}')$ in terms of the potential $w(\mathbf{r})$. Thus we obtain

$$\tilde{G}(\mathbf{r},\mathbf{p};\beta) = \frac{1}{(2\pi\hbar)^3}\exp[-\beta H(\mathbf{r},\mathbf{p})] \tag{B.3.9}$$

$$\left\{1+\hbar^2\left[-\frac{\beta^2}{8m}\nabla^2 w(\mathbf{r}) + \frac{\beta^3}{24m}|\nabla w(\mathbf{r})|^2 + \frac{\beta^3}{24m}(\mathbf{p}\cdot\nabla)(\mathbf{p}\cdot\nabla)w(\mathbf{r})\right]\right\}$$

with

$$H(\mathbf{r},\mathbf{p}) = \frac{\mathbf{p}^2}{2m} + w(\mathbf{r}) \tag{B.3.10}$$

Correspondingly, by (B.3.8)

$$\tilde{\rho}_1(\mathbf{r},\mathbf{p}) = \frac{2}{(2\pi\hbar)^3}\eta(\varepsilon_F - H(\mathbf{r},\mathbf{p}))+ \tag{B.3.11}$$

$$+\frac{2\hbar^2}{(2\pi\hbar)^3}\left\{-\frac{1}{8m}\eta^{(2)}(\varepsilon_F - H(\mathbf{r},\mathbf{p}))\nabla^2 w(\mathbf{r})\right\}+$$

$$+\frac{2\hbar^2}{(2\pi\hbar)^3}\left\{\eta^{(3)}(e_F - H(\mathbf{r},\mathbf{p})) + \left[\frac{1}{24m}|w(\mathbf{r})|^2 + \frac{1}{24m^2}(\mathbf{p}\cdot\nabla)(\mathbf{p}\cdot\nabla)w(\mathbf{r})\right]\right\}$$

where we used

$$\eta(\varepsilon_F - \varepsilon_i) = \frac{1}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty}\frac{d\beta}{\beta}e^{\beta\varepsilon_F}e^{-\beta\varepsilon_i} \tag{B.3.12}$$

and

$$\eta^{(n)}(\varepsilon_F - H(\mathbf{r},\mathbf{p})) = \frac{\partial^n}{\partial\varepsilon_F^n}\eta(\varepsilon_F - H(\mathbf{r},\mathbf{p})) \tag{B.3.13}$$

By inserting (B.3.11) into (B.3.7), we find for the total electron density,

$$\rho(\mathbf{r}) = \rho_1(\mathbf{r},\mathbf{r}) = \int\tilde{\rho}_1(\mathbf{r},\mathbf{p})d\mathbf{p} = \frac{1}{3\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2}[\varepsilon_F - w(\mathbf{r})]^{3/2}\eta(\varepsilon_F - w(\mathbf{r})) \tag{B.3.14}$$

$$\left\{1 - \frac{1}{8}\frac{\hbar^2}{2m}[\nabla^2 w(\mathbf{r})(\varepsilon_F - w(\mathbf{r}))^{-2} + \frac{1}{4}|\nabla w(\mathbf{r})|^2(\varepsilon_F - w(\mathbf{r}))^{-3}]\right\}$$

where we have used the integral formula

$$\int\eta\left(\varepsilon_F - \frac{p^2}{2m} - w(\mathbf{r})\right)d\mathbf{p} = \int\eta\left(\varepsilon_F - w(\mathbf{r}) - \frac{\mathbf{p}^2}{2m}\right)4\pi p^2 dp = \frac{4\pi}{3}p_F^3(\mathbf{r}) = \tag{B.3.15}$$

$$= \frac{4\pi}{3}[2m(\varepsilon_F - w(\mathbf{r}))]^{3/2}\eta(\varepsilon_F - w(\mathbf{r}))$$

and its derivative with respect to ε_F. The total kinetic energy is

$$T = \int\frac{p^2}{2m}\tilde{\rho}_1(\mathbf{r},\mathbf{p})d\mathbf{r}d\mathbf{p} = \int d\mathbf{r}\frac{\hbar^2}{10\pi^2 m}\left(\frac{2m}{\hbar^2}\right)^{5/2}[\varepsilon_F - w(\mathbf{r})]^{5/2}\eta(\varepsilon_F - w(\mathbf{r})) \tag{B.3.16}$$

$$\left\{1 - \frac{5}{8}\frac{\hbar^2}{2m}[\frac{5}{3}\nabla^2 w(\mathbf{r})(\varepsilon_F - w(\mathbf{r}))^{-2} - \frac{3}{4}|\nabla w(\mathbf{r})|^2(\varepsilon_F - w(\mathbf{r}))^{-3}]\right\}$$

Using (B.3.14) and (B.3.16) we can eliminate $\varepsilon_F - w(\mathbf{r})$ from (B.3.16). The result found is the TF-$\frac{1}{9}$W functional:

$$T^{(2)}[\rho] = T_{TF}[\rho] + \frac{1}{9}T_W[\rho] = \frac{\hbar^2}{m}\left[\frac{3}{10}(3\pi^2)^{2/3}\int\rho(\mathbf{r})^{5/3}d\mathbf{r} + \frac{1}{9}\cdot\frac{1}{8}\int\frac{|\nabla\rho(\mathbf{r})|^2}{\rho(\mathbf{r})}d\mathbf{r}\right] \tag{B.3.17}$$

Thus we obtained the so-called Thomas-Fermi-Dirac-Weizsacker [64, 65, 44, 46] (TFD-λW) model,

$$E_{TFD-\lambda W}[\rho] = C_F \int \rho^{5/3}(\mathbf{r})d\mathbf{r} + \lambda\frac{1}{8}\int \frac{|\nabla\rho(\mathbf{r})|^2}{\rho(\mathbf{r})}d\mathbf{r} + \int \rho(\mathbf{r})v(\mathbf{r})d\mathbf{r} + J[\rho] - C_x \int \rho^{4/3}(\mathbf{r})d\mathbf{r} \tag{B.3.18}$$

B.4 LINEAR RESPONSE CORRECTION MODEL.

The kinetic energy models above do not satisfy the exact linear response (LR) behavior described by the Lindhard function in the homogeneous limit and the simple local gradient correction used cannot reproduce the oscillatory atomic shell structure. A modification of the kinetic energy functional which is a somewhat different combination of the Thomas-Fermi terms T_{TF} and the gradient correction given by von Weizsacker T_W has been suggested as follows [79]:

$$T[\rho] = F(N)T_{TF}[\rho] + T_W[\rho] \tag{B.4.1}$$

where N is the number of electrons and the factor $F(N)$ is

$$F(N) = \left(1 - \frac{2}{N}\right)\left(1 - \frac{A_1}{N^{1/3}} + \frac{A_2}{N^{2/3}}\right) \tag{B.4.2}$$

with optimized parameter values A_1 and A_2. This kinetic energy functional is known to describe the response properties of the electron gas well and has yielded very good polarizabilities for various atomic systems. It also provides an excellent representation of the kinetic energy of atoms.

However, since only a truly nonlocal kinetic energy density functional can satisfy the exact LR condition, we have to modify the kinetic energy by nonlocal terms completely determined by the requirement that the linear response is exactly satisfied.

In 1994, Perrot [52] proposed another functional by adding a term to the Thomas-Fermi [65] and von Weizsacker [44, 46] functional with an integration kernel $K_\alpha(\mathbf{r}-\mathbf{r}')$ to incorporate the correct linear response:

$$T[\rho] + T_{TF}[\rho] + T_W[\rho] + \int\int d\mathbf{r}d\mathbf{r}' P(\mathbf{r}) K_\alpha(\mathbf{r}-\mathbf{r}') P(\mathbf{r}') \tag{B.4.3}$$

where

$$P(r) = \frac{6}{5}\frac{\rho_\alpha + v\Delta\rho}{\rho}\left(\rho^{5/6} - \rho_\alpha^{5/6}\right) \tag{B.4.4}$$

with $v = 5/\sqrt{32}$ and ρ_α the average electron density. This functional has the correct scaling for $\rho_\alpha \to 0$, and incorporates the proper linear response for perturbations both small and large.

The correction term of the Thomas-Fermi and von Weizsacker was proposed by Wang and Teter [51]. The Wang and Teter kinetic energy is determined by

$$T_{WT}[\rho] = T_{TF}[\rho] + T_W[\rho] + F_{WT}[\rho] \tag{B.4.5}$$

where

$$F_{WT}[\rho] = -\frac{32C_{TF}}{35}\int \rho^{5/3} + \frac{4C_{TF}}{5}\int \rho^{5/6} K_{WT} * \rho^{5/6} \tag{B.4.6}$$

The convolution kernel K_{WT} is given in Fourier space in terms of the Lindhard susceptibility function, which is already available for nearly free electron gas system:

$$\hat{K}_{WT}(\eta) = \left(\frac{1}{2} + \frac{1-\eta^2}{4\eta}\ln\left|\frac{1+\eta}{1-\eta}\right|\right) - 3\eta^2 + \frac{3}{5} \tag{B.4.7}$$

Furthermore, there exists an important group of kinetic energy functionals based on linear response theory which contain the Wang-Teter [51] model. These functionals take the form

$$T_s[\rho] = T_{TF}[\rho] + T_W[\rho] + T_K[\rho] \tag{B.4.8}$$

where the kernel term $T_K[\rho]$ is expressed as

$$T_K[\rho] = C_{TF}\int\int \rho^\alpha(\mathbf{r})\omega(\mathbf{r},\mathbf{r}')\rho^\beta(\mathbf{r}')d\mathbf{r}d\mathbf{r}' \tag{B.4.9}$$

Different functionals of this class are determined by constants α and β and the kernel function ω. A standard treatment based on the linear response of the noninteracting electron gas leads

to a simple expression in reciprocal space for ω when it is taken to be density independent. It is shown [62, 58, 59, 60] that the choice

$$\alpha = \frac{5+\sqrt{5}}{6}, \beta = \frac{5-\sqrt{5}}{6} \tag{B.4.10}$$

was optimal for the density independent kernel.

The latter improvements were recently developed by Wang, Govind and Carter (WGC) [58], who generalize (B.4.9) to density dependent kernels.We are not using the latter functionals in our present work. For this reason we are not discussing them here in detail.

B.5 EXCHANGE-CORRELATION FUNCTIONALS.

So far, we have looked at the major parts of the total energy of an N-electron system. In particular, we have given some approximation to the kinetic energy functional $T[\rho]$, the classical electron-electron interaction or Hartree term $J[\rho]$, the interaction of electron with the external field $\int \rho(\mathbf{r})v(\mathbf{r})d\mathbf{r}$ and the approximation to the exchange energy proposed by Dirac $-C_x \int \rho^{4/3}(\mathbf{r})d\mathbf{r}$. The last contribution, i.e. the correlation energy is the big unknown. Wigner [63] was the first to address this issue in the context of the homogeneous electron gas, by proposing the correlation energy per unit volume given by

$$C[\rho] = -0.056 \int \frac{\rho^{4/3}(\mathbf{r})}{0.079 + \rho^{1/3}(\mathbf{r})} d\mathbf{r} \tag{B.5.1}$$

The exchange energy, although well known as a function of the single-particle orbitals (Hartree-Fock), involves the calculation of computationally expensive integrals. In addition, up to date there is no approximation available where the correlation energy is treated at a comparable level of accuracy. Therefore, if exchange is treated exactly as a functional of the orbitals, it will not be able to compensate for any errors introduced when approximating the correlation term. The key issue here is that the really meaningful quantity is the sum of the two terms $K[\rho] + C[\rho]$; the division is the matter of convenience. Therefore, it seems sensible to treat both terms to a similar level of approximation.

The idea now is to look for consistent approximations to exchange and correlation where both terms are treated in a similar manner. One of the natural starting points is the homogeneous electron gas. We have seen that the exchange energy for this system is exactly given by Dirac's expression [17]:

$$K_D[\rho] = -C_x \int \rho(\mathbf{r})^{4/3} d\mathbf{r} \tag{B.5.2}$$

Excellent approximations for correlation are also available. The most accurate results are based on the quantum Monte Carlo (QMC) simulations of Ceperley and Alder [66]. This correlation functional is exact within numerical accuracy, and has been parameterized by Perdew and Zunger for spin-polarized (P) and spin-unpolarized (U) homogeneous electron gas [68, 69]:

$$C^{PZ}[\rho] = \begin{cases} A \ln r_s + B + C r_s \ln r_s + D r_s, & r_s \leq 1 \\ \frac{\gamma}{1+\beta_1\sqrt{r_s}+\beta_1 r_s}, & r_s > 1 \end{cases} \tag{B.5.3}$$

where $r_s = (3/4\pi\rho)^{1/3}$.

For $r_s \leq 1$ the above expression derives from the random phase approximation, has been calculated by Gell-Mann and Brueckner. This is valid in the limit of very dense electron systems, and fixes the values of the leading coefficients: $A^U = 0.0311$, $B^U = -0.048$. The remaining coefficients were fitted to the QMC results of Ceperley and Alder [66]: $C^U = 0.002$,$D^U = -0.0116$.

Another possible parametrization is one proposed by Vosko [67] where the correlation functional is given by

$$\frac{C^{VNW}[r_s]}{A} = \ln\left(\frac{r_s}{F(\sqrt{r_s})}\right) + \frac{2b}{\sqrt{4c-b^2}} \tan^{-1}\left(\frac{\sqrt{4c-b^2}}{2\sqrt{r_s}+b}\right) - \tag{B.5.4}$$

$$-\frac{b x_0}{F(x_0)}\left[\ln\left(\frac{\sqrt{r_s}-x_0}{F(\sqrt{r_s})}\right) + \tan^{-1}\left(\frac{\sqrt{4c-b^2}}{\sqrt{r_s}+b}\right)\right]$$

with $F(x) = x^2 + bx + c$, and where A,b,c, and x_0 are fitting constants that differ for spin-polarized and spin-unpolarized cases.

B.5.1 The local density approximation.

The local density approximation (LDA) has been for a long time the most widely used approximation to the exchange-correlation energy. The main idea is to consider a general inhomogeneous electronic system as locally homogeneous, and then to use the exchange-correlation hole (A.2.11) corresponding to the homogeneous electron gas, which is known to an excellent accuracy. In practice, energy terms local in the density are calculated by integrating over the volume of the system.

The exchange-correlation energy can be written as the average of an energy density $\epsilon_{XC}^{LDA}[\rho]$:

$$E_{XC}^{LDA}[\rho] = \int \rho(\mathbf{r})\epsilon_{XC}^{LDA}[\rho(\mathbf{r})]d\mathbf{r} \tag{B.5.5}$$

weighted with the space-dependent electronic density of the system.

While the exchange-correlation energy $E_{XC}[\rho]$ should be a local functional of ρ, there is no reason why the energy density should also be so. In fact, in general ϵ_{XC} is not a functional of the density. From its definition it is clear that it has to be a non-local object, because it reflects the fact that the probability of finding an electron at $\mathbf{r}$ depends on the presence of other electrons in the surroundings. However, in the LDA it becomes a function of the local density because it corresponds to a homogeneous system where ρ is the same everywhere.

There are a number of features of the LDA that are rather general and well established. These are the following.

- It favors electronic densities that are more homogeneous than the exact ones.
- It tends to overestimate the binding energy of molecules and the cohesive energy of solids. This trend is opposite to Hartree-Fock, which underestimates binding energies.
- Geometries of well-behaved systems, i.e. those involving strong bonds are remarkably good within the LDA. Bond lengths, bond angles, and vibrational frequencies reproduce experimental values within a few percent. Dielectric properties like the dielectric constant and piezo-electric coefficients are overestimated by about 10%.
- For weakly bound systems that involve hydrogen bonds or van der Waals closed shell interactions, bond lengths are too short. Dispersion interactions are poorly reproduced.

B.5.2 Gradient expansions.

To address the issue of inhomogeneities in the electron density, the natural road is to carry out an expansion of the density in terms of the gradient and higher order derivatives. In general, the exchange-correlation energy can be written in the following form:

$$E_{XC}[\rho] = \int \rho(\mathbf{r})\epsilon_{XC}[\rho(\mathbf{r})]F_{XC}[\rho(\mathbf{r}), \nabla\rho(\mathbf{r}), \nabla^2\rho(\mathbf{r}), \cdots]d\mathbf{r} \tag{B.5.6}$$

where the function F_{XC} is an enhancement factor that modifies the LDA expression according to the variation of the density in the vicinity of the considered point. In this sense, the gradient corrections constitute a semi-local approach, which will hardly be able to capture non-local effects at longer ranges.

The second order gradient expansion of the exchange energy introduces a term proportional to the squared gradient of the density. The fourth order gradient expansion of the exchange enhancement factor F_X is

$$F_X(p, q) = 1 + \frac{10}{81}p + \frac{146}{2025}q^2 - \frac{73}{405}qp + Dp^2 + O(\nabla\rho^6) \tag{B.5.7}$$

where

$$p = \frac{|\nabla\rho|^2}{4(3\pi^2)^{2/3}\rho^{8/3}} \tag{B.5.8}$$

and

$$q = \frac{\nabla^2\rho}{4(3\pi^2)^{2/3}\rho^{5/3}} \tag{B.5.9}$$

The first two coefficients of the expansion are exactly known. The third one is the result of a difficult many-body calculation. The fourth coefficient, has not been explicitly calculated to date, but the best numerical estimate is that it is negligible.

The second order gradient expansion corresponds to an expression of the type

$$E_{XC}[\rho] = \int A_{XC}[\rho]\rho(\mathbf{r})^{4/3}d\mathbf{r} + \int C_{XC}[\rho]\frac{|\nabla\rho(\mathbf{r})|^2}{\rho(\mathbf{r})^{4/3}}d\mathbf{r} \tag{B.5.10}$$

which is asymptotically correct for densities which vary slowly in space. The LDA retains only the leading term of (B.5.10). It is well known that a straightforward evaluation of this expansion is ill behaved, in the sense that it is not monotonically convergent, and exhibits singularities that cancel out only when an infinite number of terms is re-summed. In fact,

the first order correction worsens the results, and the second order correction is plagued with divergences.

Gradient expansion should be carried out very carefully in order to retain all the relevant contributions to the desired order. These expansions easily violate one or more of the exact conditions required for the exchange and correlation holes, such as the normalization condition, the negativity of the exchange density, or the self-interaction cancelation. Perdew showed that imposing these conditions on functionals that originally do not satisfy them results in a remarkable improvement of the quality of exchange energies. On the basis of this type of reasoning, a number of modified gradient expansions have been proposed. These have been named generalized gradient approximations (GGAs).

Normally GGAs improve over some of the drawbacks of the LDA, although this is not always the case.

B.5.2.1 Langreth-Mehl functional. The first GGA proposed in the literature assumes the following form [70]:

$$\epsilon_X = \epsilon_X^{LDA} - a\frac{|\nabla\rho(r)|^2}{\rho(r)^{4/3}}\left(\frac{7}{9} + 18f^2\right) \tag{B.5.11}$$

$$\epsilon_C = \epsilon_C^{RPA} + a\frac{|\nabla\rho(r)|^2}{\rho(r)^{4/3}}(2e^{-F} + 18f^2) \tag{B.5.12}$$

where $F = b|\nabla\rho(r)|/\rho(r)^{7/6}$, $b = (9\pi)^{1/6}$, $a = \pi/(16(3\pi^2)^{4/3})$, and $f = 0.15$.

B.5.2.2 BLYP functional. In 1988, Becke [71] proposed an exchange functional where the parameters were fitted to experimental molecular data.

$$\epsilon_X = \epsilon_X^{LDA}\left(1 - \frac{\beta}{2^{1/3}A_x}\frac{x^2}{1 + 6\beta x\sinh^{-1}(x)}\right) \tag{B.5.13}$$

for $x = 2(6\pi^2)^{1/3}, s = 2^{1/3}|\nabla\rho(r)|/\rho(r)^{4/3}$, $A_x = (3/4)(3/\pi)^{1/3}$, and $\beta = 0.0042$. This was complemented by a correlation functional derived by Lee, Yang and Parr (LYP), thus giving rise to a very widely used combination called BLYP functional:

$$\epsilon_C = -\frac{a}{1 + d\rho^{-1/3}}\left\{\rho + b\rho^{-2/3}\left[C_F\rho^{5/3} - 2t_W + \frac{1}{9}\left(t_W + \frac{1}{2}\nabla^2\rho\right)\right]e^{-c\rho^{-1/3}}\right\} \tag{B.5.14}$$

where

$$t_W = \frac{1}{8}\left(\frac{|\nabla\rho|^2}{\rho} - \nabla^2\rho\right) \tag{B.5.15}$$

$C_F = (3/10)(3\pi^2)^{2/3}$, $a = 0.04918$, $b = 0.132$, $c = 0.2533$, and $d = 0.349$. This correlation functional is not based on the LDA.

B.5.2.3 PBE functional and revisions In 1996, Perdew, Burke, and Ernzerhof (PBE) [19, 20] proposed an exchange and correlation functional that satisfies as many formal properties and limits as possible, sacrificing only those deemed to be energetically less important.

The enhancement factor $F_X(\rho, \varsigma, s)$ over the local exchange defined in (B.5.6) depends on the local density ρ, magnetization density ς (in spin-dependent case), and the dimensionless density gradient $s = |\nabla\rho(r)|/(2k_F\rho)$. The chosen expression is

$$F_X(s) = 1 + \kappa - \frac{\kappa}{1 + \mu s^2/\kappa} \tag{B.5.16}$$

where $\mu = \beta(\pi^2/3) = 0.21951$ and $\beta = 0.066725$ is related to the second order gradient expansion.This form:

- satisfies the uniform scaling condition
- recovers the correct uniform gas limit because $F_X(0) = 1$
- obeys the spin-scaling relationship
- recovers the local spin density approximation (LSDA) linear response limit for $s \to 0$
- satisfies the local Lieb-Oxford bound, $\epsilon_X(\mathbf{r}) \geq -1.679\rho(\mathbf{r})^{4/3}$

PBE choose the largest allowed value, $\kappa = 0.804$. Other author have proposed the same form, but with values of κ and μ fitted empirically to a database of atomization energies.

The correlation energy is written in a form similar to an earlier proposal of Perdew and Wang [38, 39, 40, 41, 18]. It assumes the form

$$E_C^{GGA} = \int \rho(\mathbf{r})\left[\epsilon_C^{LDA}(\rho, \varsigma) + H[\rho, \varsigma, t]\right] d\mathbf{r} \tag{B.5.17}$$

with

$$H[\rho, \varsigma, t] = \frac{e^2}{a_0}\gamma\phi^3 \ln\left\{1 + \frac{\beta}{\gamma}t^2\left[\frac{1 + At^2}{1 + At^2 + A^2t^4}\right]\right\} \tag{B.5.18}$$

Here, $t = |\nabla\rho(r)|/(2\phi k_s\rho)$ is a dimensionless density gradient, with k_s the Thomas-Fermi screening wave number, and $\phi(\varsigma) = [(1+\varsigma)^{2/3} + (1-\varsigma)^{2/3}]/2$ is a spin-scaling factor. The quantity β is the same as for the exchange term and $\gamma = (1 - \ln 2)/\pi^2 = 0.031091$. The function A has the following form:

$$A = \frac{\beta}{\gamma}\left[e^{-\epsilon_C^{LDA}[\rho]/(\gamma\phi^3 e^2/a_0)} - 1\right]^{-1} \tag{B.5.19}$$

Trends of the GGAs.

- They improve binding energies and also atomic energies.
- They improve bond lengths and angles.
- They improve energetics, geometries, and dynamic properties of water,ice, and water clusters.
- Semiconductors are marginally better described within the LDA than in GGA, except for the binding energies.

BIBLIOGRAPHY

[1] J. P. M. Lommerse, W. D. S. Motherwell, H. L. Ammon, J. D. Dunitz, A. Gavezzotti, D. W. M. Hofmann, F. J. J. Leusen, W. T. M. Mooij, S. L. Price, B. Schweizer, M. U. Schmidt, B. P. van Eijck, P. Verwer and D. E. Williams, Acta Cryst. (2000). B56, 697-714

[2] W. D. S. Motherwell, H. L. Ammon, J. D. Dunitz, A. Dzyabchenko, P. Erk, A. Gavezzotti, D. W. M. Hofmann, F. J. J. Leusen, J. P. M. Lommerse, W. T. M. Mooij, S. L. Price, H. Scheraga, B. Schweizer, M. U. Schmidt, B. P. van Eijck, P. Verwer and D. E. Williams, Acta Cryst. (2002). B58, 647-661

[3] G. M. Day, W. D. S. Motherwell, H. L. Ammon, S. X. M. Boerrigter, R. G. Della Valle, E. Venuti, A. Dzyabchenko, J. D. Dunitz, B. Schweizer, B. P. van Eijck, P. Erk, J. C. Facelli, V. E. Bazterra, M. B. Ferraro, D. W. M. Hofmann, F. J. J. Leusen, C. Liang, C. C. Pantelides, P. G. Karamertzanis, S. L. Price, T. C. Lewis, H. Nowell, A. Torrisi, H. A. Scheraga, Y. A. Arnautova, M. U. Schmidt and P. Verwer, Acta Cryst. (2005). B61, 511-527

[4] G. M. Day, T. G. Cooper, A. J. Cruz-Cabeza, K. E. Hejczyk, H. L. Ammon, S. X. M. Boerrigter, J. S. Tan, R. G. Della Valle, E. Venuti, J. Jose, S. R. Gadre, G. R. Desiraju, T. S. Thakur, B. P. van Eijck, J. C. Facelli, V. E. Bazterra, M. B. Ferraro, D. W. M. Hofmann, M. A. Neumann, F. J. J. Leusen, J. Kendrick, S. L. Price, A. J. Misquitta, P. G. Karamertzanis, G. W. A. Welch, H. A. Scheraga, Y. A. Arnautova, M. U. Schmidt, J. van de Streek, A. K. Wolf and B. Schweizer, Acta Cryst. (2009). B65, 107-125

[5] D. A. Bardwell, C. S. Adjiman, Y. A. Arnautova, E. Bartashevich, S. X. M. Boerrigter, D. E. Braun, A. J. Cruz-Cabeza, G. M. Day, R. G. Della Valle, G. R. Desiraju, B. P. van Eijck, J. C. Facelli, M. B. Ferraro, D. Grillo, M. Habgood, D. W. M. Hofmann, F. Hofmann, K. V. J. Jose, P. G. Karamertzanis, A. V. Kazantsev, J. Kendrick, L. N. Kuleshova, F. J. J. Leusen, A. V. Maleev, A. J. Misquitta, S. Mohamed, R. J. Needs, M. A. Neumann, D. Nikylov, A. M. Orendt, R. Pal, C. C. Pantelides, C. J. Pickard, L. S. Price, S. L. Price, H. A. Scheraga, J. van de Streek, T. S. Thakur, S. Tiwari, E. Venuti and I. K. Zhitkov, Acta Cryst. (2011). B67, 535-551

[6] Asmadi A, Neumann MA, Kendrick J, Girard P, Perrin MA, Leusen FJ.,J Phys Chem B. 2009 Dec 24;113(51):16303-13.

[7] Marcus A. Neumann, J. Phys. Chem. B, 2008, 112 (32), pp 9810-9829

[8] Rapcewiez, K.; Ashcroft, N. W. Phys. Rev. B 1991, 44, 4032.

[9] Andersson, Y.; Langreth, D. C.; Lundqvist, B. I. Phys. Rev. Lett. 1996, 76, 102.

[10] Dobson, J. F.; Dinte, B. P. Chem. Phys. Lett. 1996, 76, 1780.

[11] Elstner, M.; Hobza, P.; Frauenheim, T.; Suhai, S.; Kaxiras E. J. Chem. Phys. 2001, 114, 5149.

[12] Elstner, M.; Frauenheim, T.; Suhai, S. J. Mol. Struct. (THEOCHEM) 2003, 632, 29.

[13] Wu, Q.; Yang, W. J. Chem. Phys. 2002, 116, 515. Kresse, G.; Hafner, J. Phys. Rev. B 1993, 47, 558.

[14] Marcus A. Neumann and Marc-Antoine Perrin, J. Phys. Chem. B, 2005, 109 (32), pp 15531-15541.

[15] S. Grimme, J. Comp. Chem. 27, 1787 (2006).

[16] Stefan Grimme, Jens Antony, Stephan Ehrlich, and Helge Krieg, J. Chem. Phys. 132, 154104 (2010).

[17] PAM Dirac, (Mathematical) Proceedings of the Cambridge Philosophical Society 26, 376 (1930).

[18] John P. Perdew and Yue Wang, Phys. Rev. B 45, 13244-13249 (1992).

[19] John P. Perdew, Kieron Burke, and Matthias Ernzerhof, Phys. Rev. Lett. 77, 3865-3868 (1996).

[20] John P. Perdew, Kieron Burke, and Matthias Ernzerhof, Phys. Rev. Lett. 78, 1396-1396 (1997).

[21] Yan Zhao and Donald G. Truhlar, Acc. Chem. Res., 2008, 41 (2), pp 157-167.

[22] Yan Zhao and Donald G. Truhlar, J. Chem. Theory Comput., 2006, 2 (4), pp 1009-1018.

[23] Yan Zhao and Donald G. Truhlar, J. Phys. Chem. C, 2008, 112 (11), pp 4061-4067.

[24] Yan Zhao and Donald G. Truhlar, Theoretical Chemistry Accounts May 2008, Volume 120, Issue 1-3, pp 215-241.

[25] P. J. Stephens , F. J. Devlin , C. F. Chabalowski , M. J. Frisch, J. Phys. Chem., 1994, 98 (45), pp 11623-11627.

[26] Chr. Muller and M. S. Plesset, Phys. Rev. 46, 618-622 (1934).

[27] Martin Head-Gordon, John A. Pople, Chemical Physics Letters,Volume 153, Issue 6, 30 December 1988, Pages 503-506.

[28] J. Harris, Phys. Rev. B 31, 1770-1779 (1985).

[29] Kinya Kobayashi, Noriyuki Kurita, Hiroki Kumahora, and Kazutami Tago, Phys. Rev. B 45, 11299-11304 (1992).

[30] Gaussian 09, Revision A.1, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, Gaussian, Inc., Wallingford CT, 2009.

[31] S. Lehtola, ERKALE - HF/DFT from Hel (2012), http://erkale.googlecode.com.

[32] J. Lehtola et al., J. Comp. Chem. 33, 1572 (2012).

[33] T. H. Dunning Jr., J. Chem. Phys., 90 (1989) 1007-23.

[34] R. A. Kendall, T. H. Dunning Jr., and R. J. Harrison, "Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions," J. Chem. Phys., 96 (1992) 6796-806.

[35] D. E. Woon and T. H. Dunning Jr., "Gaussian-basis sets for use in correlated molecular calculations. 3. The atoms aluminum through argon," J. Chem. Phys., 98 (1993) 1358-71.

[36] K. A. Peterson, D. E. Woon, and T. H. Dunning Jr., J. Chem. Phys., 100 (1994) 7410-15.

[37] A. K. Wilson, T. van Mourik, and T. H. Dunning Jr., J. Mol. Struct. (Theochem), 388 (1996) 339-49.

[38] J. P. Perdew, Ed. P. Ziesche and H. Eschrig (Akademie Verlag, Berlin, 1991) 11.

[39] J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiolhais, "Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation," Phys. Rev. B, 46 (1992) 6671-87.

[40] J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiolhais, "Erratum: Atoms, molecules, solids, and surfaces - Applications of the generalized gradient approximation for exchange and correlation," Phys. Rev. B, 48 (1993) 4978.

[41] J. P. Perdew, K. Burke, and Y. Wang, Phys. Rev. B, 54 (1996) 16533-39.

[42] K. Burke, J. P. Perdew, and Y. Wang, in Electronic Density Functional Theory: Recent Progress and New Directions, Ed. J. F. Dobson, G. Vignale, and M. P. Das (Plenum, 1998).

[43] Carlos J. Garcia-Cervera, Commun. Comput. Phys., 2 (2007), pp. 334-357.

[44] Acharya PK, Bartolotti LJ, Sears SB, Parr RG.,Proc Natl Acad Sci U S A. 1980 Dec;77(12):6978-82.

[45] L Goodwin et al 1990 J. Phys.: Condens. Matter 2 351-365.

[46] W Jones and W H Young 1971 J. Phys. C: Solid State Phys. 4 1322.

[47] Y. A. Wang and E. A. Carter, edited by S. D. Schwartz, pp. 117-184 (Kluwer, Dordrecht, 2000).

[48] Gregory S. Ho, Vincent L. Ligneres, Emily A. Carter, Computer Physics Communications Volume 179, Issue 11, 1 December 2008, Pages 839-854.

[49] Stuart C. Watson, Emily A. Carter,Computer Physics Communications Volume 128, Issues 1-2, 9 June 2000, Pages 67-92.

[50] P. Hohenberg and W. Kohn, Phys. Rev. 136, B864, 1964.

[51] Lin-Wang Wang, Michael P. Teter,Phys. Rev. B 45, 13196-13220 (1992).

[52] F Perrot 1994 J. Phys.: Condens. Matter 6 431.

[53] Enrico Smargiassi and Paul A. Madden, Phys. Rev. B 49, 5220-5226 (1994).

[54] David Garcia-Aldea and J. E. Alvarellos, J. Chem. Phys. 129, 074103 (2008).

[55] Michael Foley and Paul A. Madden, Phys. Rev. B 53, 10589-10598 (1996).

[56] E. Chacon,J. E. Alvarellos,P. Tarazona,Phys. Rev. B 32, 7868-7877 (1985).

[57] P. Garcia-Gonzalez, J. E. Alvarellos, and E. Chacon, Phys. Rev. A 54, 1897-1905 (1996).

[58] Yan Alexander Wang, Niranjan Govind, and Emily A. Carter, Phys. Rev. B 58, 13465-13471 (1998).

[59] Yan Alexander Wang, Niranjan Govind, Emily A Carter, Phys. Rev. B 60:16350-16358.

[60] Baojing Zhou, Vincent L Ligneres, Emily A Carter, J. Chem. Phys. 122, 044103 (2005).

[61] Baojing Zhou and Emily A. Carter,J. Chem. Phys. 122, 184108 (2005).

[62] Chen Huang, Emily A. Carter, Phys. Rev. B 81, 045206 (2010).

[63] E. Wigner, Phys. Rev. 46, 1002-1011 (1934).

[64] Thomas, L. H. (1927). Proc. Cambridge Phil. Soc. 23 (5): 542-548.

[65] Fermi, Enrico (1927). Rend. Accad. Naz. Lincei 6: 602-607.

[66] D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. 45, 566-569 (1980).

[67] S. H. Vosko, L. Wilk, M. Nusair, Canadian Journal of Physics, 1980, 58(8).

[68] J. P. Perdew , Alex Zunger, Phys. Rev. B 23, 5048-5079 (1981).

[69] J. P. Perdew , Alex Zunger, Phys. Rev. B 23, 5048-5079 (1981).

[70] D. C. Langreth and M. J. Mehl, Phys. Rev. Lett. 47, 446 (1981).

[71] A. D. Becke, Phys. Rev. A, 38 (1988) 3098-100.

[72] Steven G. Johnson, The NLopt nonlinear-optimization package, http://ab-initio.mit.edu/nlopt.

[72] Miguel A. L. Marques, Micael J. T. Oliveira, and Tobias Burnus, Libxc: a library of exchange and correlation functionals for density functional theory, Comput. Phys. Commun. 183, 2272 (2012)

[73] Schlick T. (2002). Molecular Modeling and Simulation: An Interdisciplinary Guide Springer-Verlag Interdisciplinary Applied Mathematics, Mathematical Biology, Vol. 21. New York, NY.

[74] Sherman, Jack; Morrison, Winifred J. (1949). Annals of Mathematical Statistics 20.

[75] Broyden, C. G. (1970), Journal of the Institute of Mathematics and Its Applications 6: 76-90.

[76] Fletcher, R. (1970),Computer Journal 13 (3): 317-322.

[77] Goldfarb, D. (1970), Mathematics of Computation 24 (109): 23-26.

[78] Shanno, David F. (July 1970), Math. Comput. 24 (111): 647-656.

[79] J. L. Gozquez and J. Robles,J. Chem. Phys. 76, 1467 (1982).

[80] M. K. HARBOLA, Phys. Rev. A, 48 (1993), pp. 2696-2699.

[81] Jianmin Tao,John P. Perdew,Viktor N. Staroverov,Gustavo E. Scuseria,Phys. Rev. Lett. 91, 146401 (2003).

[82] E. Teller, Rev. Mod. Phys., 34 (1962), p. 627.

[83] J.C. Slater, H.M. Krutter, Phys. Rev., 47 (1935), p. 559.

[84] C.F. von Weizsacker, Z. Phys., 96 (1935), p. 431.

[85] C. Herring, Phys. Rev. A, 34 (1986), p. 2614.

[86] E.H. Lieb, Int. J. Quant. Chem., 24 (1983), p. 243.

[87] R.G. Parr, W. Yang, Density-Functional Theory of Atoms and Molecules,Oxford University Press, Oxford (1989).

[88] V. Shah, D. Nehete, D.G. Kanhere, J. Phys.: Condens. Matter, 6 (1994), p. 10773.

[89] N. Govind, J. Wang, H. Guo, Phys. Rev. B, 50 (1994), p. 11175.

[90] H. Jiang, W. Yang, J. Chem. Phys., 121 (2004), p. 2030.

[91] J. Lindhard, Kgl. Danske Videnskab. Selskab Mat.-Fys. Medd., 28 (1954), p. 8.

[92] W. Yang, Phys. Rev. A, 34 (1986), p. 4575.

[93] Wan-Jian Yin, Xin-Gao Gong, Physics Letters A, Volume 373, Issue 4, 19 January 2009, Pages 480-483.

[94] M. Frigo, S. G. Johnson, Proceedings of the IEEE, Vol. 93, No. 2. (February 2005), pp. 216-231.

[95] Sherman, Jack; Morrison, Winifred J. (1949), Annals of Mathematical Statistics 20: 621.

[96] Wolfe, P. (1969), SIAM Review 11 (2): 226-000.

[97] Wolfe, P. (1971), SIAM Review 13 (2): 185-000.

[98] Erin R. Johnson and Axel D. Becke,J. Chem. Phys. 124, 174104 (2006).

[99] Axel D. Becke and Erin R. Johnson, J. Chem. Phys. 123, 154101 (2005).

[100] Axel D. Becke and Erin R. Johnson, J. Chem. Phys. 122, 154104 (2005).

[101] Axel D. Becke and Erin R. Johnson, J. Chem. Phys. 124, 014104 (2006).

[102] Axel D. Becke and Erin R. Johnson, J. Chem. Phys. 127, 154108 (2007).

[103] Ashok Kumar, G. R. G. Fairley, and William J. Meath, J. Chem. Phys. 83, 70 (1985).

[104] J. F. Stanton, Phys. Rev. A 49, 1698 (1994).

[105] M. A. Spackman, J. Chem. Phys. 94, 1295 (1991).

[106] M. Elstner, P. Hobza, S. Suhai, and E. Kaxiras, J. Chem. Phys. 114, 5149 (2001).

[107] U. Zimmerli, M. Parrinello, and P. Koumoutsakos, J. Chem. Phys. 120, 2693 (2004).

[108] K. J. Miller, J. Am. Chem. Soc. 112, 8533 (1990).

[109] Q. Wu and W. Yang, J. Chem. Phys. 116, 515 (2002).

[110] John A. Pople, Peter M.W. Gill and Benny G. Johnson, Chemical Physics Letters, Volume 199, Issue 6, 20 November 1992, Pages 557-560.

[111] Rokhlin, Vladimir (1985), J. Computational Physics Vol. 60, pp. 187-207.

[112] Spall, J. C., Johns Hopkins APL Technical Digest, vol. 19(4), pp. 482-492.

[113] John L. Maryak and Daniel C. Chin, IEEE Transactions on Automatic Control, vol. 53, No. 3, April 2008.

[114] G. te Velde, F.M. Bickelhaupt, S.J.A. van Gisbergen, C. Fonseca Guerra, E.J. Baerends, J.G. Snijders and T. Ziegler. Journal of Computational Chemistry 22, 931 (2001