

# Range-Separated Hybrid Functionals in the Density Functional-Based Tight-Binding Method



Dissertation

zur Erlangung des Doktorgrades  
der Naturwissenschaften (Dr. rer. nat.)  
der Fakultät für Physik  
der Universität Regensburg

vorgelegt von  
Vitalij Lutsker  
aus Ushgorod (Ukraine)  
im Jahr 2015

Promotionsgesuch eingereicht am: 5. Mai 2015

Die Arbeit wurde angeleitet von: Prof. Dr. Thomas Niehaus

Prüfungsausschuss:

Vorsitzender:	Prof. Dr. Jascha Repp
1. Gutachter:	Prof. Dr. Thomas Niehaus
2. Gutachter:	Prof. Dr. Klaus Richter
Weiterer Prüfer:	Prof. Dr. Andreas Schäfer

Termin Promotionskolloquium: 23. Juli 2015

## TABLE OF CONTENTS

<b>Introduction</b>	<b>v</b>
<b>1 Density functional theory and approximations</b>	<b>1</b>
1.1 Hohenberg-Kohn density functional theory . . . . .	2
1.2 Practical DFT: the Kohn-Sham approach . . . . .	3
1.2.1 Local density approximation (LDA) . . . . .	5
1.2.2 Generalized gradient approximation (GGA) . . . . .	5
1.3 DFTB: the approximate PBE-based DFT . . . . .	6
<b>2 Delocalization problem and range-separated hybrid functionals</b>	<b>9</b>
2.1 Self-interaction error . . . . .	9
2.2 Note on the meaning of the Kohn-Sham eigenvalues . . . . .	12
2.3 Solution: Hybrid functionals . . . . .	14
2.3.1 Global hybrids . . . . .	14
2.3.2 Range-separated functionals . . . . .	14
2.3.3 Adiabatic connection functional . . . . .	16
2.4 Matrix equations . . . . .	17
<b>I Implementation</b>	<b>21</b>
<b>3 LC-DFTB: the approximate LC-DFT</b>	<b>23</b>
3.1 Choice of the xc-functional . . . . .	23
3.1.1 Correlation functional . . . . .	24
3.1.2 Short-range exchange functional . . . . .	25
3.2 Second-order energy expansion . . . . .	26
3.3 Two-center approximation and the zeroth-order LC-DFTB . . . . .	28
3.4 Basis set . . . . .	29
3.4.1 Minimal pseudo-atomic basis set (MPA set) . . . . .	30
3.4.2 Basis for LC-DFTB . . . . .	31
3.5 Extension of parametrization tools and computational performance . . . . .	32

TABLE OF CONTENTS

---

3.6	Eigenvalues from the zeroth-order LC-DFTB . . . . .	32
3.7	Self-consistent field LC-DFTB . . . . .	35
3.8	Hubbard parameters . . . . .	39
3.8.1	Influence of the long-range HF exchange term . . . . .	40
3.8.2	Comment on the decay constant . . . . .	43
3.9	Run-time Hamiltonian evaluation . . . . .	43
3.9.1	Direct SCF and thresholding algorithm . . . . .	43
3.9.2	Neighbor list-based algorithm . . . . .	47
3.10	Evaluation of energy gradients . . . . .	48
<b>4</b>	<b>Evaluation of two-electron integrals</b>	<b>53</b>
4.1	Choice of the integrator . . . . .	54
4.2	Description of the scheme and extension to Yukawa interaction . . . . .	54
4.3	Scheme for one-center integrals . . . . .	58
4.4	Scheme for two-center integrals . . . . .	61
<b>5</b>	<b>Parametrization of repulsive potentials for hydrocarbons</b>	<b>65</b>
5.1	Why is the new parametrization needed? . . . . .	65
5.2	Repulsive potential fit procedure . . . . .	66
5.3	Parametrization and test . . . . .	68
<b>6</b>	<b>Summary</b>	<b>71</b>
<b>II</b>	<b>Applications</b>	<b>75</b>
<b>7</b>	<b>Quasi-particle energies from LC-DFTB</b>	<b>77</b>
7.1	Ionization potentials . . . . .	78
7.2	Fundamental gaps . . . . .	81
7.3	Photoemission spectrum . . . . .	83
7.4	Comment on the choice of the range-separation parameter . . . . .	86
7.5	Execution times . . . . .	88
7.6	Summary . . . . .	90
<b>8</b>	<b>Polymers</b>	<b>91</b>
8.1	Bond length alternation . . . . .	92
8.2	Emerging polaron signatures . . . . .	93
8.3	Response to electric field . . . . .	96
8.4	Proteins in gas-phase . . . . .	99
8.5	Summary . . . . .	100

<b>9</b>	<b>Conclusions and open questions</b>	<b>103</b>
<b>III</b>	<b>Appendix</b>	<b>105</b>
<b>A</b>	<b><math>\gamma</math>-Integral over Yukawa interaction</b>	<b>107</b>
A.1	Reduction to one-dimensional integral . . . . .	107
A.2	Analytical evaluation of the off-site $\gamma$ -integral . . . . .	108
A.3	Analytical evaluation of the on-site $\gamma$ -integral . . . . .	109
<b>B</b>	<b>Average potential</b>	<b>111</b>
<b>C</b>	<b>Algorithm for the corrected decay constants</b>	<b>113</b>
<b>D</b>	<b>Geometry data for hydrocarbons</b>	<b>115</b>
<b>E</b>	<b>Neighbor list-based evaluation of forces</b>	<b>121</b>
	<b>Bibliography</b>	<b>123</b>



## INTRODUCTION

The understanding of the electronic structure of atoms, molecules and solids is indispensable for the understanding of the fascinating properties of these physical systems, which justly awaken the interest of both academia and industry. This gives rise to a truly broad spectrum of research activities, including a variety of theoretical approaches. By introducing specific concepts and formalizing them mathematically, the model descriptions arise, which help to deepen the understanding of the condensed matter physics.

Remarkably, the exact electronic structure of a many-electron system can be in principle obtained from the solution of a boundary value problem with the many-body Schrödinger equation. This is the starting point for the methods, based on the concept of the wave function. Together with density functional theory (DFT) they constitute the class of first-principles (*ab initio*) methods, which basically allow to obtain *exact* description of the electronic structure. In practice, however, the exact solution of many-electron systems from these both approaches is not feasible. The reasons in both cases, though are different. While the accuracy of the wave-function based methods is in principle limited by the available computational resources, the computationally low priced DFT in its Kohn-Sham formulation (KS-DFT) relies on inevitable approximations to the exchange-correlation (xc) energy functional. Thus, what practically really matters is the computational effort to accuracy ratio. The result of which quality can be obtained at a given computational cost? From this point of view, the methods, which allow approximate solutions to the exact problem, but at highly reduced cost become increasingly interesting. The density functional-based tight-binding method (DFTB) is one of such methods, which is derived from a specific KS-DFT with a particular approximate exchange-correlation energy functional by applying a set of approximations. These approximations allow to drastically reduce the computational effort, while keeping the accuracy at acceptable level. The reduction of the computational cost is mainly due to the simple structure of the DFTB equations, which depend on a small set of precomputed parameters.

Usually the exchange-correlation functional in the KS-DFT is modeled within the so called local density approximation (LDA) and its generalizations. The simple structure and remarkable accuracy of the exchange-correlation functionals, which emerge from this approach and usually called local functionals, makes them important for practical calculations. However, the KS-DFT with local functionals exhibits systematic failures. This behavior is known as the delocalization problem and is usually connected to the concept of self-interaction error (SIE). In depth investiga-

tion of the delocalization problem resulted in a variety of correction schemes, which successfully deal with it and thus improve the performance of the KS-DFT albeit with moderate increase in computational cost. In particular, the class of range-separated hybrid exchange-correlation functionals deals with the delocalization problem to a great extent. These functionals require the inclusion of a range-dependent Hartree-Fock exchange term, which results in the hybrid Hartree-Fock-DFT calculation.

It is known, that the traditional DFTB, which is derived from the KS-DFT with a local exchange-correlation functional inherits the curse of the delocalization problem [43]. The success of the KS-DFT with range-separated hybrid functionals gives rise to the expectation, that the approximate methods, which are derived from it will be less prone to the delocalization problem. In ref. [137] the usual DFTB approximations have been formally applied to a DFT with a long-range corrected functional (LC-DFT), which is a special case of a range-separated functional (compare also section 2.3 of this thesis). The legitimacy of such approximations with respect to the new functional has been provided. This is the starting point of the work, presented in this thesis, which is organized as follows.

In chapter 1 we introduce the DFT as initially formulated by Hohenberg and Kohn and its formulation as a single-particle theory, given by Kohn and Sham (KS-DFT). We give brief comment on the local density approximation and generalized gradient approximation. We also briefly introduce the traditional DFTB method.

In chapter 2 we discuss the delocalization problem and comment on the hybrid functionals, which are usually able to deal with the delocalization problem. In particular, the definition of the range-separated functionals is given.

The technical part I of this thesis is dedicated to the implementation and parametrization of the new scheme. The specific parametrization, which is used in this thesis is called the long-range corrected DFTB (LC-DFTB). The in depth description of the method, discussion on the choice of the underlying exchange-correlation functional, specific aspects of the parametrization and numerical efficiency of the current run-time implementation are found in chapter 3. The DFTB method by construction relies on a set of precomputed parameters. The modification of parametrization programs is necessary in order to perform the parametrization of the DFTB method with a new functional. We address the aspects of these modifications in chapter 4, where the integration routines for the evaluation of specific two-electron integrals are described. In chapter 5 the parametrization of the repulsive potential, in particular important for the geometry optimization for the elements C and H is described.

In the part II we apply the LC-DFTB method to a series of problems, which are usually insufficiently described by the DFT with local functionals and by traditional DFTB. In chapter 7 for the first time the LC-DFTB method is applied to a series of organic molecules. We discuss the frontier orbital energies as obtained from the new method and compare them to the experimental data (where available), the traditional DFTB method and the first-principles DFT with local



---

and with long-range corrected functionals. We furthermore discuss the possibility of obtaining photoemission spectra from LC-DFTB. The comparison of computational times of LC-DFTB against first-principles approaches and traditional DFTB is addressed as well. In chapter 8 we apply the LC-DFTB to *trans*-polyacetylene oligomers and study again quantities, which are problematic for the traditional DFTB method. This includes the difference in single and double bond lengths also known as bond length alternation, response to the electric field (specifically the static longitudinal polarizability) and the formation of polaronic defects in the doped polymer. Finally, the application of LC-DFTB to the two selected proteins in zwitterionic conformation in gas-phase is performed.

In the appendix, the supplementary information to the chapters of the main text can be found. In appendix A the analytical formulas for the long-range  $\gamma$ -integral, which is important in the new method, are derived. In appendix B the algorithm for the evaluation of the average potential from converged orbitals, which is used for a calculation in chapter 2 is described. This method was initially mentioned in ref. [9]. In appendix C the numerical algorithm for the correction of decay constants (see also section 3.8) is presented. The appendix D accompanies the description of the repulsive potential parametrization in chapter 5 and contains the summary of bond lengths and angles of a benchmark set of selected hydrocarbons as obtained from the LC-DFTB and other methods, compared to experiment. The vibrational frequencies of the reference molecules, used for the parametrization, are presented as well. Finally in appendix E the pseudo code and corresponding equations for the neighbor list-based evaluation of the energy gradients (section 3.10) are presented, which accompany the source code of the implementation in the DFTB+ code [6].



## DENSITY FUNCTIONAL THEORY AND APPROXIMATIONS

In quantum theory the central concept is the wave function  $\Psi$ . The energy of a physical system is the functional of the wave function and the variational principle holds

$$E[\Psi] \geq E[\Psi_0], \quad (1.1)$$

where the wave function  $\Psi_0$  represents the ground state and the equality holds for  $\Psi = \Psi_0$ . The wave function is the solution of the stationary Schrödinger boundary value problem

$$\hat{H}\Psi = E\Psi. \quad (1.2)$$

Here  $\hat{H}$  is the Hamilton operator, which can be constructed in an intuitive way from the classical pendants via correspondence principle. For example, a  $N$ -electron system in non-relativistic limit and in Born-Oppenheimer approximation is described by the many-body Hamiltonian in the real space representation <sup>1</sup>

$$\hat{H} = \underbrace{-\frac{1}{2} \sum_{i=1}^N \nabla_i^2}_{\hat{T}} + \underbrace{\frac{1}{2} \sum_{i \neq j}^N \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}}_{\hat{V}_{ee}} + \underbrace{\sum_{i=1}^N \sum_{A=1}^M \frac{-Z_A}{|\mathbf{r}_i - \mathbf{R}_A|}}_{\hat{V}_{\text{ext}}}, \quad (1.3)$$

where  $\mathbf{r}_i$  is the coordinate of the  $i$ -th electron,  $Z_A$  and  $\mathbf{R}_A$  are the charge and coordinate of the nuclei  $A$ . The corresponding wave function  $\Psi = \Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)$  depends on coordinates  $\mathbf{x}_i$  of the  $i$ -th electron, which comprise space coordinates  $\mathbf{r}_i$  and spin coordinates  $\omega_i$ . From the wave function the expectation value of each observable can be obtained. The energy functional is the expectation value of the Hamilton operator, provided  $\langle \Psi | \Psi \rangle = 1$

$$E[\Psi] = \langle \Psi | \hat{H} | \Psi \rangle = \langle \Psi | \hat{T} | \Psi \rangle + \langle \Psi | \hat{V}_{ee} | \Psi \rangle + \langle \Psi | \hat{V}_{\text{ext}} | \Psi \rangle = T[\Psi] + V_{ee}[\Psi] + \int \rho(\mathbf{r}) v^{\text{ext}}(\mathbf{r}) d\mathbf{r}. \quad (1.4)$$

<sup>1</sup>in this thesis we use atomic units if not stated otherwise.

Note, that the expectation value of the kinetic energy operator  $T[\Psi] = \langle \Psi | \hat{T} | \Psi \rangle$  and the electron-electron interaction energy operator  $V_{ee}[\Psi] = \langle \Psi | \hat{V}_{ee} | \Psi \rangle$  depend only on the number of electrons  $N$ . The variational principle, the particle number  $N$  and the external potential determine the ground state wave function.

This approach to the exact description of many-electron systems like atoms or molecules, however, turns out to be problematic in practice. For large electron numbers it faces, what W. Kohn calls the exponential wall [97]. The effort to store and construct an accurate approximation to the exact wave function, as well as to minimize the corresponding energy functional grows exponentially for increasing system size. Thus usually different levels of approximations are introduced. One of the most famous in this context is the Hartree-Fock (HF) approximation. For a given many-body Hamiltonian it is the wave function which is approximated. One assumes that the system is described by an antisymmetric product of single-particle wave functions (Slater determinant). The HF approximation results in the theory which includes the classical electron repulsion and quantum mechanical Pauli-principle. However, the purely quantum mechanical electron-electron repulsion, known as correlation energy, is not included in the HF theory.<sup>2</sup> Nevertheless, it is the starting point for the correlated wave function-based methods, generally called post-Hartree-Fock methods. This includes the configuration interaction (CI) [181], perturbation theory (MP2) [128], Hedin approximation (GW) [7, 66] and coupled cluster (CC) [130] methods. Despite their success and high accuracy these methods still have tremendous computational demands. They are not practicable for extended systems. In fact their high computational requirements are still related to the fact, that the approximate representation of a highly complex exact wave-function beyond single Slater determinant is exploited. The alternative approach due to Hohenberg and Kohn is described in following.

## 1.1 HOHENBERG-KOHN DENSITY FUNCTIONAL THEORY

As has been already stated in the wave function-based methods the knowledge of the wave function implies knowledge over every observable of the system in the corresponding state. So, for example the density of a system, which can be directly measured, is obtained from the wave function according to

$$\rho(\mathbf{r}_1) = N \int |\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)|^2 d\omega_1 d\mathbf{x}_2 \dots d\mathbf{x}_N. \quad (1.5)$$

Hohenberg and Kohn asked, whether the knowledge of the ground state density implies knowledge over every observable in the ground state? With other words, does the density uniquely determine the ground state of a physical system? They showed, that the description of a  $N$ -electron system by a wave function and corresponding Schrödinger equation can be replaced by an equivalent theory, based on the electronic density [72].

---

<sup>2</sup>The correlation energy  $E_c$  is often defined as the difference of exact and Hartree-Fock energies.

The theory is based on two theorems. The first one states: The external potential  $v^{\text{ext}}(\mathbf{r})$  is determined by the ground state electron density  $\rho_0(\mathbf{r})$  up to a trivial additive constant. Thus a ground state density, which yields  $N = \int \rho(\mathbf{r}) d\mathbf{r}$  determines the wave function and thus all ground state properties of an  $N$ -electron system. The proof of the theorem can be found in ref. [72]. The total energy of the system is the functional of density

$$E[\rho] = F[\rho] + \int \rho(\mathbf{r})v^{\text{ext}}(\mathbf{r}) d\mathbf{r}, \quad (1.6)$$

where the functional  $F[\rho] = T[\rho] + V_{\text{ee}}[\rho]$  is composed of kinetic and electron interaction energies, which itself are functionals of density. This functional is called the universal functional for a  $N$ -electron system, since it does not depend on the external potential. The second theorem constitutes the variational principle. Given a density  $\rho'$  and the ground state density  $\rho_0$ , for the fixed external potential the energy functional obeys the inequality

$$E[\rho'] \geq E[\rho_0], \quad (1.7)$$

where the equality holds for the case  $\rho' = \rho_0$ . Thus, ground state energy and the corresponding density can be obtained from

$$\delta \left\{ E[\rho] - \mu \int \rho(\mathbf{r}) d\mathbf{r} \right\} = 0 \quad (1.8)$$

with some Lagrange multiplier  $\mu$ , which is characteristic to the system and can be identified as the chemical potential of the system [148]. In this case the density corresponds to the global minimum of the energy functional.

The simplicity of the Hohenberg-Kohn theory faces a serious challenge, which is still not overcome. The explicit form of the universal functional is not known and is suspected to be complicated. In this respect the universal functional plays the role of the exact wave function in the wave function-based methods.

In following we present the approach due to Kohn and Sham [98]. They show an elegant way how to relate an effective single particle theory to the Hohenberg-Kohn theory. The “unknown and complicated” in their approach is placed into a small part of the universal energy functional, while the rest is treated on the same footing as in the wave function-based single particle theory (like HF). Despite it’s single particle form, it is exact.

## 1.2 PRACTICAL DFT: THE KOHN-SHAM APPROACH

We now show how to come from the Hohenberg-Kohn energy functional to a practical scheme. The Hohenberg-Kohn energy functional can be defined in the following way [111, 112]

$$E[\rho] = \underbrace{\min_{\Psi \rightarrow \rho} \langle \Psi | \hat{T} + \hat{V}_{ee} | \Psi \rangle}_{F[\rho]} + \int \rho(\mathbf{r})v^{\text{ext}}(\mathbf{r}) d\mathbf{r}, \quad (1.9)$$

where the  $F[\rho]$  is the universal functional, which depends only on the number of particles. Here the notation  $\Psi \rightarrow \rho$  means that the minimization process runs over all valid many-body wave functions, which yield the density  $\rho$ . The universal functional is not known. The idea is to represent the functional by some known model, which contributes the most to the energy and the rest, which is then approximated. The universal functional for a non-interacting fermionic system is

$$F^0[\rho] = \min_{\Phi \rightarrow \rho} \langle \Phi | \hat{T} | \Phi \rangle = T^0[\rho], \quad (1.10)$$

where the wave function  $\Phi = |\phi_1 \phi_2 \dots \phi_N|$  is the Slater determinant. The density can be expressed in terms of the single-particle orbitals  $\rho(\mathbf{r}) = \sum_i^N |\phi_i(\mathbf{r})|^2$ . The key point in this procedure is that the minimization is performed over the Slater determinants, which yield the same density as the interacting wave functions in Eq. 1.9. For more detailed discussion of the foundations of the Hohenberg-Kohn DFT and Kohn-Sham DFT (KS-DFT) we refer to ref. [143]. The energy is then expressed in terms of classical Coulomb interaction (Hartree term)  $E_H[\rho] = \int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}d\mathbf{r}'$ , the non-interacting kinetic energy  $T^0 = \sum_i \int \phi_i(\mathbf{r})(-\frac{1}{2}\nabla^2)\phi_i(\mathbf{r})d\mathbf{r}$ , the energy due to the interaction with the external potential  $v^{\text{ext}}(\mathbf{r})$  and the rest, which is referred to as the exchange-correlation (xc) energy  $E_{\text{xc}}$

$$E[\rho] = T^0[\rho] + E_H[\rho] + \int v^{\text{ext}}(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} + E_{\text{xc}}[\rho]. \quad (1.11)$$

The  $E_{\text{xc}}$  accounts for quantum mechanical electron-electron interaction effects. We rewrite the universal energy functional

$$F[\rho] = T^0[\rho] + E_H[\rho] + \underbrace{[T[\rho] + V[\rho] - T^0[\rho] - E_H[\rho]]}_{E_{\text{xc}}[\rho]}, \quad (1.12)$$

where the last term defines the exchange-correlation functional. The minimization due to the density in Eq. 1.8 can be replaced by an equivalent minimization due to a set of orbitals, which constitute a Slater determinant. This gives the set of  $N$  Kohn-Sham equations

$$\left(-\frac{1}{2}\nabla^2 + v_{\text{KS}}(\mathbf{r})\right)\phi_i(\mathbf{r}) = \epsilon_i\phi_i(\mathbf{r}) \quad (1.13)$$

with a local effective Kohn-Sham potential

$$v_{\text{KS}}(\mathbf{r}) = v^{\text{ext}}(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' + v_{\text{xc}}(\mathbf{r}), \quad (1.14)$$

where the exchange-correlation potential is defined as the functional derivative of the exchange-correlation energy  $v_{\text{xc}}(\mathbf{r}) = \frac{\delta E_{\text{xc}}}{\delta \rho}$ . Thus the interacting system can be mapped onto the non-interacting with modified local effective potential. We again emphasize, that the KS-DFT is by construction exact. However, in practice the unknown exchange-correlation functional has still to be approximated. These approximations lead to some spectacular failures of DFT, which we will discuss in chapter 2.

### 1.2.1 Local density approximation (LDA)

Kohn and Sham proposed a simple approximate form of the exchange-correlation functional, expressed in terms of density  $\rho(\mathbf{r})$  and exchange-correlation energy density per particle  $\epsilon_{xc}^{\text{LDA}}(\mathbf{r})$  of an uniform electron gas with density  $\rho$

$$E_{xc}^{\text{LDA}} = \int \epsilon_{xc}^{\text{LDA}}[\rho](\mathbf{r})\rho(\mathbf{r}) d\mathbf{r}. \quad (1.15)$$

The latter is split into exchange and correlation parts,

$$\epsilon_{xc}^{\text{LDA}}[\rho](\mathbf{r}) = \epsilon_x^{\text{LDA}}[\rho](\mathbf{r}) + \epsilon_c^{\text{LDA}}[\rho](\mathbf{r}). \quad (1.16)$$

The exchange part is given by the Slater formula for uniform electron gas

$$\epsilon_x^{\text{LDA}}[\rho](\mathbf{r}) = -\frac{3}{4} \left( \frac{3\rho(\mathbf{r})}{\pi} \right)^{1/3}. \quad (1.17)$$

The explicit expression for the correlation part  $\epsilon_c^{\text{LDA}}[\rho](\mathbf{r})$  is not known. Instead it is parametrized from the available numerical quantum Monte-Carlo data [30].

### 1.2.2 Generalized gradient approximation (GGA)

While the LDA functionals depend only on the density at the evaluation point, the generalized gradient approximation functionals use the information about the density behavior around the evaluation point. This is done by introducing the explicit dependence on the density gradients in addition to the density. This additional information leads to a significant qualitative improvement over the LDA functionals in terms of atomization energies, exchange and correlation energies, reaction barriers. The GGA functionals are usually formulated as the extended LDA functional

$$E_{xc}^{\text{GGA}} = \int \rho(\mathbf{r})\epsilon_{xc}^{\text{LDA}}[\rho](\mathbf{r})F(s) d\mathbf{r} \quad (1.18)$$

with the GGA enhancement factor  $F(s)$ , which depends on the reduced density gradient  $s = \frac{|\nabla\rho|}{\rho} \frac{1}{2(3\pi^2\rho)^{1/3}}$ . The GGA functionals in general outperform the LDA functionals (see for example refs. [1, 34, 106]).

The methods, presented in this thesis have been developed based on the LDA functional and GGA functional of Perdew, Burke and Ernzerhof (PBE) [146]. The latter is defined for the case of spin unpolarized system as

$$E_x^{\text{PBE}} = \int \rho(\mathbf{r})\epsilon_x^{\text{LDA}}(\mathbf{r}) \left[ 1 + \kappa - \frac{\kappa}{1 + (\mu/\kappa)s^2} \right] d\mathbf{r} \quad (1.19)$$

$$E_c^{\text{PBE}} = \int \rho(\mathbf{r}) \left[ \epsilon_c^{\text{LDA}}(\mathbf{r}) + \gamma \ln \left( 1 + \frac{\beta}{\gamma} \frac{t^2 + At^4}{1 + At^2 + A^2t^4} \right) \right] d\mathbf{r}, \quad (1.20)$$

where the constants  $\mu = 0.21951$ ,  $\kappa = 0.804$ ,  $\beta = 0.066725$ ,  $\gamma = 0.031091$  appear and  $t = |\nabla\rho|(2k_s\rho)^{-1}$ ,  $k_s = \sqrt{\frac{4k_F}{\pi a_0}}$ ,  $k_F = (3\pi^2\rho)^{1/3}$  and  $A = \frac{\beta}{\gamma} [\exp(-\epsilon_c^{\text{LDA}}/(\gamma e^2/a_0)) - 1]^{-1}$ . The functional is developed to fulfill a set of exact conditions of the exchange-correlation functional and contains no empirical parameters.

## 1.3 DFTB: THE APPROXIMATE PBE-BASED DFT

The Kohn-Sham method is implemented in nearly every general purpose quantum chemistry program package. The program packages use highly optimized algorithms which are the result of a half a century of intensive scientific research. Still, the applicability of the Kohn-Sham DFT, dependent on the desired information extent and its precision (provided we are interested in qualitatively meaningful results), is restricted to the systems up to a several hundreds of atoms in size. In order to access the systems which are well beyond the computational limits of the DFT at the given moment of time the approximate methods are developed.

In following we introduce the method, which can be classified as an approximate Kohn-Sham DFT, the density functional-based tight-binding method (DFTB) [42, 46, 155, 175]. Formally it is similar to the empirical tight-binding schemes [56, 76], but in contrast to these it is derived from the DFT with LDA or GGA xc-functional. Thus in fact it is the approximated DFT, rather than a model, parametrized from empirical data.

The DFTB approximations lead to the energy, which is formulated in terms of two-center parameters, which depend on the interatomic distance  $R_{AB}$

$$E = \sum_i^{\text{occ basis}} \sum_{\mu\nu} c_{\mu i} c_{\nu i} H_{\mu\nu}^0(R_{AB}) + \frac{1}{2} \sum_{AB}^{\text{atoms}} \gamma_{AB}(R_{AB}) \Delta q_A \Delta q_B + \frac{1}{2} \sum_{AB}^{\text{atoms}} V_{AB}(R_{AB}). \quad (1.21)$$

The coefficients  $c_{\mu i}$  represent the molecular Kohn-Sham orbital  $\psi_i$ , which is approximated by a linear combination of atom-centered functions  $\phi_\mu$  (LCAO ansatz)

$$\psi_i(\mathbf{r}) = \sum_{\mu}^{\text{basis}} c_{\mu i} \phi_{\mu}(\mathbf{r}). \quad (1.22)$$

These functions are orthogonal on each atom, but not orthogonal if they are on distinct atoms. Their overlap integral is denoted by  $S_{\mu\nu} = \int \phi_{\mu}(\mathbf{r}) \phi_{\nu}(\mathbf{r}) d\mathbf{r}$ . The molecular orbital (MO) coefficients  $c_{\mu i}$  are obtained by solving the generalized eigenvalue problem

$$\sum_{\nu} \left( H_{\mu\nu}^0 + \frac{1}{2} S_{\mu\nu} \sum_C^{\text{atoms}} (\gamma_{AC} + \gamma_{BC}) \Delta q_C \right) c_{\nu i} = \epsilon_i \sum_{\nu} S_{\mu\nu} c_{\nu i}, \quad (1.23)$$

which can be derived by applying the variational principle to the DFTB energy expression with respect to the MO coefficients. Here the Mulliken atomic charges are defined as

$$\Delta q_A = \frac{1}{2} \sum_{\mu \in A} \sum_{\nu}^{\text{occ}} (c_{\mu i} c_{\nu i} S_{\mu\nu} + c_{\nu i} c_{\mu i} S_{\nu\mu}) - q_A^0, \quad (1.24)$$

where  $q_A^0$  is the number of valence electrons of the atom  $A$ . The parameters  $\gamma_{AB}$  for the case  $A \neq B$  are parametrized as [42]

$$\gamma_{AB}(R) = \frac{1}{R} - \left[ e^{-\tau_A R} \left( \frac{\tau_B^4 \tau_A}{2(\tau_A^2 - \tau_B^2)^2} - \frac{\tau_B^6 - 3\tau_B^4 \tau_A^2}{(\tau_A^2 - \tau_B^2)^3 R} \right) + e^{-\tau_B R} \left( \frac{\tau_A^4 \tau_B}{2(\tau_B^2 - \tau_A^2)^2} - \frac{\tau_A^6 - 3\tau_A^4 \tau_B^2}{(\tau_B^2 - \tau_A^2)^3 R} \right) \right], \quad (1.25)$$



where the parameter  $\tau_A$  is obtained from the condition that the Hubbard derivative  $U$  for a single atom  $A$  calculated from the reference DFT-PBE calculation should be equal to the Hubbard parameter, obtained from the DFTB calculation

$$U = \frac{\partial^2 E^{\text{atom,PBE}}}{\partial n^2} \stackrel{!}{=} \frac{\partial^2 E^{\text{atom,DFTB}}}{\partial \Delta q^2} = \gamma_{AA} = \frac{5}{16} \tau_A, \quad (1.26)$$

where  $\gamma_{AA}$  is the on-site  $\gamma$ -parameter and the derivative  $\frac{\partial^2 E^{\text{atom,PBE}}}{\partial n^2}$  is performed numerically with respect to the occupation number  $n$  of the highest occupied molecular orbital (HOMO). For the DFTB the Hubbard derivative is known analytically and equals to the on-site  $\gamma$ -parameter  $\gamma_{AA} = \frac{5}{16} \tau_A$ . The distance dependent  $\gamma$ -parameters represent the approximated electron repulsion integrals and effectively contain the contributions due to the exchange-correlation potential. The zeroth-order Hamiltonian  $H_{\mu\nu}^0$  is defined as

$$H_{\mu\nu}^0 = \begin{cases} \epsilon_{\mu}^{\text{free atom}} & \mu = \nu \\ \langle \phi_{\mu} | H^{\text{DFT-PBE}}[\rho_A + \rho_B] | \phi_{\nu} \rangle & \mu \in A, \nu \in B \\ 0 & \text{else,} \end{cases} \quad (1.27)$$

where  $\epsilon_{\mu}^{\text{free atom}}$  is the eigenvalue of a spherically symmetric free pseudo-atom, obtained from the DFT-PBE calculation, and  $H^{\text{DFT-PBE}}[\rho_A + \rho_B]$  is the DFT-PBE Hamiltonian evaluated at the reference density  $\rho = \rho_A + \rho_B$ , which is a simple superposition of atomic densities  $\rho_A, \rho_B$ . Finally, the repulsive pair-potentials  $V_{AB}$  for a pair of atomic species are obtained by a fit to the total energies of a reference DFT method.

The DFTB method operates with precomputed parameters and simple analytical formulas. These quantities, apart from the repulsive potential, are directly connected to the quantities in KS-DFT, which is being approximated by the DFTB. This will become more evident in chapter 3, where we derive the DFTB method as an approximation of a long-range corrected hybrid DFT (section 2.3), which contains the PBE-based DFTB (standard DFTB) as a special case. Together with small number of basis functions per atom the efficient large scale calculations ( $\approx 10000$  atoms) are possible. This opens the possibility to study large biological systems, perform the potential energy scans, do the molecular dynamics simulations for both extended systems and large time scales [35, 42, 43, 50].



## DELOCALIZATION PROBLEM AND RANGE-SEPARATED HYBRID FUNCTIONALS

The GGA and LDA functionals, mentioned in previous chapter have an excellent accuracy to computational cost ratio and are usually easy to implement. They depend on the density and its gradients in an explicit way and usually called local or semilocal functionals. However, the DFT with this class of exchange-correlation functionals exhibits remarkable failures. Incorrect dissociation limits of molecules [11, 205], instability of anions [184, 199], absence of Rydberg series [9], underestimation of reaction barriers [34], overestimation of the electric field response [196, 205], generally bad description of localized states [45, 192] are often mentioned in this context. The approximations to the DFT with local xc-functionals, such as the DFTB method (section 1.3) inherit this erroneous behavior [77, 118, 137, 158].

### 2.1 SELF-INTERACTION ERROR

The aforementioned failures have a common root: the inability of the DFT with local xc-functionals to deal with the self-interaction error (SIE), also known as delocalization problem. The standard illustration of this behavior is the dissociation of a  $\text{H}_2^+$  molecule ion [11]. Local DFT predicts two protons at infinite separation with half an electron around each of them. The self-interaction of the only electron, which is by construction included in the Hartree energy should be cancelled by the exact exchange-correlation energy. In the HF-theory it is the case and the dissociation limit is correct. The approximate form of the xc-functional in the local DFT, however, gives rise to spurious, unphysical self-interaction. Extending this reasoning to the systems with many electrons the functional is called one-electron self-interaction free, if it fulfills the condition

$$E_{\text{xc}}[\rho_i] + E_H[\rho_i] = 0, \quad (2.1)$$

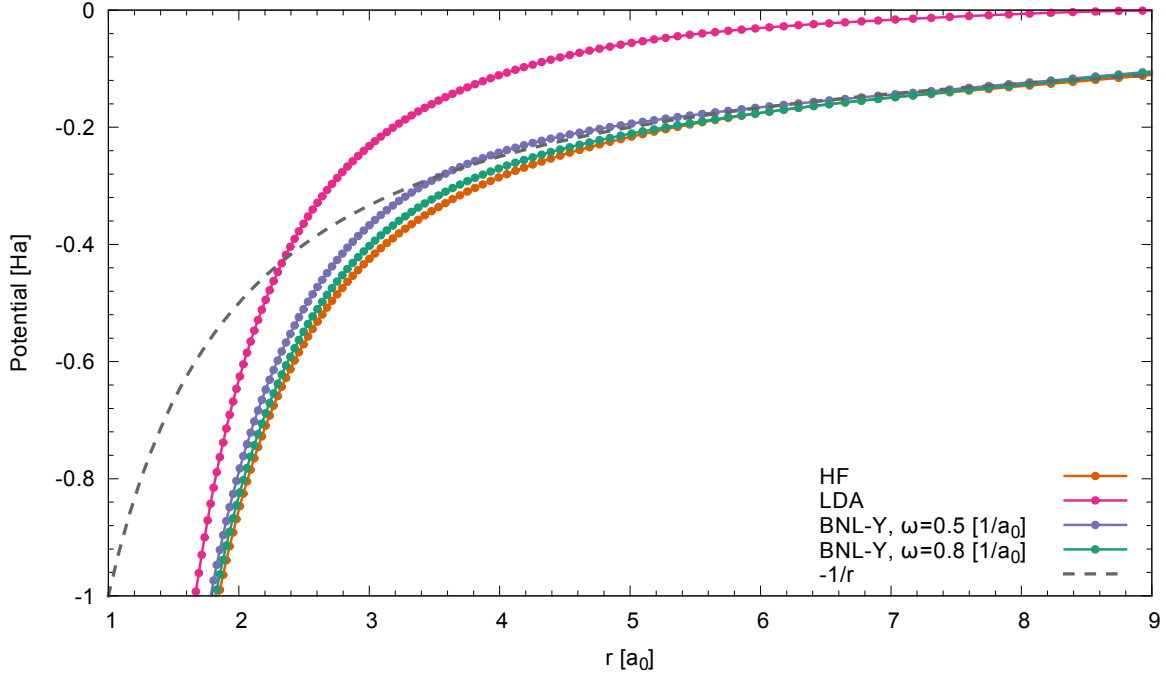


Figure 2.1: The LDA potential and averaged HF and BNL-Y (section 3.1) potentials for a single argon atom as a function of distance from nuclei. The gray dashed line indicates the correct  $-1/r$  limit.

where  $\rho_i = |\psi_i|^2$  is the one-electron density [150]. It is in general expected, that the self-interaction error should be connected to the properties of the xc-potential. For example, physically it is expected, that an electron of a finite charge-neutral system, if removed to infinity (the system is ionized), interacts with the hole, which is created in this electron removal process. This implies, that the potential, which is seen by the electron should exhibit the  $-1/r$  asymptotics [3]

$$\lim_{|\mathbf{r}|\rightarrow\infty} v_{\text{xc}}(\mathbf{r}) = -\frac{1}{|\mathbf{r}|}. \quad (2.2)$$

In Fig. 2.1 the LDA potential of an argon atom as a function of distance from nucleus is depicted. It decays exponentially, much faster than the correct  $-1/r$  limit, indicated here by the gray dashed line. On contrast, the average potentials from the Hartree-Fock theory and the long-range corrected functional BNL-Y, which will be introduced in section 3.1 show correct long-range behavior. The potentials have been obtained according to the procedure, suggested in ref. [9] and the details of the calculation can be found in appendix B. The correct asymptotic decay of the xc-potential is an important condition for the improvement of the theory [9, 14, 29, 193].

It has been found, however, that the functionals, which are one-electron self-interaction free, in the case of many-electron systems can still exhibit failures, related to the self-interaction error [17, 34, 131]. While the concept of one-electron self-interaction error is simple from the physical point, it is not the best choice to understand and mathematically formalize the self-interaction

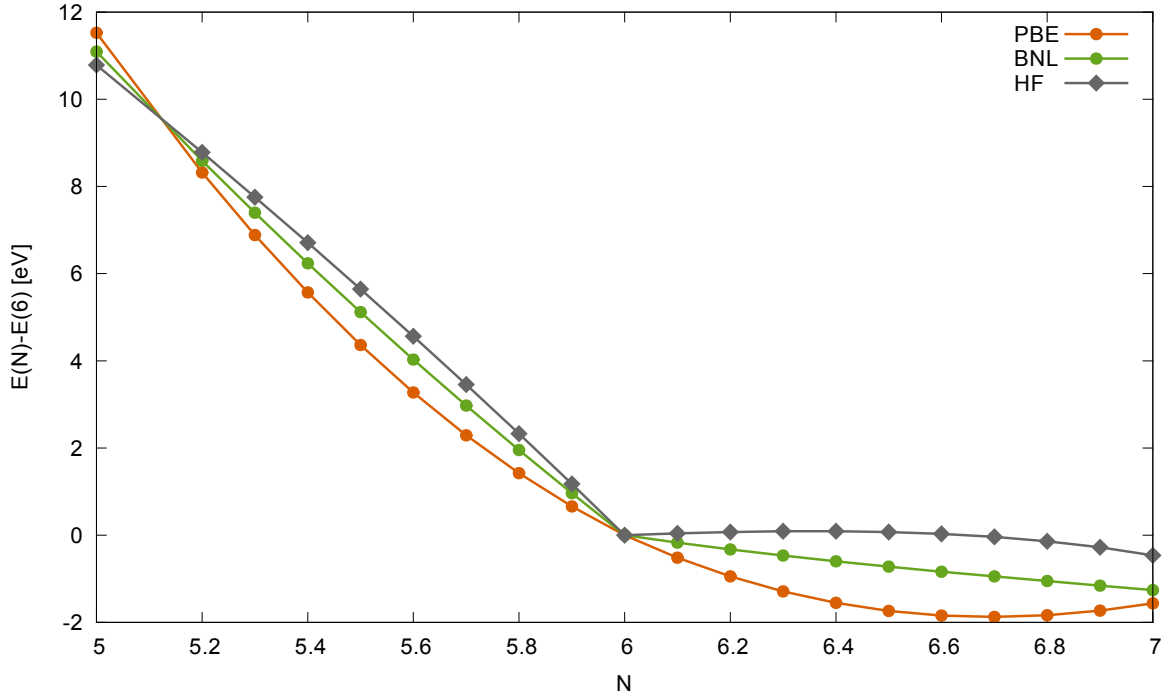


Figure 2.2: Total energy of a carbon atom as a function of electron number  $N$  for DFT with local PBE and long-range corrected BNL functionals and HF theory. Calculations have been performed with NWChem package [195] in spin unrestricted formalism with aug-cc-pVDZ basis set. The number of up-electrons was varied from 3 to 5, while number of down-electrons was kept equal 2. Note, that the curves have been shifted such that  $E(6) = 0$  for each method.

error in many-electron systems. Another approach to this problem can be made by considering fractionally occupied systems. While such systems are not physical, their study gives deeper insights into the delocalization problem and allows to formulate conditions for the functional design. The main result in this respect has been provided by Perdew et al. [148]. They showed that the ground-state energy of a statistical mixture of an  $N$ -electron pure state  $\Psi_N$  with an  $(N+1)$ -electron pure state  $\Psi_{N+1}$  with corresponding density  $\rho(\mathbf{r}) = (1-\omega)\rho_N(\mathbf{r}) + \omega\rho_{N+1}(\mathbf{r})$  is given by

$$E = (1-\omega)E_N + \omega E_{N+1}. \quad (2.3)$$

Here  $E_N$  and  $E_{N+1}$  are the ground state energies of the  $N$  and  $(N+1)$ -particle system respectively and  $0 \leq \omega \leq 1$ . This mixture is then a  $(N+\omega)$ -particle system. The exact ground state energy as a function of fractional occupation is piece-wise linear. It is linear between two adjacent integer particle numbers and exhibits jumps in the derivative at integer occupations. In Fig. 2.2 an example calculation for a carbon atom is carried out. The total energy of a fractionally occupied carbon atom as a function of electron number  $N$  for the local PBE functional [146], the HF theory and the long-range corrected BNL functional [10, 116] are plotted. The curves have been shifted

such that the energy of a neutral carbon with number of electrons  $N = 6$  for each theory is at 0. We clearly observe the non-linear behavior for PBE and HF theories. The local DFT gives usually good estimates to the ground state energies of systems with integer particle numbers. However, fractionally occupied systems have too low energies and thus are favoured by the local DFT. This is the reason for the wrong dissociation of the  $\text{H}_2^+$  ion. The energy of two infinitely separated fractionally occupied  $\text{H}^{0.5+}$  atoms is lower than the energy of the  $\text{H}^+ + \text{H}$  system. One could also say, that the local DFT tends to overly delocalize the density. At the same time the HF theory gives too high energies for fractionally occupied systems [58, 149, 205]. For the molecule dissociation this is not so crucial, since the integer occupations are favoured. The HF shows tendency to overly localize the density.

From the relation between the energy derivative with respect to the occupation  $n_i$  of a single particle Kohn-Sham (KS) orbital and the energy of this orbital  $\epsilon_i$

$$\frac{\partial E^{\text{total}}}{\partial n_i} = \epsilon_i, \quad (2.4)$$

also known as Janak's theorem, one can identify the highest occupied molecular orbital (HOMO) as the chemical potential of the system [87, 150]. The negative of the HOMO eigenvalue from the exact KS-theory equals to the ionization potential (IP) of the system. Indeed from Fig. 2.2 we see that the slope of the energy curve at  $N = 6$  from the left derivative is too low for the local DFT and too high for the HF theory if compared to the assumed linear behavior (we assume that the energies for integer particle numbers are correct). This means, that the HOMO eigenvalue from PBE underestimates the IP and the HF in contrast overestimates it. It has been observed, that for many cases, where the local DFT systematically underestimates a quantity, the HF overestimates it (and vice versa). For example the bond-length alternation in the conjugated polymer chains is overestimated by the HF theory, while the local DFT usually underestimate it (compare also results of section 8.1 in this thesis). The long-range corrected (range-separated) functionals, like BNL show almost linear energy dependence on fractional charge and have balanced behavior. They usually outperform both, local DFT and HF in cases, where delocalization problem gives rise to massive errors.

The currently accepted definition of the self-interaction free functional is the one which fulfills the linearity condition [34]. For this reason the condition  $-\epsilon_{\text{HOMO}} = \text{IP}$  is an important indicator for the correction of the self-interaction error.

## 2.2 NOTE ON THE MEANING OF THE KOHN-SHAM EIGENVALUES

By construction the KS-eigenvalues do not have any physical meaning, with HOMO eigenvalue being an important exception. Despite this fact the connection between the Kohn-Sham HOMO-LUMO gap ( $E_{\text{gap}}^{\text{KS}} = \epsilon_{\text{LUMO}} - \epsilon_{\text{HOMO}}$ ) and experimental gaps has been extensively studied [86, 101, 169, 203]. Surprisingly it has been found, that the optical gap in some cases is very accurately described by the HOMO-LUMO gap from hybrid functionals like B3LYP [101, 168, 169, 203].

Studies on different systems, however, reveal that such findings are rather fortuitous [80, 86, 132, 172, 189]. Thus in general there is no evident correlation between the band gap and the HOMO-LUMO gap. The interpretation is based on a misconception.

Physically there are two distinct experimentally accessible gaps. The fundamental gap is the difference of the ionization potential (IP) and the electron affinity (EA). Experimentally it can be determined by measuring both quantities separately using photoemission spectroscopy (ionization potential) and inverse photoemission spectroscopy (electron affinity). Both processes are particle addition/removal processes, thus the particle number is not conserved. Such a process can not be described by a single ground state DFT calculation. However, it can be described by two ground state calculations, for example one for the  $N$ - and another for the  $(N + 1)$ -particle systems. Using the Janak's theorem, provided the functional is exact, the  $N$ -particle calculation gives the IP and the  $(N + 1)$ -particle calculation the EA. The proper fundamental gap from the KS-theory is

$$E_{\text{gap}} = \text{IP} - \text{EA} = E_{\text{gap}}^{\text{KS}} + \Delta_{\text{xc}}, \quad (2.5)$$

where the first term is the HOMO-LUMO gap from the KS-calculation and  $\Delta_{\text{xc}}$  is the derivative discontinuity [113]. The latter can be evaluated only from two distinct ground state calculations with different particle numbers. For example for range-separated (long-range corrected) functionals (section 2.3.2) the derivative discontinuity is equal to the difference of the HOMO for the  $(N + 1)$ -system and LUMO of the  $N$ -system [9]. For this reason a single ground state calculation will give an exact IP but the EA from LUMO eigenvalue will be in general wrong. Nevertheless, practical calculations show, that HOMO-LUMO gaps from nearly SIE-free functionals (e.g. range-separated functionals, sections 2.3.2 and 2.3.3) compare very well to experimental fundamental gaps [103].

The optical gap is defined as the difference of the lowest dipole-allowed excited state and the ground state. It can be measured by absorption spectroscopy. In this process the number of electrons is unchanged. It can be thought of as the creation of quasi-electron and quasi-hole. The interaction energy of these particles (binding energy of the exciton) leads to the optical gap which is smaller than the fundamental gap. The optical gap can be obtained in a proper way only using the time-dependent DFT [167], which is beyond the scope of this work. The local functionals, however, violating the linearity condition underestimate the IP and in the same way the EA (calculated from LUMO eigenvalue). Thus statistically they can yield a good estimate of the optical gaps as has been stated above.

Finally, we point out, that according to the Görling-Levy perturbation theory [57, 113], the physical foundation to the occupied KS-eigenvalues can be given. The KS-eigenvalue differences, have been shown to be the approximations to the excitation energies of the zeroth order in electron-electron interaction. Thus it can be expected, that the single-particle eigenvalues can be good estimates to the vertical IP's of the system, as measured by photoemission experiments. In fact, the references [33, 194] provide the evidence for this assumption.

## 2.3 SOLUTION: HYBRID FUNCTIONALS

### 2.3.1 Global hybrids

Hartree-Fock theory or the exact exchange functional in the KS theory (EXX) correct the self-interaction error, which stems from Hartree term in the exact manner. However, these functionals do not contain the correlation. On contrast, local approximate xc-functionals employ the exchange and correlation in a consistent way and take the advantage of mutual error cancellation of the exchange and correlation contributions. The combination of the two theories is usually referred to as a (global) hybrid Hartree-Fock-DFT theory or (global) hybrid xc-functionals. A fixed fraction  $\alpha$  of the exact exchange is mixed to the usual DFT functional

$$E_{xc}^{\text{hybrid}} = E_{xc}^{\text{DFT}} + \alpha(E_x^{\text{HF}} - E_x^{\text{DFT}}). \quad (2.6)$$

The mixing parameter  $\alpha$  can be determined empirically [16] or using the adiabatic connection theorem [147]. The HF exchange functional is explicitly orbital dependent. This leads to the problem, that the KS-potential Eq. 1.14 can not be obtained by explicit differentiation of the exact exchange with respect to the density. This problem can be solved with optimized effective potential (OEP) approach [105, 179, 191]. This procedure gives the KS-potential which corresponds to the orbitals, which minimize the total energy of the orbital-dependent functional. However, this approach is computationally very costly and is rarely used in practical calculations on extended systems.

An alternative way of employing the hybrid functionals, which is widely used in practical calculations is the generalized Kohn-Sham (GKS) approach. Instead of mapping the interacting system to the fully non-interacting, like in the pure Kohn-Sham approach, Seidl et. al. proposed to map the interacting system to an interacting one, which can be still described by a single Slater determinant (for example HF theory) [174]. This is the formal basis for the hybrid HF-DFT methods. Practically, the calculations are performed as in HF theory with additional local DFT potential.

The hybrid functionals correct many failures of local functionals and usually give very accurate results for a wide range of properties. However, they do not completely remove the self-interaction error.

### 2.3.2 Range-separated functionals

The asymptotic decay of the potential from the global hybrid functionals is  $-\alpha/r$ , i.e. proportional to the fraction of the exact exchange. To obtain the correct asymptotic decay of the potential the functionals have been proposed, which include the exact HF exchange for the long-range interaction and employ the local DFT exchange for the short-range interaction. This is achieved



by partitioning of the electron-electron interaction

$$\frac{1}{r} = \underbrace{\frac{\xi_{\text{sr}}^{\omega}(r)}{r}}_{\text{DFT}} + \underbrace{\frac{1 - \xi_{\text{sr}}^{\omega}(r)}{r}}_{\text{HF}} \quad (2.7)$$

by some smooth range-separation function  $\xi_{\text{sr}}^{\omega}(r)$ . It is usually chosen to be a complementary error function  $\xi_{\text{sr}}^{\omega}(r) = 1 - \text{erf}(\omega r)$  or exponential function  $\xi_{\text{sr}}^{\omega}(r) = e^{-\omega r}$ . In the latter case the screened Coulomb potential is of Yukawa type, thus we refer to it as Yukawa screening. The particular form of the range-separation function turns out to be of minor importance for the scheme. The long-range HF exchange then enforces the correct asymptotic decay of the xc-potential, while the theory still benefits from mutual error cancellation of the local exchange and correlation functionals at the short range. These functionals are called long-range corrected (LC) functionals.<sup>1</sup> The more general form of such functionals, also known as Coulomb-attenuating method (CAM) functionals reads

$$\frac{1}{r} = \underbrace{\frac{1 - \alpha - \beta(1 - \xi_{\text{sr}}^{\omega}(r))}{r}}_{\text{DFT}} + \underbrace{\frac{\alpha + \beta(1 - \xi_{\text{sr}}^{\omega}(r))}{r}}_{\text{HF}}, \quad (2.8)$$

where  $\alpha$  is the fraction of the global exact HF exchange and  $\alpha + \beta$  is the fraction of the HF exchange for the long-range [2, 200]. The  $\alpha$  and  $\beta$  parameters should satisfy the relations  $0 \leq \alpha + \beta \leq 1$ ,  $0 \leq \alpha \leq 1$ ,  $0 \leq \beta \leq 1$ . The xc-energy of such functional can be written as

$$E_{\text{xc}}^{\text{CAM}} = [1 - (\alpha + \beta)] E_{\text{x}}^{\text{DFT}} + \alpha E_{\text{x}}^{\text{HF}} + \beta (E_{\text{x, sr}}^{\omega, \text{DFT}} + E_{\text{x, lr}}^{\omega, \text{HF}}) + E_{\text{c}}^{\text{DFT}}, \quad (2.9)$$

where  $E_{\text{x, sr}}^{\omega, \text{DFT}}$  is the short-range exchange functional in LDA or GGA approximation, where the Coulomb interaction is screened by a range-separation function and  $E_{\text{x, lr}}^{\omega, \text{HF}}$  is the accompanying long-range HF exchange with descreened Coulomb interaction  $(1 - \xi_{\text{sr}}^{\omega}(r))/r$ . This way of mixing the DFT and HF naturally includes different range-separated and global hybrid functionals as limiting cases [2, 53, 54, 163, 198, 200]. The long-range corrected functional is for example obtained for  $\beta = 1, \alpha = 0$ , while a global hybrid is obtained for  $\alpha \neq 0, \beta = 0$ .

The parameters in the CAM functionals are essentially determined in an empirical way. As in the case of global hybrids, usual procedure in obtaining the parameters for such functionals is the fit of a standard benchmark set of molecules to experimental thermochemical data. The parameters  $\alpha, \beta, \omega$ , obtained in this way are then used for all systems.

The range-separated functionals have been shown to fulfill the linearity condition to a great extent [99]. They generally improve over the local and global hybrid functionals in description of response properties [70, 90, 176], photoemission spectra [99, 161], bond length alternation in conjugated polymers [100] to name some examples.

<sup>1</sup>Note also, that terms range-separated and long-range corrected are often used as synonyms.

### 2.3.3 Adiabatic connection functional

The range-separated functionals and the long-range corrected functionals as their special case have been so far introduced on rather empirical grounds. However, they can be directly derived using adiabatic connection theorem. Assume there exists a one-parameter smooth representation of an electron-electron interaction  $\hat{V} \rightarrow \hat{V}^\omega$ ,  $a \leq \omega \leq b$  such that  $\hat{V}^a = 0$  and  $\hat{V}^b = \hat{V}$ . The universal functional for the system with such interaction reads

$$F^\omega[\rho] = \min_{\Psi^\omega \rightarrow \rho} \langle \Psi^\omega | \hat{T} + \hat{V}^\omega | \Psi^\omega \rangle. \quad (2.10)$$

Then by construction  $F^a = T^0$  and  $F^b = F$  and the xc-functional can be written as

$$E_{\text{xc}}[\rho] = \int_a^b \frac{dF^{\omega'}}{d\omega'} d\omega' - E_H[\rho]. \quad (2.11)$$

This is known as adiabatic connection theorem [60, 107, 201]. It can be used to derive a long-range corrected functional. Assume the following parametrization of the electron-electron interaction  $\hat{V}^\omega = \frac{1}{2} \sum_{i \neq j}^N (1 - \exp(-\omega |\mathbf{r}_i - \mathbf{r}_j|)) |\mathbf{r}_i - \mathbf{r}_j|^{-1}$ . The conditions above are fulfilled for  $a = 0$ ,  $b \rightarrow \infty$  and the exchange-correlation energy yields

$$E_{\text{xc}}[\rho] = \int_0^\infty \frac{dF^{\omega'}}{d\omega'} d\omega' - E_H[\rho] = \int_0^\infty \langle \Psi^{\omega'} | \frac{d\hat{V}^{\omega'}}{d\omega'} | \Psi^{\omega'} \rangle d\omega' - E_H[\rho]. \quad (2.12)$$

This kind of adiabatic connection along with the practical approximations, which we present below has been suggested by Baer and Neuhauser [10]. To approximately evaluate the adiabatic connection integral they proceed as follows. Given the value of parameter  $\omega$ , the wave function in Eq. 2.12 for the case  $\omega' < \omega$  is assumed to be the Slater determinant  $\Phi$  and for  $\omega' > \omega$  the full interacting wave function  $\Psi$

$$\begin{aligned} \int_0^\infty \langle \Psi^{\omega'} | \frac{d\hat{V}^{\omega'}}{d\omega'} | \Psi^{\omega'} \rangle d\omega' &\approx \int_0^\omega \langle \Phi | \frac{d\hat{V}^{\omega'}}{d\omega'} | \Phi \rangle d\omega' + \int_\omega^\infty \langle \Psi | \frac{d\hat{V}^{\omega'}}{d\omega'} | \Psi \rangle d\omega' \\ &= \langle \Phi | \hat{V}^\omega | \Phi \rangle + \langle \Psi | \hat{V} - \hat{V}^\omega | \Psi \rangle \\ &= \underbrace{\frac{1}{2} \int \frac{1 - e^{-\omega |\mathbf{r} - \mathbf{r}'|}}{|\mathbf{r} - \mathbf{r}'|} \rho(\mathbf{r}) \rho(\mathbf{r}') d\mathbf{r} d\mathbf{r}'}_{E_H^\omega[\rho]} - \underbrace{\frac{1}{2} \int \frac{1 - e^{-\omega |\mathbf{r} - \mathbf{r}'|}}{|\mathbf{r} - \mathbf{r}'|} \rho^2(\mathbf{r}, \mathbf{r}') d\mathbf{r} d\mathbf{r}'}_{E_{\text{x,lr}}^{\omega, \text{HF}}[\rho]} + \langle \Psi | \hat{V} - \hat{V}^\omega | \Psi \rangle, \end{aligned} \quad (2.13)$$

where  $\rho(\mathbf{r}, \mathbf{r}')$  is the first-order reduced density matrix. The integral is evaluated as

$$\int_a^b \langle \Psi | \frac{d\hat{V}^{\omega'}}{d\omega'} | \Psi \rangle d\omega' = \langle \Psi | \int_a^b \frac{d\hat{V}^{\omega'}}{d\omega'} d\omega' | \Psi \rangle = \langle \Psi | \frac{1}{2} \sum_{i \neq j}^N \left( -\frac{e^{-\omega |\mathbf{r}_i - \mathbf{r}_j|}}{|\mathbf{r}_i - \mathbf{r}_j|} \right) \Big|_a^b | \Psi \rangle. \quad (2.14)$$

The Slater determinant  $\Phi$ , applied to the descreened interaction operator gives the long-range Hartree term  $E_H^\omega$  and the long-range Fock term  $E_{\text{x,lr}}^{\omega, \text{HF}}$ . With this the Baer-Neuhauser (BN)

exchange-correlation functional can be written as

$$\begin{aligned}
 E_{\text{xc}}^{\text{BN}}[\rho] &= E_{\text{x,lr}}^{\omega,\text{HF}}[\rho] + \langle \Psi | \hat{V} - \hat{V}^\omega | \Psi \rangle - E_H[\rho] + E_H^\omega[\rho] \\
 &= E_{\text{x,lr}}^{\omega,\text{HF}}[\rho] + \underbrace{\left[ \langle \Psi | \hat{V} - \hat{V}^\omega | \Psi \rangle - \frac{1}{2} \int \frac{e^{-\omega|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} \rho(\mathbf{r})\rho(\mathbf{r}') d^3\mathbf{r}d^3\mathbf{r}' \right]}_{E_{\text{xc}}^\omega}. \tag{2.15}
 \end{aligned}$$

The long-range HF exchange  $E_{\text{x,lr}}^{\omega,\text{HF}}$  is non-local and accompanies the Hartree term in the total energy expression Eq. 1.11 at large separations, correcting the potential asymptotically. The screened part is further approximated as a local DFT functional

$$E_{\text{xc}}^\omega[\rho] \approx \int \epsilon_{\text{xc}}^\omega[\rho](\mathbf{r})\rho(\mathbf{r})d\mathbf{r}, \tag{2.16}$$

where the xc-energy per particle  $\epsilon_{\text{xc}}^\omega$  can be evaluated within the LDA or GGA approximation derived for the screened Coulomb interaction [53, 67, 81, 110, 171]. We outline a possible way of doing this in section 3.1. To sum up, the full energy functional reads

$$E[\rho] = T^0[\rho] + \int v^{\text{ext}}\rho + E_H[\rho] + E_{\text{x,lr}}^{\omega,\text{HF}}[\rho] + E_{\text{xc}}^\omega[\rho]. \tag{2.17}$$

Note, that for  $\omega = 0$  the scheme reduces to the usual Kohn-Sham approach.

We consider now the difference of the functional  $E_{\text{xc}}^{\text{BN}}$  and the exact Kohn-Sham xc-functional

$$\Delta E_{\text{xc}} = [T + \langle \Psi | \hat{V} | \Psi \rangle - T_s - E_H] - [\langle \Phi | \hat{V}^\omega | \Phi \rangle - \langle \Psi | \hat{V}^\omega | \Psi \rangle + \langle \Psi | \hat{V} | \Psi \rangle - E_H[\rho]] \tag{2.18}$$

$$= [T + \langle \Psi | \hat{V}^\omega | \Psi \rangle] - [T_s + \langle \Phi | \hat{V}^\omega | \Phi \rangle]. \tag{2.19}$$

For  $\omega \rightarrow 0$ ,  $\Delta E_{\text{xc}} = T - T^0 \geq 0$ . This can be seen from variational principle. For a given density  $\rho$  the Slater determinant  $\Phi$ , which yields this density minimizes the functional  $T^0 = \langle \Phi | \hat{T} | \Phi \rangle$ . The interacting wave function, which yields the same density minimizes the functional  $E = \langle \Psi | \hat{T} + \hat{V} | \Psi \rangle$ , thus  $T^0 = \langle \Phi | \hat{T} | \Phi \rangle \leq \langle \Psi | \hat{T} | \Psi \rangle = T$ . For the case  $\omega \rightarrow \infty$ ,  $\Delta E_{\text{xc}} = T + \langle \Psi | \hat{V} | \Psi \rangle - (T_s + \langle \Phi | \hat{V} | \Phi \rangle) = E - E^{\text{HF}} = E_c < 0$ . Note, that the correlation energy  $E_c$  is defined as the difference of the exact energy and the Hartree-Fock energy. Assuming the continuity one concludes that there is an  $\omega$  with  $0 \leq \omega \leq \infty$  such that  $\Delta E_{\text{xc}} = 0$  and the BN-functional is exact [10].

## 2.4 MATRIX EQUATIONS

In this section we derive the matrix equations, which can be directly used for the implementation of a long-range corrected DFT and approximations to it. The energy expression Eq. 2.17 defines the long-range corrected functional. The density is represented by a Slater-determinant with spin-orbitals  $\psi_i$ , such that  $\rho = \sum_i n_i |\psi_i|^2$ , where  $n_i$  are the occupations of the spin-orbitals. Thus

the energy can be rewritten in terms of the spin-orbitals

$$\begin{aligned}
 E &= \sum_i^N n_i \int \psi_i(\mathbf{x}) \left( -\frac{1}{2} \nabla^2 \right) \psi_i(\mathbf{x}) d\mathbf{x} + \sum_i^N n_i \int |\psi(\mathbf{x})|^2 v^{\text{ext}}(\mathbf{r}) d\mathbf{x} + E_{\text{xc}}^\omega[\rho] \\
 &+ \sum_{ij}^N \frac{n_i n_j}{2} \int \frac{\psi_i(\mathbf{x}) \psi_i(\mathbf{x}) \psi_j(\mathbf{x}') \psi_j(\mathbf{x}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{x} d\mathbf{x}' \\
 &- \sum_{ij}^N \frac{n_i n_j}{2} \int \psi_i(\mathbf{x}) \psi_j(\mathbf{x}) \frac{1 - e^{-\omega|\mathbf{r} - \mathbf{r}'|}}{|\mathbf{r} - \mathbf{r}'|} \psi_i(\mathbf{x}') \psi_j(\mathbf{x}') d\mathbf{x} d\mathbf{x}' + E_{NN}, \tag{2.20}
 \end{aligned}$$

where  $E_{NN}$  is the nuclei-nuclei repulsion energy, which is constant for a given geometry. As in the usual KS-approach the variation is performed with respect to the orbitals. For the numerical treatment of the problem on a computer the discretization has to be done. The direct way of discretization in the real space on a grid (or set of grids) leads to a class of grid-based numerical methods (for example [102]). Another way is usually referred to as Roothaan-Hall method, where the discretization is performed with respect to an auxiliary basis set [62, 164–166]. Although this method has been initially developed for the Hartree-Fock theory, it is widely used in the implementations of the Kohn-Sham method and the hybrid Hartree-Fock-DFT approaches, since the mathematical procedure in all cases is essentially the same. Each spin orbital is represented by a linear combination of some appropriate finite set of functions  $\{\phi_\mu\}$ , which constitute a basis

$$\psi_i(\mathbf{x}) = \sum_\mu c_{\mu i} \phi_\mu(\mathbf{r}) \sigma_i(\omega), \tag{2.21}$$

where  $\sigma_i(\omega)$  is the spin part. Inserting this definition in the energy expression Eq. 2.20 and integrating out the spin degrees of freedom  $\int \sigma_i(\omega) \sigma_j(\omega) d\omega = \delta_{\sigma_i \sigma_j}$  we obtain

$$\begin{aligned}
 E &= \sum_{\mu\nu} \sum_i^N n_i c_{\mu i} c_{\nu i} \int \phi_\mu(\mathbf{r}) \left( -\frac{1}{2} \nabla^2 + v^{\text{ext}}(\mathbf{r}) \right) \phi_\nu(\mathbf{r}) d\mathbf{r} + E_{\text{xc}}^\omega[\rho] + E_{NN} \\
 &+ \frac{1}{2} \sum_{\mu\nu\kappa\lambda} (\mu\nu|\kappa\lambda) \sum_i^N n_i c_{\mu i} c_{\nu i} \sum_j^N n_j c_{\kappa j} c_{\lambda j} - \frac{1}{2} \sum_{\mu\nu\kappa\lambda} (\mu\nu|\kappa\lambda)^{\text{lr}} \sum_{ij}^N n_i n_j \delta_{\sigma_i \sigma_j} c_{\mu i} c_{\kappa i} c_{\nu j} c_{\lambda j} \\
 &= \sum_{\mu\nu} P_{\mu\nu}^{\text{total}} h_{\mu\nu} + E_{\text{xc}}^\omega[\rho] + \frac{1}{2} \sum_{\mu\nu\kappa\lambda} P_{\mu\nu}^{\text{total}} P_{\kappa\lambda}^{\text{total}} (\mu\nu|\kappa\lambda) \\
 &- \frac{1}{2} \sum_{\mu\nu\kappa\lambda} \left( P_{\mu\kappa}^\alpha P_{\nu\lambda}^\alpha + P_{\mu\kappa}^\beta P_{\nu\lambda}^\beta \right) (\mu\nu|\kappa\lambda)^{\text{lr}} + E_{NN}. \tag{2.22}
 \end{aligned}$$

Here following one- and two-electron integrals (matrix elements) appear

$$h_{\mu\nu} = \int \phi_\mu(\mathbf{r}) \left( -\frac{1}{2} \nabla^2 + v^{\text{ext}}(\mathbf{r}) \right) \phi_\nu(\mathbf{r}) d\mathbf{r} \tag{2.23}$$

$$(\mu\nu|\kappa\lambda) = \int \frac{\phi_\mu(\mathbf{r}) \phi_\nu(\mathbf{r}) \phi_\kappa(\mathbf{r}') \phi_\lambda(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' \tag{2.24}$$

$$(\mu\nu|\kappa\lambda)^{\text{lr}} = \int \phi_\mu(\mathbf{r}) \phi_\nu(\mathbf{r}) \frac{1 - e^{-\omega|\mathbf{r} - \mathbf{r}'|}}{|\mathbf{r} - \mathbf{r}'|} \phi_\kappa(\mathbf{r}') \phi_\lambda(\mathbf{r}') d\mathbf{r} d\mathbf{r}'. \tag{2.25}$$

We assume, that the orbitals with  $1 \leq i \leq N_\alpha$  are spin-up or  $\alpha$ -electrons and the rest  $N_\alpha < i \leq N$  are the spin-down or  $\beta$ -electrons, where  $N_\alpha$  is the number of alpha electrons. The spin density matrix for  $\alpha$ -electrons is defined as  $P_{\mu\nu}^\alpha = \sum_i^{N_\alpha} n_i c_{\mu i} c_{\nu i}$ . For  $\beta$ -electrons the expression reads  $P_{\mu\nu}^\beta = \sum_{i=N_\alpha+1}^N n_i c_{\mu i} c_{\nu i}$ . The total density matrix is the sum of spin-up and spin-down density matrices  $P^{\text{total}} = P^\alpha + P^\beta$ . For the closed shell case  $N_\alpha = N_\beta = N/2$  and  $P = P^{\text{total}} = 2P^\alpha$ . In this case the equations read

$$E = \sum_{\mu\nu} P_{\mu\nu} h_{\mu\nu} + \frac{1}{2} \sum_{\mu\nu} \sum_{\alpha\beta} P_{\mu\nu} P_{\alpha\beta} (\mu\nu|\alpha\beta) - \frac{1}{4} \sum_{\mu\nu} \sum_{\alpha\beta} P_{\mu\nu} P_{\alpha\beta} (\mu\alpha|\beta\nu)^{\text{lr}} + E_{\text{xc}}^\omega[\rho] + E_{\text{NN}}. \quad (2.26)$$

The variation of the energy with respect to the molecular orbital coefficients  $c_{\mu i}$  subject to the condition  $\int \psi_i(\mathbf{r}) \psi_j(\mathbf{r}) d\mathbf{r} = \delta_{ij}$  yields the matrix equations

$$\sum_{\mu\nu} H_{\mu\nu} c_{\nu i} = \epsilon_i \sum_{\mu\nu} S_{\mu\nu} c_{\nu i}, \quad (2.27)$$

where  $\epsilon_i$  are the eigenvalues,  $S_{\mu\nu} = \int \phi_\mu(\mathbf{r}) \phi_\nu(\mathbf{r}) d\mathbf{r}$  is the overlap of the basis functions and the Hamiltonian reads

$$H_{\mu\nu} = h_{\mu\nu} + v_{\mu\nu}^{\text{xc}}[\rho] + \sum_{\alpha\beta} P_{\alpha\beta} (\mu\nu|\alpha\beta) - \frac{1}{2} \sum_{\alpha\beta} P_{\alpha\beta} (\mu\alpha|\beta\nu)^{\text{lr}}, \quad (2.28)$$

where  $v_{\mu\nu}^{\text{xc}}[\rho] = \int \phi_\mu(\mathbf{r}) v_{\text{xc}}^\omega[\rho](\mathbf{r}) \phi_\nu(\mathbf{r}) d\mathbf{r}$  is the representation of the xc-potential in this basis. The xc-functional, which is non-linear in density is usually evaluated numerically on a set of atom-centered grids. Other matrix elements for the given geometry can be precomputed.



**Part I**

**Implementation**





## LC-DFTB: THE APPROXIMATE LC-DFT

The delocalization problem, which is typical for the local DFT arises also in the density functional-based tight-binding method (DFTB), which we briefly introduced in section 1.3. The proposal to extend the DFTB method to the case of hybrid exchange-correlation functionals with special emphasis on the long-range corrected functionals (section 2.3.2) has been done by Niehaus and Della Sala [137]. Their paper covers the basic formalism of the new scheme. In this chapter we present the practical implementation of the DFTB method, based on a long-range corrected xc-functional. We denote this method as the long-range corrected density functional-based tight-binding method (LC-DFTB). First we briefly motivate the choice of the particular exchange-correlation functional and provide its definition and implementation details in section 3.1. The approximations and the resulting expressions for the total energy and Hamiltonian for single point calculations are presented in sections 3.2 to 3.8. In section 3.9 the efficiency of the run-time algorithms is addressed. Finally, we cover the evaluation of energy gradients (forces) for the new scheme in section 3.10.

### 3.1 CHOICE OF THE XC-FUNCTIONAL

The main conclusion from the work of ref. [137] is that DFTB approximations can formally be applied to a general range-separated hybrid DFT. It is undoubtedly tempting to include the variety of functionals into the DFTB method. However, one should keep in mind, that the main advantage of the DFTB is its computational efficiency. It relies on the extensive use of precomputed and optimized parameter sets, which simplify significantly the computational procedure, as compared to the full *ab initio* approach. On the one hand it is desirable to keep the number of parameter

sets manageable and on the other their generation and optimization can be very time consuming<sup>1</sup>. Thus we keep the method as simple as possible and use the long-range corrected scheme, which has been derived using adiabatic connection in section 2.3.3. In this method the range-separation parameter  $\omega$  is a free parameter and for each value of  $\omega$  a separate parameter set for the DFTB has to be generated. The long-range corrected hybrid exchange-correlation functional takes the form

$$E_{xc} = E_{xc}^{\omega} + E_{x,lr}^{\omega,HF}, \quad (3.1)$$

where  $E_{x,lr}^{\omega,HF}$  is the long-range Hartree-Fock exchange term (compare Eq. 2.13) and the choice of the screened local DFT part  $E_{xc}^{\omega}$  of the functional is discussed in following.

### 3.1.1 Correlation functional

As has been already stated in section 2.3.3 the long-range corrected functional can be derived from the adiabatic connection theorem by imposing additional approximations. The last ingredient for the practical implementation of the theory, the screened local DFT exchange-correlation functional, has to be further approximated. Usual path is to use some combination of LDA or GGA exchange functional, derived for the case of screened Yukawa (or error-function-based) interaction and standard LDA or GGA correlation functional. Although there are quantitative differences in the benchmark tests for different choices of local DFT exchange-correlation part, the qualitative improvement over the LDA/GGA and global hybrid functionals is always observed. Since the main aim of this work is a proof of concept, we consider the particular choice of the DFT functional to be of minor importance.

We thus decide to use a variant of the BNL functional, based on the aforementioned work by Neuhauser and Baer (section 2.3.3) and further developed by Livshits and Baer [116]. For efficiency reasons in that work they switched to the error-function-based range separation, since the functional was implemented in the code, based on the Gaussian-type orbitals (GTO). On contrast, we still use the Yukawa-type range separation as in the original work of Neuhauser and Baer.<sup>2</sup> For  $E_{xc}^{\omega}$  Livshits and Baer suggested to combine a standard GGA correlation functional with a short-range LDA exchange, according to

$$E_{xc}^{\omega} = E_c^{GGA} + \eta E_{x,sr}^{\omega,LDA}, \quad (3.2)$$

where the parameter  $0 \leq \eta \leq 1$  is determined empirically. Both parameters  $\omega$  and  $\eta$  are obtained by minimization of error for some molecule benchmark set with respect to thermochemical data. In the same way more general CAM functionals are fine-tuned, as has been already outlined. We do

<sup>1</sup> In last time in the DFTB community the automatic parameter optimization tools get developed [21]. However, at present time they are not generally available and their use requires some routine.

<sup>2</sup>This choice is mainly due to the lack of a proper integration routine for the efficient evaluation of electron repulsion integrals with error-function-based screened interaction over Slater-type basis functions, see also section 4.2.

not attempt any fine-tuning of this functional for the method, presented in this thesis. We decide to use the PBE [146] correlation functional, which is used in traditional DFTB parametrization and choose  $\eta = 1$  in Eq. 3.2.

In general, different local xc-functionals can be obtained for example from the LIBXC library [124]. However, since the library was designed for GTO-based packages, it still does not contain the Yukawa-screened exchange functionals. For this reason we modify the existent routine for the evaluation of the LDA exchange as described in the following section. The standard correlation functional (PBE) is obtained from the LIBXC library.

### 3.1.2 Short-range exchange functional

We decide to use the simplest short-range exchange functional. We rederive the formula for the exchange energy of the uniform electron gas with Yukawa interaction. The exact first-order spinless density matrix of a uniform electron gas reads

$$\rho(\mathbf{r}_1, \mathbf{r}_2) = 3\rho(\mathbf{r}) \left( \frac{\sin(t) - t \cos(t)}{t^3} \right), \quad (3.3)$$

where  $k_F = (3\pi^2\rho)^{1/3}$ ,  $\mathbf{r} = \frac{1}{2}(\mathbf{r}_1 + \mathbf{r}_2)$ ,  $t = sk_F$ ,  $s = |\mathbf{r}_1 - \mathbf{r}_2|$ . The exchange energy is obtained from

$$\begin{aligned} E_{\text{x,SR}}^{\omega,\text{LDA}}[\rho] &= -\frac{1}{4} \int \rho^2(\mathbf{r}, s) \frac{\exp(-\omega s)}{s} d\mathbf{r} ds = -\pi \int \rho^2(\mathbf{r}, s) s \exp(-\omega s) ds d\mathbf{r} \\ &= -\frac{9\pi}{4} \int \frac{\rho^2(\mathbf{r})}{k_F^2(\mathbf{r})} \underbrace{\left[ \int_0^\infty \frac{4(\sin t - t \cos t)^2}{t^5} \exp(-\alpha t) dt \right]}_{P(\alpha)} d\mathbf{r} \\ &= -\frac{3}{4} \left( \frac{3}{\pi} \right)^{1/3} \int \rho^{4/3}(\mathbf{r}) P(\alpha) d\mathbf{r}, \end{aligned} \quad (3.4)$$

where we define the quantity  $\alpha = \omega/k_F(\mathbf{r})$ . The factor  $P(\alpha)$  reads [2, 162]

$$P(\alpha) = 1 - \frac{4}{3}\alpha \left[ \text{atan} \frac{2}{\alpha} + \frac{\alpha}{8} - \frac{\alpha}{8} \left( \frac{\alpha^2}{4} + 3 \right) \ln \left( 1 + \frac{4}{\alpha^2} \right) \right]. \quad (3.5)$$

We note, that for the limit  $\omega \rightarrow 0$ ,  $\alpha \rightarrow 0$  and  $P(\alpha) = 1$ , which gives the result for the standard Slater exchange. The extension to the short-range version of the given LDA functional is thus performed by multiplication of the exchange energy per particle by the screening factor  $P(\alpha)$ .

The direct numerical implementation of the factor  $P(\alpha)$  according to the formula Eq. 3.5 shows instabilities for large  $\alpha$  (small densities). In Fig. 3.1 the absolute value of the screening factor  $P(\alpha)$  is plotted for the wide range of values  $\alpha$ . The circles denote the direct evaluation from the Eq. 3.5, and the sign of the screening factor is encoded by blue (positive) and red (negative) color. We observe large fluctuations of the screening factor for  $\alpha > 200$ . This can be corrected, using Taylor expansion of  $P(1/\alpha)$  around zero, from which we obtain for large  $\alpha$

$$P(\alpha) \approx \frac{4}{9} \frac{1}{\alpha^2} - \frac{8}{15} \frac{1}{\alpha^4} + \mathcal{O}\left(\frac{1}{\alpha^6}\right). \quad (3.6)$$

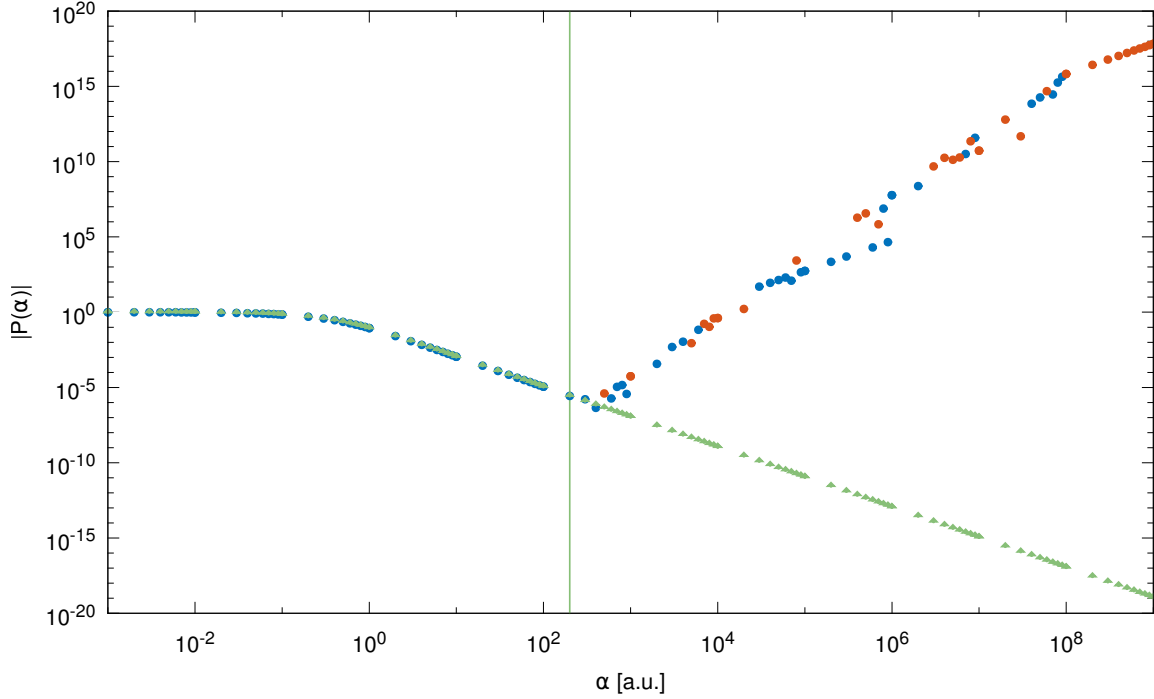


Figure 3.1: The numerical instability in the Yukawa screening factor  $P(\alpha)$ . The absolute value of the function  $P(\alpha)$  evaluated directly using Eq. 3.5 is plotted with circles, where the sign is coded by blue (positive) and red (negative) color. The numerical instability occurs for  $\alpha > 200$ . The selective algorithm (see main text) uses the Taylor expansion Eq. 3.6 for values  $\alpha > 200$ , which solves the problem (green triangles).

Within a selective algorithm, this formula is then applied for values  $\alpha > 200$  and for  $\alpha \leq 200$  the direct formula Eq. 3.5 is used. The triangles in Fig. 3.1 show the result for selective algorithm, which clearly corrects the numerical instability. At this point we mention, that the same problem (and solution) have been reported in supplementary material of ref. [177].

We now perform the functional derivative of the exchange energy Eq. 3.4 and obtain the expression for the short-range exchange potential with Yukawa-type range separation

$$\frac{\delta E_{x,\text{sr}}^{\omega,\text{LDA}}}{\delta \rho} = - \left( \frac{3}{\pi} \right)^{1/3} \rho^{1/3}(\mathbf{r}) P(\alpha) + \frac{\omega}{4\pi} \frac{dP}{d\alpha}. \quad (3.7)$$

We call the functional, defined in this section BNL-Y functional to distinguish it from the original BNL functional.

### 3.2 SECOND-ORDER ENERGY EXPANSION

In following we will derive the total energy and Hamiltonian for the LC-DFTB method by applying approximations, known from the standard DFTB method to the total energy of the first-principles LC-DFT. This essentially parallels the derivation in ref. [137].

Given a local DFT exchange-correlation functional  $E_{\text{xc}}^\omega$  (for example one, defined in section 3.1), the energy of a LC-DFT theory in the given finite basis reads (Eq. 2.26)

$$E = \sum_{\mu\nu} P_{\mu\nu} h_{\mu\nu} + \frac{1}{2} \sum_{\mu\nu} \sum_{\alpha\beta} P_{\mu\nu} P_{\alpha\beta} (\mu\nu|\alpha\beta) - \frac{1}{4} \sum_{\mu\nu} \sum_{\alpha\beta} P_{\mu\nu} P_{\alpha\beta} (\mu\alpha|\beta\nu)^{\text{lr}} + E_{\text{xc}}^\omega[\rho] + E_{\text{NN}}. \quad (3.8)$$

The first approximation is the linearization of the local part of the exchange-correlation potential, which is done by expansion of the exchange-correlation energy functional around some reference density matrix  $\rho_0$  up to the second order in the difference density matrix  $\delta\rho = \rho - \rho_0$

$$\begin{aligned} E_{\text{xc}}^\omega[\rho_0 + \delta\rho] &= E_{\text{xc}}^\omega[\rho_0] + \int \left. \frac{\delta E_{\text{xc}}^\omega}{\delta\rho(\mathbf{r})} \right|_{\rho_0} \delta\rho(\mathbf{r}) d\mathbf{r} + \frac{1}{2} \int \left. \frac{\delta^2 E_{\text{xc}}^\omega}{\delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')} \right|_{\rho_0} \delta\rho(\mathbf{r})\delta\rho(\mathbf{r}') d\mathbf{r}d\mathbf{r}' + \mathcal{O}(\delta\rho^3) \\ &= E_{\text{xc}}^\omega[\rho_0] + \int v_{\text{xc}}^\omega[\rho](\mathbf{r})\delta\rho(\mathbf{r}) d\mathbf{r} + \frac{1}{2} \int f_{\text{xc}}^\omega[\rho](\mathbf{r}, \mathbf{r}')\delta\rho(\mathbf{r})\delta\rho(\mathbf{r}') d\mathbf{r}d\mathbf{r}' + \mathcal{O}(\delta\rho^3). \end{aligned} \quad (3.9)$$

This expansion is expected to be an appropriate approximation to the exchange-correlation energy for the density matrices  $\rho$  which are close to the reference density matrix. We note, that the non-local long-range Hartree-Fock exchange potential and the Hartree potential are already linear in density matrix. For convenience and more clarity we rewrite the expansion Eq. 3.9 in terms of the finite auxiliary basis

$$\begin{aligned} E_{\text{xc}}^\omega[\rho] &= E_{\text{xc}}^\omega[\rho_0] + \sum_{\mu\nu} \Delta P_{\mu\nu} \int v_{\text{xc}}^\omega[\rho_0](\mathbf{r})\phi_\mu(\mathbf{r})\phi_\nu(\mathbf{r}) d\mathbf{r} \\ &\quad + \frac{1}{2} \sum_{\mu\nu\alpha\beta} \Delta P_{\mu\nu} \Delta P_{\alpha\beta} \int f_{\text{xc}}^\omega[\rho_0](\mathbf{r}, \mathbf{r}')\phi_\mu(\mathbf{r})\phi_\nu(\mathbf{r})\phi_\alpha(\mathbf{r}')\phi_\beta(\mathbf{r}') d\mathbf{r}d\mathbf{r}' + \mathcal{O}(\Delta P^3) \\ &= E_{\text{xc}}^\omega[\rho_0] + \sum_{\mu\nu} \Delta P_{\mu\nu} v_{\mu\nu}^{\text{xc}} + \frac{1}{2} \sum_{\mu\nu\alpha\beta} \Delta P_{\mu\nu} \Delta P_{\alpha\beta} f_{\mu\nu\alpha\beta}^{\text{xc}} + \mathcal{O}(\Delta P^3), \end{aligned} \quad (3.10)$$

where the difference density matrix in the finite basis representation is denoted by  $\Delta P_{\mu\nu} = P_{\mu\nu} - P_{\mu\nu}^0$  and the first and second functional derivatives of the xc-functional are denoted by  $v_{\mu\nu}^{\text{xc}}$  and  $f_{\mu\nu\alpha\beta}^{\text{xc}}$ . The goal is now to factor out the Hamiltonian  $H^0$ , which depends only on the reference density matrix. We follow the procedure in [42] for standard DFTB and rearrange the terms in the total energy expression Eq. 3.8

$$\begin{aligned} E &= \sum_{\mu\nu} P_{\mu\nu} \left[ h_{\mu\nu} + v_{\mu\nu}^{\text{xc}}[P^0] + \sum_{\alpha\beta} P_{\alpha\beta}^0 (\mu\nu|\alpha\beta) - \frac{1}{2} \sum_{\alpha\beta} P_{\alpha\beta}^0 (\mu\alpha|\beta\nu)^{\text{lr}} \right] - \sum_{\mu\nu} P_{\mu\nu} v_{\mu\nu}^{\text{xc}}[P^0] \\ &\quad + \sum_{\mu\nu\alpha\beta} P_{\mu\nu} \Delta P_{\alpha\beta} (\mu\nu|\alpha\beta) - \frac{1}{2} \sum_{\mu\nu\alpha\beta} P_{\mu\nu} \Delta P_{\alpha\beta} (\mu\alpha|\beta\nu)^{\text{lr}} - \frac{1}{2} \sum_{\mu\nu\alpha\beta} P_{\mu\nu} P_{\alpha\beta} (\mu\nu|\alpha\beta) \\ &\quad + \frac{1}{4} \sum_{\mu\nu\alpha\beta} P_{\mu\nu} P_{\alpha\beta} (\mu\nu|\alpha\beta)^{\text{lr}} + E_{\text{xc}}^\omega[\rho] + E_{\text{NN}}. \end{aligned} \quad (3.11)$$

Here we inserted (added and subtracted) the exchange-correlation potential  $v_{\text{xc}}^\omega[\rho_0]$ , evaluated at the reference density matrix  $P_{\mu\nu}^0$ . The zeroth-order Hamiltonian is the expression in the square brackets in the first line of Eq. 3.11

$$H_{\nu\mu}^0 = h_{\mu\nu} + v_{\mu\nu}^{\text{xc}}[\rho^0] + \sum_{\alpha\beta} P_{\alpha\beta}^0 (\mu\nu|\alpha\beta) - \frac{1}{2} \sum_{\alpha\beta} P_{\alpha\beta}^0 (\mu\alpha|\beta\nu)^{\text{lr}}. \quad (3.12)$$

We proceed further by inserting the expansion of the xc-functional Eq. 3.10 and after some term rearrangement we arrive at the energy expression

$$\begin{aligned}
 E = & \sum_{\mu\nu} P_{\mu\nu} H_{\mu\nu}^0 + \frac{1}{2} \sum_{\mu\nu\alpha\beta} \Delta P_{\mu\nu} \Delta P_{\alpha\beta} \left[ (\mu\nu|\alpha\beta) + f_{\mu\nu\alpha\beta}^{\text{xc}} \right] - \frac{1}{4} \sum_{\mu\nu\alpha\beta} \Delta P_{\mu\nu} \Delta P_{\alpha\beta} (\mu\alpha|\beta\nu)^{\text{lr}} \\
 & + \underbrace{E_{\text{xc,sr}}^{\omega}[\rho_0] - \sum_{\mu\nu} P_{\mu\nu}^0 v_{\mu\nu}^{\text{xc}}[P^0] - \frac{1}{2} \sum_{\mu\nu\alpha\beta} P_{\mu\nu}^0 P_{\alpha\beta}^0 (\mu\nu|\alpha\beta) + \frac{1}{4} \sum_{\mu\nu\alpha\beta} P_{\mu\nu}^0 P_{\alpha\beta}^0 (\mu\alpha|\beta\nu)^{\text{lr}} + E_{\text{NN}}}_{E_{\text{rep}}}. \quad (3.13)
 \end{aligned}$$

The second line in the Eq. 3.13 depends only on the reference density matrix and is referred to as the DFTB repulsive energy  $E_{\text{rep}}$ . In standard DFTB this contribution to the total energy is approximated by a sum of fast decaying pair potentials, which depend only on the atom pair and distance between them

$$E_{\text{rep}} \approx \sum_{AB} V_{AB}(|\mathbf{R}_{AB}|). \quad (3.14)$$

In the present method one can use the same assumption as has been already discussed in ref. [137]. For the electronic structure calculations at fixed geometry this term is of no importance. This covers for example the eigenvalue spectrum, the single-particle orbitals and response properties without geometry relaxation. The parametrization of the repulsive potential for the new method for the species carbon and hydrogen is presented in chapter 5.

### 3.3 TWO-CENTER APPROXIMATION AND THE ZERO-ORDER LC-DFTB

The energy expression in Eq. 3.13 has been derived by the expansion of the total energy of the LC-DFT around a reference density matrix  $\rho_0$ . If we assume, that it is a sufficiently good approximation to the ground state density matrix and neglect the terms, which depend on  $\Delta P_{\mu\nu}$  we obtain the energy of the LC-DFTB scheme in zeroth order, in analogy to the standard DFTB

$$E = \sum_{\mu\nu} P_{\mu\nu} H_{\mu\nu}^0 + E_{\text{rep}}. \quad (3.15)$$

The variation with respect to the molecular orbital coefficients  $c_{\mu i}$  results in the generalized eigenvalue problem

$$\sum_{\mu} H_{\nu\mu}^0 c_{\mu i} = \epsilon_i \sum_{\mu} S_{\nu\mu} c_{\mu i} \quad (3.16)$$

with zeroth-order Hamiltonian, given by equation 3.12. This problem is solved by one diagonalization. So obtained coefficients  $c_{\mu i}$  constitute the density matrix  $P_{\mu\nu}$  and the energy Eq. 3.15 can be evaluated. So far this calculation would correspond to a first diagonalization in the *ab initio* procedure for some initial density guess, which is the reference density matrix  $\rho_0$ . The standard DFTB method uses the superposition of atomic densities as an initial guess  $\rho_0 = \sum_A^{\text{atoms}} \rho_A$ . The

reference density matrix in the atomic basis yields then  $P_{\mu\nu} = \delta_{\mu\nu}n_{\mu}$ , where  $n_{\mu}$  is the occupation of the atomic orbital  $\phi_{\mu}$ . This initial density guess is often used as default in *ab initio* quantum chemistry packages. For the LC-DFTB method we use the same reference density.

The zeroth-order Hamiltonian within the (LC-)DFTB method is evaluated in two-center approximation. Assume without loss of generality  $\mu \in A$  (read orbital  $\mu$  is on atom  $A$ ) and  $\nu \in B$ . Then for the off-site matrix elements  $A \neq B$  the zeroth-order Hamiltonian yields

$$\begin{aligned} H_{\nu\mu}^0 &= h_{\mu\nu} + v_{\mu\nu}^{\text{xc}} \left[ \sum_A^{\text{atoms}} \rho_A \right] + \sum_{\alpha} n_{\alpha}(\mu\nu|\alpha\alpha) - \frac{1}{2} \sum_{\alpha} n_{\alpha}(\mu\alpha|\alpha\nu)^{\text{lr}} \\ &= h_{\mu\nu} + v_{\mu\nu}^{\text{xc}} \left[ \rho_A + \rho_B + \sum_{C \neq \{A,B\}}^{\text{atoms}} \rho_C \right] + \sum_{\alpha \in \{A,B\}} n_{\alpha}(\mu\nu|\alpha\alpha) - \frac{1}{2} \sum_{\alpha \in \{A,B\}} n_{\alpha}(\mu\alpha|\alpha\nu)^{\text{lr}} \\ &+ \sum_{C \neq \{A,B\}} \sum_{\alpha \in C} n_{\alpha}(\mu\nu|\alpha\alpha) - \frac{1}{2} \sum_{C \neq \{A,B\}} \sum_{\alpha \in C} n_{\alpha}(\mu\alpha|\alpha\nu)^{\text{lr}} \end{aligned} \quad (3.17)$$

$$\approx h_{\mu\nu} + v_{\mu\nu}^{\text{xc}} [\rho_A + \rho_B] + \sum_{\alpha \in \{A,B\}} n_{\alpha}(\mu\nu|\alpha\alpha) - \frac{1}{2} \sum_{\alpha \in \{A,B\}} n_{\alpha}(\mu\alpha|\alpha\nu)^{\text{lr}}, \quad (3.18)$$

where the three-center terms have been neglected. This expression is exact for a dimer. For the on-site case  $A = B$ , the Hamiltonian is approximated as  $H_{\mu\nu}^0 \approx \epsilon_{\mu}^{\text{free atom}} \delta_{\mu\nu}$ . Thus the prescription for the LC-DFTB zeroth-order Hamiltonian is

$$H_{\mu\nu}^0[\rho_0] \approx \begin{cases} \epsilon_{\mu}^{\text{free atom}} & \nu = \mu \\ H_{\mu\nu}^0[\rho_A + \rho_B] & \mu \in A, \nu \in B \\ 0 & \text{else.} \end{cases} \quad (3.19)$$

In this way the approximate Hamiltonian can be constructed from a small set of precomputed matrix elements using the Slater-Koster rules [183]. If only s,p and d-orbitals are involved in the calculation, which is the case for the (LC-)DFTB method, only 10 Slater-Koster integrals as a function of the interatomic distance for each pair of involved atomic species are needed to construct an arbitrary molecular Hamiltonian.

### 3.4 BASIS SET

Up to now we did not mention how to choose the auxiliary basis set  $\{\phi_{\mu}\}$ . In the current approach the total energy, the Hamiltonian and Kohn-Sham molecular orbitals are represented in terms of finite auxiliary basis. The finiteness of the basis leads in general to an approximate representation of the Kohn-Sham orbitals and for this reason the densities, obtained in this way will be not the exact ground state densities. On the other hand the size of the basis defines the size of the matrices involved in the generalized eigenvalue problem Eq. 3.16 and thus the computational effort grows if approaching the basis set limit. Therefore it is a common practice to find an optimum with respect to the basis size and accuracy in the representation of the wave function.

### 3.4.1 Minimal pseudo-atomic basis set (MPA set)

A meaningful atom-centered basis set should consist of at least one radial function per atomic shell. Such a choice of basis is called minimal. For example in the case of carbon, the minimal set consists of 1s, 2s, 2p radial functions. Thus the overall number of basis functions is five (2p orbital is three-fold degenerate).

We obtain basis functions from the KS-orbitals of an atomic (LC-)DFT problem with additional confinement potential (we denote  $r = |\mathbf{r}|$ )

$$V^{\text{conf}}(r) = \left(\frac{r}{r_0}\right)^2, \quad (3.20)$$

where the compression radius  $r_0$  is usually proportional to the covalent radius of the respective atom. The compression radius is the optimization parameter of the basis set. The atoms are forced to be spherically symmetric. The electrons of the highest occupied shell are equally distributed over the symmetry states of that shell. For example the carbon atom has the occupation of 1/3 electrons in each spin orbital of the p-shell. The optimal choice of the compression radius can be obtained by a fit to some reference band structures [42]. We use the values of compression radii, which were obtained for the standard DFTB parametrization (the values labeled  $r_0^{\text{basis}}$  in Tab. 3.1), and call the minimal pseudo-atomic basis for this choice of the compression radii the MPA basis or DFTB basis.

We want now to assess the quality of the minimal basis, obtained in such a way, compared to standard Gaussian-type orbital (GTO) basis sets. We demonstrate the total energy for dimers  $N_2$ ,  $C_2$ ,  $O_2$  and  $CO$  as a function of interatomic distance as obtained from an all-electron LDA-DFT calculation with different basis sets in Fig. 3.2. The calculations with MPA set have been performed using the two-center all-electron (LC-)DFT code, developed by author for testing purposes. The code is based on the numerical integration routines, described in chapter 4. The calculations with the GTO basis set have been performed with NWCHEM package [195]. The MPA basis is compared to the minimal GTO basis (STO-3G), small double zeta GTO basis (3-21G) and large double zeta GTO basis (cc-pVDZ). Inspection of the figure Fig. 3.2 shows, that the minimal DFTB basis gives total energies which are better than that of the 3-21G basis set. Remarkable is the fact, that the minimal GTO basis (STO-3G) gives the energies, which are approximately 1 Hartree too high, compared to the minimal DFTB basis. We conclude, that although the DFTB basis set is a minimal one, it is comparable to small GTO double zeta basis. This finding will be reconfirmed also for the entire LC-DFTB method later on, despite the fact, that it employs further approximations. There are, however, cases where the minimal basis set is not sufficient. For sulfur [139] and phosphorus the extension of minimal basis set to include the polarization functions is necessary in order to correctly describe the hypervalent molecules (for example sulfur hexafluoride or phosphorus pentachloride). We note also, that to our best knowledge up to now there were no implementations of DFTB method with double zeta basis.



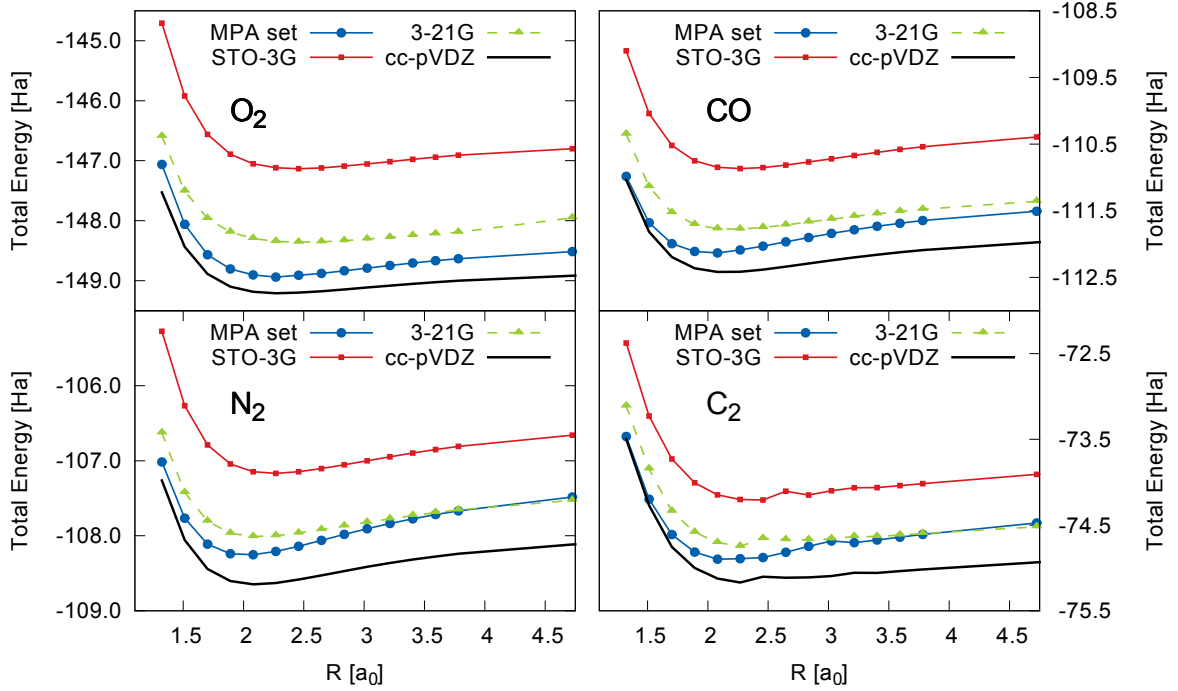


Figure 3.2: The total energy of  $N_2$ ,  $C_2$ ,  $O_2$  and  $CO$  molecules as a function of interatomic distance  $R$  from all-electron LDA-DFT with different basis sets. The minimal DFTB basis performs comparable or better than the small double zeta basis (3-21G) for the dimers.

### 3.4.2 Basis for LC-DFTB

In the standard DFTB method the pseudo-atomic basis set, described in the previous section is further truncated and contains only the valence orbitals. For example carbon atom would contain only 2s and 2p orbitals. This is based on the assumption, that the contribution to the electronic structure due to the core electrons can be neglected.

As has been already stated, the zeroth-order Hamiltonian is evaluated according to the prescription Eq. 3.19. Usually the density matrix, which enters the off-site Hamiltonian is evaluated using the pseudo-atomic basis with weaker compression. It has been found that this generally improves the description of electronic structure. Thus the zeroth-order Hamiltonian and overlap matrix elements are optimized with respect to both the basis compression radius  $r_0^{\text{basis}}$  and the density compression radius  $r_0^{\text{density}}$ . For the LC-DFTB we use the same compression radii for all values of the range-separation parameter  $\omega$ . We rely on the values, optimized for the standard DFTB [42, 139]. This choice is motivated by the requirement, that the method should approach the standard DFTB solution in the limit  $\omega \rightarrow 0$  on the one hand. On the other hand, the reoptimization of the compression radii for each value of  $\omega$  would overparametrize the method. The compression radii used in the calculations for this thesis are summarized in Tab. 3.1.

Species	$r_0^{\text{basis}} [a_0]$			$r_0^{\text{density}} [a_0]$		
	s	p	d	s	p	d
H	3.0			2.5		
C	2.7	2.7		14.0	14.0	
N	2.7	2.7		14.0	14.0	
O	2.3	2.3		9.0	9.0	
S	3.8	3.8	4.4	9.0	9.0	9.0

Table 3.1: The compression radii  $[a_0]$  for the basis and density compression for the elements H,C,N,O,S, used in this thesis. These are the values of the standard DFTB parametrization, referred to as mio-1-1 parameter set.

### 3.5 EXTENSION OF PARAMETRIZATION TOOLS AND COMPUTATIONAL PERFORMANCE

The parametrization of the zeroth-order LC-DFTB requires the extension of the atomic DFT code and of the two-center code, which are part of the DFTB parametrization toolkit, to include the long-range corrected functional. For both, the local short-range xc-functional is implemented according to the description in section 3.1. The methods for the evaluation of the long-range HF exchange integrals are outlined in section 4.4 (two-center code) and section 4.3 (atomic DFT code). The modified atomic DFT code provides the pseudo-atomic eigenvalues  $\epsilon_\mu^{\text{free atom}}$ , the basis  $\{\phi_\mu(\mathbf{r})\}$  and the density matrix  $\rho_A(\mathbf{r}, \mathbf{r}')$ . The two-center code evaluates then the off-site Hamiltonian  $H_{\mu\nu}^0$ ,  $\mu \in A, \nu \in B, A \neq B$  and the overlap integrals  $S_{\mu\nu}$  for the given pseudo-atomic basis and pseudo-atomic density matrix. The parameters are then tabulated in the Slater-Koster files as a function of interatomic distance. We note, that for the zeroth-order LC-DFTB the only changes are due to the zeroth-order Hamiltonian and overlap matrix elements. For this reason the computational efficiency of the zeroth-order LC-DFTB (LC-DFTB-0) is the same as that of standard DFTB.

### 3.6 EIGENVALUES FROM THE ZERO-ORDER LC-DFTB

In order to test the performance of the new parametrization we calculate the eigenvalues of small closed-shell molecules benzene, ethylene,  $C_2$  and  $N_2$  as a function of range-separation parameter  $\omega$ . We compare the results to the all-electron first principles LC-DFT with BNL functional. The BNL eigenvalues have been calculated using the minimal basis (STO-3G) and the large double zeta basis (cc-pVDZ) with NWCHEM package. The molecular geometries have been optimized on the standard DFTB level. The HOMO and LUMO eigenvalues for zeroth-order LC-DFTB (LC-DFTB-0), the BNL/STO-3G and BNL/cc-pVDZ are depicted in Fig. 3.3. Additionally, we plot the results from the self-consistent LC-DFTB (red dashed lines), which will be introduced later. We recognize that the LC-DFTB-0 theory does not provide the necessary qualitative effect of pushing down the HOMO eigenvalue and opening the HOMO-LUMO gap if the range-separation

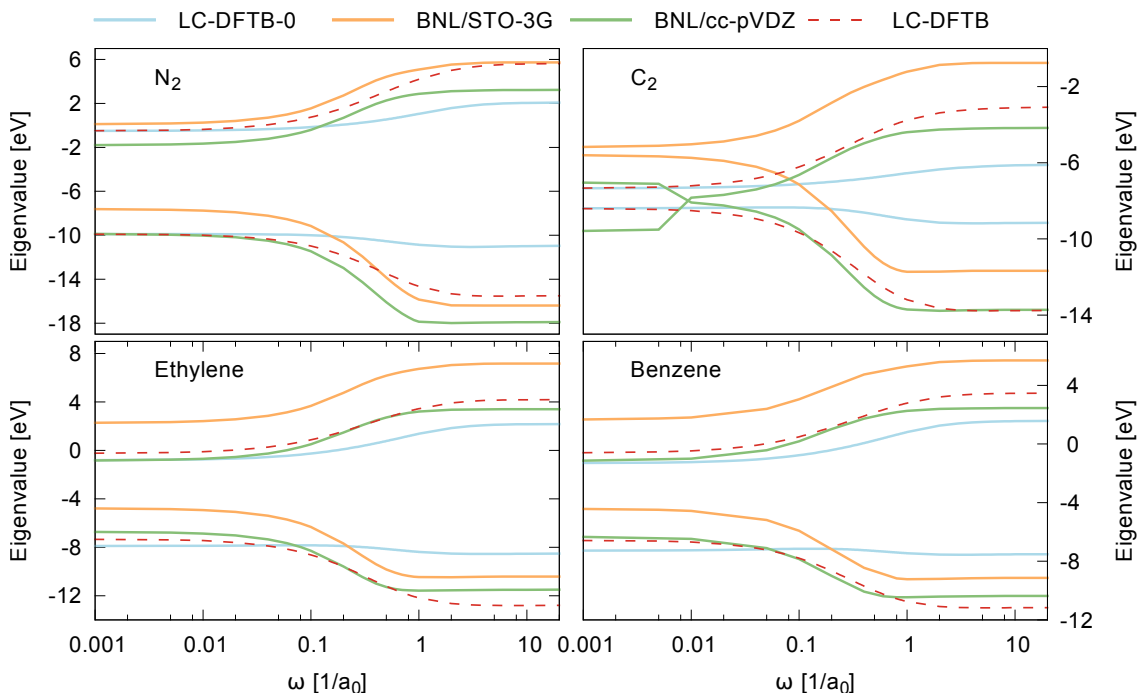


Figure 3.3: The HOMO and LUMO of benzene, ethylene,  $C_2$  and  $N_2$  molecules, calculated with LC-DFTB-0 theory and full *ab initio* BNL/STO-3G and BNL/cc-pVDZ theories. The HOMO eigenvalue of the LC-DFTB-0 is not properly pushed down in energy as the range-separation parameter  $\omega$  is increased. The result for self-consistent LC-DFTB, which is introduced in following sections is included for comparison.

parameter is increased as it is the case for the first-principles methods. Although the gap gets larger, the effect is rather small, compared to the first-principles reference. The same is true for the HOMO eigenvalue. As has been discussed in chapter 2 the HOMO eigenvalue for the exact functional equals up to a sign to the ionization potential. Since it is usually underestimated by the local DFT (limit  $\omega \rightarrow 0$  in this case) the drop of the HOMO eigenvalue is the signature of the self-interaction error correction, which is not observed for LC-DFTB-0. Contrary to this the DFTB limit  $\omega \rightarrow 0$  gives gaps and positions of HOMO and LUMO levels, which compare better to the corresponding *ab initio* theory (this is in fact no surprise, since the standard DFTB has been optimized to do this).

The problem of insufficient description of the eigenvalues for the typical values of the range-separation parameter  $0.1a_0^{-1} < \omega < 1.0a_0^{-1}$  and in the HF+c limit ( $\omega \rightarrow \infty$ ) does not seem to be the basis set effect, since the minimal STO-3G basis shows the proper gap opening and lowering of the HOMO eigenvalue. The two-center approximation can not explain the problem either, since even for the case of dimers  $N_2, C_2$  this problem occurs. We conclude, that the reason should be the non-self-consistent evaluation of the eigenvalues within the LC-DFTB-0 method.

The LC-DFTB-0 method performs only one diagonalization. For this reason, our aim is to

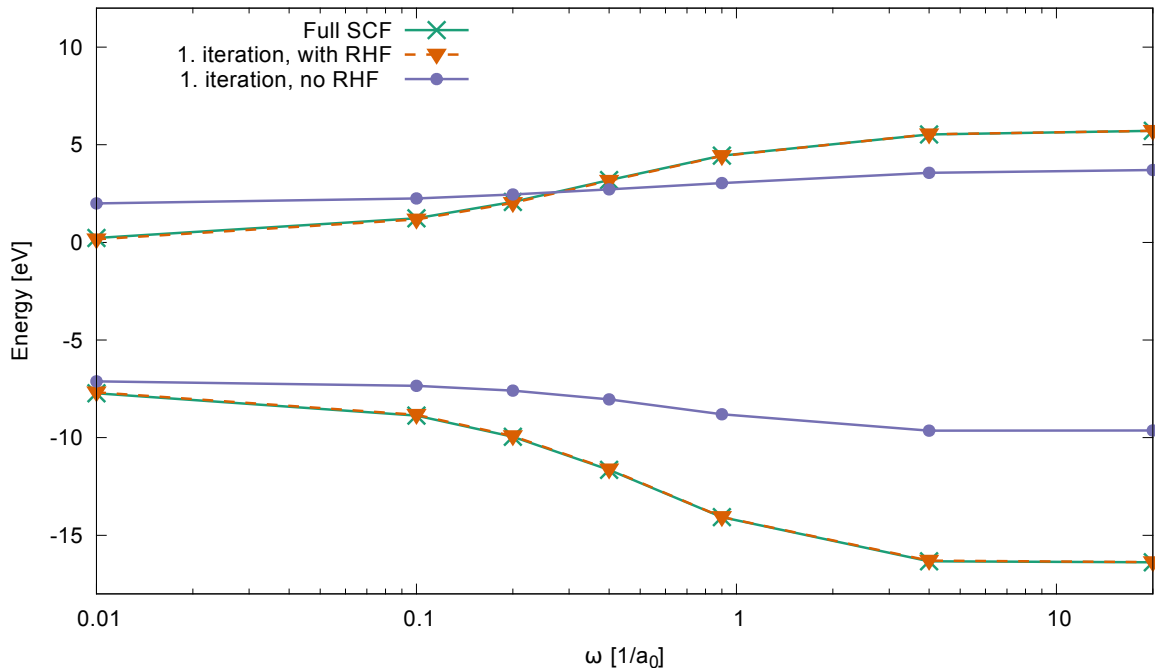


Figure 3.4: The HOMO and LUMO eigenvalues of a  $N_2$  molecule as a function of the range-separation parameter  $\omega$ . The calculation have been performed on the BNL/STO-3G theory level using one diagonalization only (no RHF), one diagonalization and RHF pre-diagonalization (with RHF) and full SCF procedure until convergence.

perform only one diagonalization for a reference all-electron *ab initio* method. The NWCHEM [195] code, which we extensively use, supports the superposition of atomic densities guess. However, for the efficiency reasons, it diagonalizes the initial Hamiltonian on the restricted Hartree-Fock (RHF) level of theory and uses the resulting density matrix as the input for the first LC-DFT self-consistent field (SCF) cycle. Thus performing only one SCF iteration with NWCHEM gives in fact the second SCF iteration (RHF  $\rightarrow$  1. DFT iteration). For this reason we can not use the NWCHEM for the verification of our conjecture without patching it.

Instead, we use the aforementioned two-center all-electron code, which is based on the numerical basis sets. It is thus independent on the type of the basis and can use the MPA basis and GTO basis sets. We calculate the HOMO and LUMO of  $N_2$  dimer within the all-electron BNL/STO-3G theory using the superposition of atomic densities guess as a function of range-separation parameter  $\omega$ . We calculate the eigenvalues from one diagonalization with and without preceding RHF diagonalization. The eigenvalues from the full SCF procedure are given as well. The results are presented in Fig. 3.4. In the case of one diagonalization (no RHF) we observe the same qualitative behavior as in the case of LC-DFTB-0 theory. The deviation between the eigenvalues from the full SCF calculation and the ones from one diagonalization (no RHF) gets larger if range-separation parameter  $\omega$  is increased. On contrast the eigenvalues from one

diagonalization with preceding RHF diagonalization are very close to that from full SCF at least for dimer. If the first diagonalization is carried out within the RHF theory, then the resulting density matrix is not anymore diagonal in atomic basis. It gives then a better initial guess for a Hamiltonian, which contains the HF exchange term. The exploitation of such reference density for the parametrization of the LC-DFTB method requires, however, a careful revision of the employed DFTB approximations.

We conclude that for the superposition of atomic densities guess the zeroth-order theory does not show the behavior, expected from a successful self-interaction error correction. For this reason the SCF extension, described in next section is required.

### 3.7 SELF-CONSISTENT FIELD LC-DFTB

The implementation of the LC-DFTB-0 and analysis of the results presented in the previous section suggested the necessity of the self-consistency in the solution of the LC-DFTB equations. In following, the results of ref. [137] are rederived and further extended to the practical computational scheme.

We come back to the energy expansion Eq. 3.13 and include now the terms, which are of the second order in the difference density matrix  $\Delta P_{\mu\nu}$

$$E^{2\text{nd}} = \frac{1}{2} \sum_{\mu\nu\alpha\beta} \Delta P_{\mu\nu} \Delta P_{\alpha\beta} \left[ (\mu\nu|\alpha\beta) + f_{\mu\nu\alpha\beta}^{\text{xc}} \right] - \frac{1}{4} \sum_{\mu\nu\alpha\beta} \Delta P_{\mu\nu} \Delta P_{\alpha\beta} (\mu\alpha|\beta\nu)^{\text{lr}}. \quad (3.21)$$

Here the first term contains the Hartree energy and the energy contribution due to the linearized xc-potential and the second term is the long-range HF exchange term.

As in the case of zeroth-order contributions the aim is to apply the dimer (or two-center) approximation in order to reduce the complexity of the method. A widely used integral approximation which allows to reduce the general four-center integrals to the sum of two-center integrals is known as the Mulliken approximation [8]. Assume the charge distribution  $\phi_\mu(\mathbf{r})\phi_\nu(\mathbf{r})$ ,  $\mu \in A$ ,  $\nu \in B$ , generated by two orbitals, which are located at centers  $A$  and  $B$ . We assume that the orbitals  $\phi_\alpha$  form a complete basis at a given center. Then each of these orbitals can be expanded at the different center

$$\phi_\mu(\mathbf{r}) = \sum_{\sigma \in B} \left[ \int \phi_\mu(\mathbf{r}') \phi_\sigma(\mathbf{r}') d\mathbf{r}' \right] \phi_\sigma(\mathbf{r}) = \sum_{\sigma \in B} S_{\mu\sigma} \phi_\sigma(\mathbf{r}) \quad (3.22)$$

$$\phi_\nu(\mathbf{r}) = \sum_{\sigma \in A} \left[ \int \phi_\nu(\mathbf{r}') \phi_\sigma(\mathbf{r}') d\mathbf{r}' \right] \phi_\sigma(\mathbf{r}) = \sum_{\sigma \in A} S_{\nu\sigma} \phi_\sigma(\mathbf{r}). \quad (3.23)$$

The product of the orbitals can be then expressed in the symmetrized form

$$\begin{aligned}\phi_\mu(\mathbf{r})\phi_\nu(\mathbf{r}) &= \frac{1}{2} \left[ \left( S_{\nu\mu}\phi_\mu(\mathbf{r}) + \sum_{\substack{\sigma \in A \\ \sigma \neq \mu}} S_{\nu\sigma}\phi_\sigma(\mathbf{r}) \right) \phi_\mu(\mathbf{r}) + \left( S_{\mu\nu}\phi_\nu(\mathbf{r}) + \sum_{\substack{\sigma \in B \\ \sigma \neq \nu}} S_{\mu\sigma}\phi_\sigma(\mathbf{r}) \right) \phi_\nu(\mathbf{r}) \right] \\ &= \frac{1}{2} S_{\mu\nu} (|\phi_\mu(\mathbf{r})|^2 + |\phi_\nu(\mathbf{r})|^2) + \frac{1}{2} \left( \sum_{\substack{\sigma \in A \\ \sigma \neq \mu}} S_{\nu\sigma}\phi_\sigma(\mathbf{r})\phi_\mu(\mathbf{r}) + \sum_{\substack{\sigma \in B \\ \sigma \neq \nu}} S_{\mu\sigma}\phi_\sigma(\mathbf{r})\phi_\nu(\mathbf{r}) \right).\end{aligned}\quad (3.24)$$

The first term of this expansion is known as the Mulliken approximation. Note, that if the orbitals are located on the same atom ( $A = B$ ), the charge distribution within this approximation is non-zero only if the orbitals are equal. We approximate the two-electron integrals using this approximation

$$(\mu\nu|\alpha\beta) \approx \frac{1}{4} S_{\mu\nu} S_{\alpha\beta} ((\mu\mu|\alpha\alpha) + (\mu\mu|\beta\beta) + (\nu\nu|\alpha\alpha) + (\nu\nu|\beta\beta)). \quad (3.25)$$

In this way the four-index quantities of the second-order term Eq. 3.21 can be written as

$$\begin{aligned}(\mu\nu|\alpha\beta) + f_{\mu\nu\alpha\beta}^{\text{xc}} &= \int \phi_\mu(\mathbf{r})\phi_\nu(\mathbf{r})\phi_\alpha(\mathbf{r}')\phi_\beta(\mathbf{r}') \left[ \frac{1}{|\mathbf{r}-\mathbf{r}'|} + f_{\text{xc}}^\omega[\rho_0](\mathbf{r},\mathbf{r}') \right] d\mathbf{r}d\mathbf{r}' \\ &\approx \frac{1}{4} S_{\mu\nu} S_{\alpha\beta} (\gamma_{\mu\alpha}^{\text{fr}} + \gamma_{\nu\alpha}^{\text{fr}} + \gamma_{\mu\beta}^{\text{fr}} + \gamma_{\nu\beta}^{\text{fr}})\end{aligned}\quad (3.26)$$

$$\begin{aligned}(\mu\nu|\alpha\beta)^{\text{lr}} &= \int \phi_\mu(\mathbf{r})\phi_\nu(\mathbf{r})\phi_\alpha(\mathbf{r}')\phi_\beta(\mathbf{r}') \frac{1 - \exp(-\omega|\mathbf{r}-\mathbf{r}'|)}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}d\mathbf{r}' \\ &\approx \frac{1}{4} S_{\mu\nu} S_{\alpha\beta} (\gamma_{\mu\alpha}^{\text{lr}} + \gamma_{\nu\alpha}^{\text{lr}} + \gamma_{\mu\beta}^{\text{lr}} + \gamma_{\nu\beta}^{\text{lr}}),\end{aligned}\quad (3.27)$$

where the full-range (fr) and long-range (lr)  $\gamma$ -integrals are introduced

$$\gamma_{\mu\nu}^{\text{fr}} = \int |\phi_\mu(\mathbf{r})|^2 |\phi_\nu(\mathbf{r}')|^2 \left[ \frac{1}{|\mathbf{r}-\mathbf{r}'|} + f_{\text{xc}}^\omega[\rho_0] \right] d\mathbf{r}d\mathbf{r}' \quad (3.28)$$

$$\gamma_{\mu\nu}^{\text{lr}} = \int |\phi_\mu(\mathbf{r})|^2 |\phi_\nu(\mathbf{r}')|^2 \frac{1 - \exp(-\omega|\mathbf{r}-\mathbf{r}'|)}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}d\mathbf{r}'. \quad (3.29)$$

The charge distributions  $|\phi_\mu(\mathbf{r})|^2$  under the integral constitute the averaged charge distributions

$$F_A(\mathbf{r}) = \frac{1}{(l+1)^2} \sum_{\mu \in A} |\phi_\mu(\mathbf{r})|^2, \quad (3.30)$$

which are assumed to have the form

$$F_A(\mathbf{r}) = \frac{\tau_A^3}{8\pi} e^{-\tau_A|\mathbf{r}-\mathbf{R}_A|}, \quad (3.31)$$

where  $\mathbf{R}_A$  is the position of the atom  $A$  and the decay constant  $\tau_A$  has still to be determined. We obtain the final expressions for the integrals by replacing the initial charge distributions with the averaged ones

$$\gamma_{\mu\nu}^{\text{fr}} = \gamma_{AB}^{\text{fr}} = \frac{\tau_A^3 \tau_B^3}{(8\pi)^2} \int e^{-\tau_A|\mathbf{r}-\mathbf{R}_A|} e^{-\tau_B|\mathbf{r}'-\mathbf{R}_B|} \left[ \frac{1}{|\mathbf{r}-\mathbf{r}'|} + f_{\text{xc}}^\omega[\rho_0] \right] d\mathbf{r}d\mathbf{r}' \quad (3.32)$$

$$\gamma_{\mu\nu}^{\text{lr}} = \gamma_{AB}^{\text{lr}} = \frac{\tau_A^3 \tau_B^3}{(8\pi)^2} \int e^{-\tau_A|\mathbf{r}-\mathbf{R}_A|} e^{-\tau_B|\mathbf{r}'-\mathbf{R}_B|} \frac{1 - \exp(-\omega|\mathbf{r}-\mathbf{r}'|)}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}d\mathbf{r}'. \quad (3.33)$$

This is also known as monopole approximation [42]. We evaluate these integrals in the same way as in the standard DFTB method. For the off-site elements  $A \neq B$  the contribution to the full-range integral due to the exchange-correlation kernel  $f_{xc}^\omega$  is assumed to vanish. The full-range integral  $\gamma^{\text{fr}}$  is then just a two-center Coulomb integral over spherically symmetric charge distribution with Slater-type profile Eq. 3.31. The long-range integral  $\gamma^{\text{lr}}$  differs only in the interaction operator, which can be separated into the Coulomb and Yukawa parts. For both cases analytical formulas are available. The case of Coulomb interaction has been covered in the original paper on the self-consistent extension of the DFTB method [42]. We extend the formula for the more general case of Yukawa interaction. The  $\gamma$ -integral over the Yukawa interaction yields (see chapter A for details)

$$\begin{aligned} \gamma_{AB}^{Y,\omega} &= \frac{\tau_A^3 \tau_B^3}{(8\pi)^2} \int e^{-\tau_A |\mathbf{r}-\mathbf{R}_A|} e^{-\tau_B |\mathbf{r}'-\mathbf{R}_B|} \frac{\exp(-\omega |\mathbf{r}-\mathbf{r}'|)}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r} d\mathbf{r}' \\ &= \frac{\tau_A^4 \tau_B^4}{(\tau_A^2 - \omega^2)^2 (\tau_B^2 - \omega^2)^2} \frac{e^{-\omega R}}{R} \\ &\quad - \left[ e^{-\tau_A R} \left( \frac{\tau_A^2}{\tau_A^2 - \omega^2} \frac{\tau_A \tau_B^4}{2(\tau_B^2 - \tau_A^2)^2} - \frac{\tau_A^4}{(\omega^2 - \tau_A^2)^2} \frac{(\tau_B^6 - 3\tau_A^2 \tau_B^4 + 2\omega^2 \tau_B^4)}{(\tau_A^2 - \tau_B^2)^3 R} \right) \right. \\ &\quad \left. + e^{-\tau_B R} \left( \frac{\tau_B^2}{\tau_B^2 - \omega^2} \frac{\tau_B \tau_A^4}{2(\tau_A^2 - \tau_B^2)^2} - \frac{\tau_B^4}{(\omega^2 - \tau_B^2)^2} \frac{(\tau_A^6 - 3\tau_B^2 \tau_A^4 + 2\omega^2 \tau_A^4)}{(\tau_B^2 - \tau_A^2)^3 R} \right) \right]. \end{aligned} \quad (3.34)$$

It contains the result of Elstner et al. [42] as the special case ( $\omega \rightarrow 0$ ). Thus  $\gamma_{AB}^{\text{lr}} = \gamma_{AB}^{Y,0} - \gamma_{AB}^{Y,\omega}$ . The long-range  $\gamma$ -integral for different values of the range-separation parameter  $\omega$  and for the case of carbon-nitrogen interaction is plotted as a function of interatomic distance in Fig. 3.5.

In the standard DFTB method one usually requires the Hubbard derivative  $U_A^{\text{DFTB}}$  for a single atom  $A$  to be equal to the Hubbard derivative  $U_A^{\text{DFT}}$  from a reference DFT calculation (compare Eq. 1.26)

$$U_A^{\text{DFT}} = U_A^{\text{DFTB}}. \quad (3.35)$$

This fixes the values of the on-site  $\gamma$ -integrals and decay constants  $\tau_A$  for each atomic species [42]. We require that the LC-DFTB method for the case  $\omega \rightarrow 0$  resembles the DFTB method, based on the same local functional. For this reason we impose the condition Eq. 3.35 to the LC-DFTB. The presence of the long-range HF exchange term requires, however, the correction of the scheme as described in section 3.8. In short, the decay constants for each atomic species have to be corrected.

With the approximations and definitions described above, the total energy of the LC-DFTB method reads

$$\begin{aligned} E &= \sum_{\mu\nu} P_{\mu\nu} H_{\mu\nu}^0 + \frac{1}{8} \sum_{\mu\nu\alpha\beta} \Delta P_{\alpha\beta} \Delta P_{\mu\nu} S_{\mu\nu} S_{\alpha\beta} \left( \gamma_{\mu\alpha}^{\text{fr}} + \gamma_{\mu\beta}^{\text{fr}} + \gamma_{\nu\alpha}^{\text{fr}} + \gamma_{\nu\beta}^{\text{fr}} \right) \\ &\quad - \frac{1}{16} \sum_{\mu\nu\alpha\beta} \Delta P_{\alpha\beta} \Delta P_{\mu\nu} S_{\mu\alpha} S_{\beta\nu} \left( \gamma_{\mu\beta}^{\text{lr}} + \gamma_{\mu\nu}^{\text{lr}} + \gamma_{\alpha\beta}^{\text{lr}} + \gamma_{\alpha\nu}^{\text{lr}} \right) + E_{\text{rep}}. \end{aligned} \quad (3.36)$$

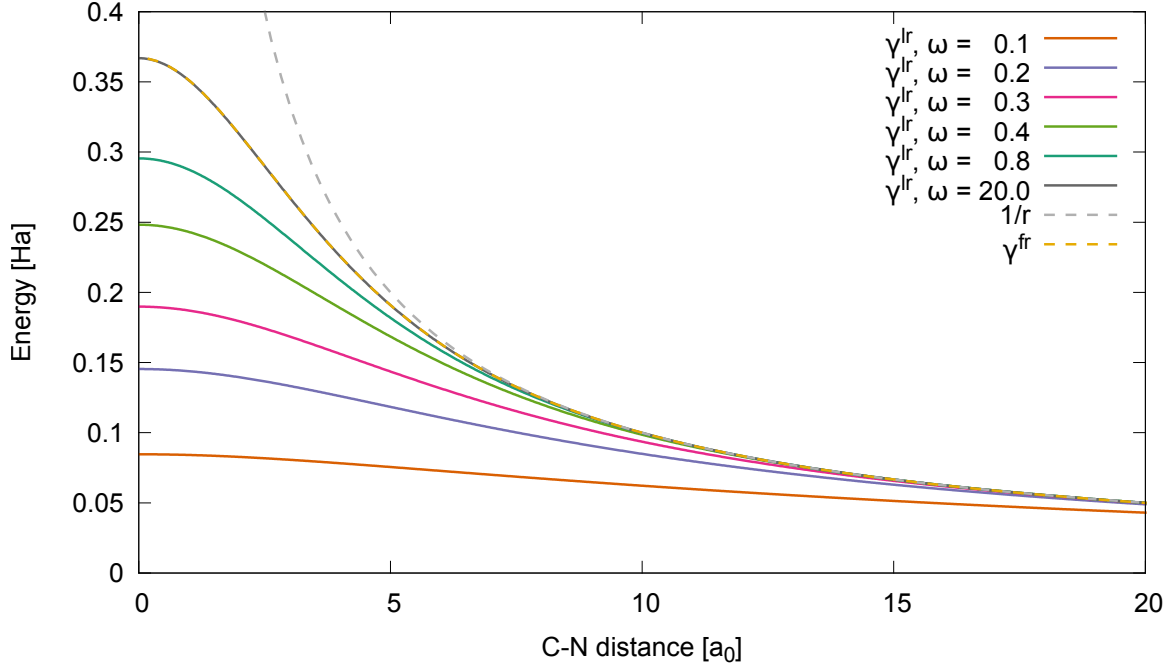


Figure 3.5: The long-range  $\gamma$ -integral for the carbon-nitrogen interaction as a function of interatomic distance for different values of the range-separation parameter  $\omega$  (all values are in  $\text{\AA}^{-1}$ ). The gray dashed line indicates the  $1/r$  limit. The full-range integral  $\gamma^{\text{fr}}$  is depicted with orange dashed line.

Applying the variational principle with respect to the molecular orbital coefficients  $c_{\mu i}$  to the energy expression we obtain the generalized eigenvalue problem as before with the Hamiltonian

$$\begin{aligned}
 H_{\mu\nu} = H_{\mu\nu}^0 + \frac{1}{4} \sum_{\alpha\beta} \Delta P_{\alpha\beta} S_{\mu\nu} S_{\alpha\beta} \left( \gamma_{\mu\alpha}^{\text{fr}} + \gamma_{\mu\beta}^{\text{fr}} + \gamma_{\nu\alpha}^{\text{fr}} + \gamma_{\nu\beta}^{\text{fr}} \right) \\
 - \frac{1}{8} \sum_{\alpha\beta} \Delta P_{\alpha\beta} S_{\mu\alpha} S_{\beta\nu} \left( \gamma_{\mu\beta}^{\text{lr}} + \gamma_{\mu\nu}^{\text{lr}} + \gamma_{\alpha\beta}^{\text{lr}} + \gamma_{\alpha\nu}^{\text{lr}} \right). \quad (3.37)
 \end{aligned}$$

Since the long-range exchange term can not be formulated in terms of Mulliken charges as it is the case for the Hartree term the self-consistency in the LC-DFTB method is achieved with respect to the density matrix and not with respect to the Mulliken charges as in the standard DFTB. This is the case even in the limit  $\omega \rightarrow 0$ . The convergence optimizing algorithms (density mixing and level-shifting) as implemented for the DFTB method can be used with minor modifications also for the LC-DFTB.



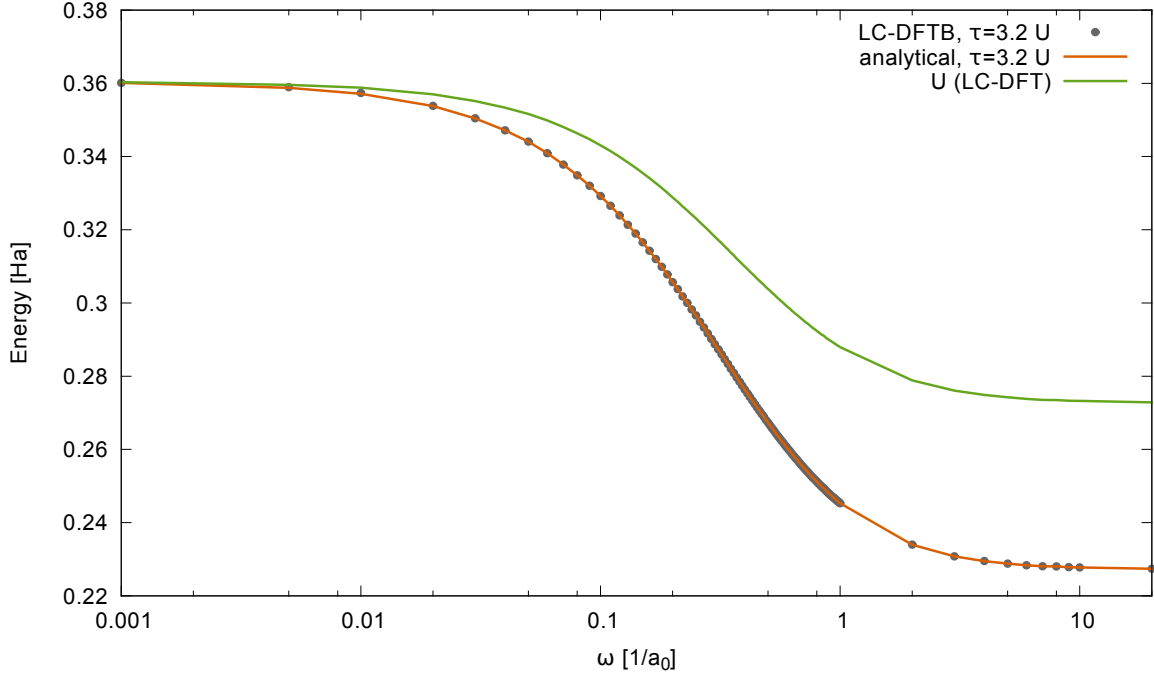


Figure 3.6: The atomic Hubbard parameter  $U = \frac{\partial^2 E^{\text{atom}}}{\partial n^2}$  as a function of the range-separation parameter  $\omega$  for a carbon atom obtained from LC-DFT (green line). The Hubbard parameter from the LC-DFTB method (gray dots) was calculated with the decay constant  $\tau = 3.2U$ . We observe the violation of the condition Eq. 3.35. The analytical formula Eq. (3.52) (orange line) is plotted for completeness.

### 3.8 HUBBARD PARAMETERS

In the standard DFTB method parameters  $\tau_A$  for the  $\gamma$ -integrals are fixed according to the condition (Eq. 1.26)

$$\tau_A = \frac{16}{5} U_A^{\text{DFT}} = 3.2 U_A^{\text{DFT}}. \quad (3.38)$$

If we use this simple definition for the LC-DFTB we find that the initial requirement Eq. 3.35

$$U_A^{\text{LC-DFT}} = U_A^{\text{LC-DFTB}} \quad (3.39)$$

is violated. We demonstrate this for the case of a carbon atom in Fig. 3.6. We plot the atomic Hubbard derivative, calculated with the LC-DFT method (green line) as the function of the range-separation parameter  $\omega$ . This is compared to the Hubbard derivative, obtained by the LC-DFTB method, if the naive definition of decay constant Eq. 3.38 is used (gray dots). The two curves increasingly differ if the parameter  $\omega$  is increased.

### 3.8.1 Influence of the long-range HF exchange term

The  $\gamma$ -integral over the Yukawa interaction with screening parameter  $\omega$  can be reduced to the one-dimensional integral (section A.1)

$$\gamma_{AB}^{Y,\omega} = \frac{2\tau_A^4 \tau_B^4}{\pi R_{AB}} \int_0^\infty \frac{q \sin(qR_{AB})}{(q^2 + \tau_A^2)^2 (q^2 + \tau_B^2)^2 (q^2 + \omega^2)} dq. \quad (3.40)$$

To obtain the expression for the on-site value  $A = B$  of this integral, where  $\tau_A = \tau_B$  and  $R_{AB} \rightarrow 0$ , we expand the sine function around 0 and apply the residue theorem (section A.3). This results in the expression

$$\gamma_{AA}^{Y,\omega} = \frac{\tau_A^8}{(\tau_A^2 - \omega^2)^4} \left[ \frac{5\tau_A^6 + 15\tau_A^4 \omega^2 - 5\tau_A^2 \omega^4 + \omega^6}{16\tau_A^5} - \omega \right]. \quad (3.41)$$

For the case  $\omega \rightarrow 0$  the integral Eq. 3.40 is taken over the Coulomb interaction and this expression reduces to the already mentioned DFTB result, which in LC-DFTB is equal to the full-range  $\gamma$ -integral

$$\gamma_{AA}^{\text{fr}} = \lim_{\omega \rightarrow 0} \gamma_{AA}^{Y,\omega} = \frac{5}{16} \tau_A. \quad (3.42)$$

The on-site value of the long-range  $\gamma$ -integral  $\gamma_{AA}^{\text{lr}} = \gamma_{AA}^{Y,0} - \gamma_{AA}^{Y,\omega}$  reads

$$\gamma_{AA}^{\text{lr}} = \frac{5}{16} \tau_A - \frac{\tau_A^8}{(\tau_A^2 - \omega^2)^4} \left[ \frac{5\tau_A^6 + 15\tau_A^4 \omega^2 - 5\tau_A^2 \omega^4 + \omega^6}{16\tau_A^5} - \omega \right]. \quad (3.43)$$

In order to ensure the basis set consistency, the parameters  $\tau_A$  employed in long-range and full-range  $\gamma$ -integrals should be related. We require, that the decay constant of a particular atom for both full-range and long-range  $\gamma$ -integral should be the same (see also subsection 3.8.2).

We derive the analytical expression for the Hubbard derivative as obtained from the LC-DFTB. We consider the total energy of the LC-DFTB method for the case of a single atom. Note that in this case  $\mu, \nu \in A$ ,  $S_{\mu\nu} = \delta_{\mu\nu}$ ,  $H_{\mu\nu}^0 = \delta_{\mu\nu} \epsilon_\mu^{\text{free atom}}$ ,  $\gamma_{\mu\nu}^{\text{lr/fr}} = \gamma_{AA}^{\text{lr/fr}} = \gamma^{\text{lr/fr}}$  and the total energy expression reads

$$E^{\text{atom}} = \sum_{\mu} P_{\mu\mu} \epsilon_{\mu}^{\text{free atom}} + \frac{1}{2} \gamma^{\text{fr}} \sum_{\mu\alpha} \Delta P_{\mu\mu} \Delta P_{\alpha\alpha} - \frac{1}{4} \gamma^{\text{lr}} \sum_{\mu\nu} \Delta P_{\mu\nu} \Delta P_{\mu\nu}. \quad (3.44)$$

The difference density matrix can be expressed as

$$\Delta P_{\mu\nu} = P_{\mu\nu} - P_{\mu\nu}^0 = \sum_{i=1}^N n_i c_{\mu i} c_{\nu i} - \sum_{i=1}^N n_i^0 \delta_{\mu\nu}, \quad (3.45)$$

where the  $n_i^0$  are the occupations of the reference density matrix  $P_{\mu\nu}^0$ . We note, that in the case of one atom, the orthogonality of the molecular orbitals implies

$$\delta_{ij} = \int \psi_i(\mathbf{r}) \psi_j(\mathbf{r}) d\mathbf{r} = \sum_{\mu\nu} c_{\mu i} c_{\nu j} \int \phi_{\mu}(\mathbf{r}) \phi_{\nu}(\mathbf{r}) d\mathbf{r} = \sum_{\mu\nu} c_{\mu i} c_{\nu j} S_{\mu\nu} = \sum_{\mu} c_{\mu i} c_{\mu j}. \quad (3.46)$$

Using these definitions we obtain

$$\begin{aligned} \sum_{\mu\alpha} \Delta P_{\mu\mu} \Delta P_{\alpha\alpha} &= \sum_{\mu} \sum_{i=1}^N (n_i c_{\mu i} c_{\mu i} - n_i^0) \sum_{\alpha} \sum_{j=1}^N (n_j c_{\alpha j} c_{\alpha j} - n_j^0) = \sum_{i=1}^N (n_i - \sum_{\mu} n_i^0) \sum_{j=1}^N (n_j - \sum_{\alpha} n_j^0) \\ &= \sum_{i=1}^N \sum_{j=1}^N n_i n_j + \text{terms linear in } n_i \end{aligned} \quad (3.47)$$

$$\begin{aligned} \sum_{\mu\nu} \Delta P_{\mu\nu} \Delta P_{\mu\nu} &= \sum_{\mu\nu} \sum_{i=1}^N (n_i c_{\mu i} c_{\nu i} - n_i^0 \delta_{\mu\nu}) \sum_{j=1}^N (n_j c_{\mu j} c_{\nu j} - n_j^0 \delta_{\mu\nu}) \\ &= \sum_{\mu\nu} \sum_{i=1}^N \sum_{j=1}^N n_i n_j c_{\mu i} c_{\nu i} c_{\mu j} c_{\nu j} + \text{terms linear in } n_i \\ &= \sum_{i=1}^N \sum_{j=1}^N n_i n_j \delta_{ij} + \text{terms linear in } n_i \\ &= \sum_{i=1}^N n_i^2 + \text{terms linear in } n_i. \end{aligned} \quad (3.48)$$

This gives the total energy of a single atom in the LC-DFTB method in terms of occupation numbers  $n_i$

$$E^{\text{atom}} = \frac{1}{2} \gamma^{\text{fr}} \sum_{i=1}^N \sum_{j=1}^N n_i n_j - \frac{1}{4} \gamma^{\text{lr}} \sum_{i=1}^N n_i^2 + \text{terms linear in } n_i, \quad (3.49)$$

where we write out only the terms, quadratic in occupation numbers. The Hubbard derivative is practically performed using finite difference method by varying the charge of the atom. The variation is done by rescaling the density matrix on the initial step of the calculation. The SCF converges then to the ground state, where the excess charge is distributed over the highest occupied shell.<sup>3</sup> Thus we obtain the change in the occupations of the highest occupied shell only and it suffices to consider only the contributions to the energy quadratic in the occupation of the orbital in that shell. The occupation of each orbital in the shell is  $n/(2l+1)$ , where  $n$  is (fractional) number of electrons in the shell and  $l$  its angular momentum. The total energy is then

$$\begin{aligned} E^{\text{atom}} &= \frac{1}{2} \gamma^{\text{fr}} \sum_i^{\text{shell}} \sum_j^{\text{shell}} \left( \frac{n}{2l+1} \right)^2 - \frac{1}{4} \gamma^{\text{lr}} \sum_i^{\text{shell}} \left( \frac{n}{2l+1} \right)^2 + \text{terms linear in } n \\ &= \frac{1}{2} \gamma^{\text{fr}} (2l+1)^2 \left( \frac{n}{2l+1} \right)^2 - \frac{1}{4} \gamma^{\text{lr}} (2l+1) \left( \frac{n}{2l+1} \right)^2 + \text{terms linear in } n \\ &= \frac{1}{2} \gamma^{\text{fr}} n^2 - \frac{1}{4} \gamma^{\text{lr}} \frac{1}{2l+1} n^2 + \text{terms linear in } n. \end{aligned} \quad (3.50)$$

Here the summation is performed over the highest occupied shell. Performing the second derivative with respect to the occupation  $n$  we obtain

$$U^{\text{LC-DFTB}} = \frac{\partial^2 E^{\text{atom}}}{\partial n^2} = \gamma^{\text{fr}} - \frac{1}{2} \gamma^{\text{lr}} \frac{1}{2l+1}. \quad (3.51)$$

This, together with Eqn. 3.43 and 3.42 provides the analytical expression for the value of the

<sup>3</sup>This is known as aufbau-principle [150].

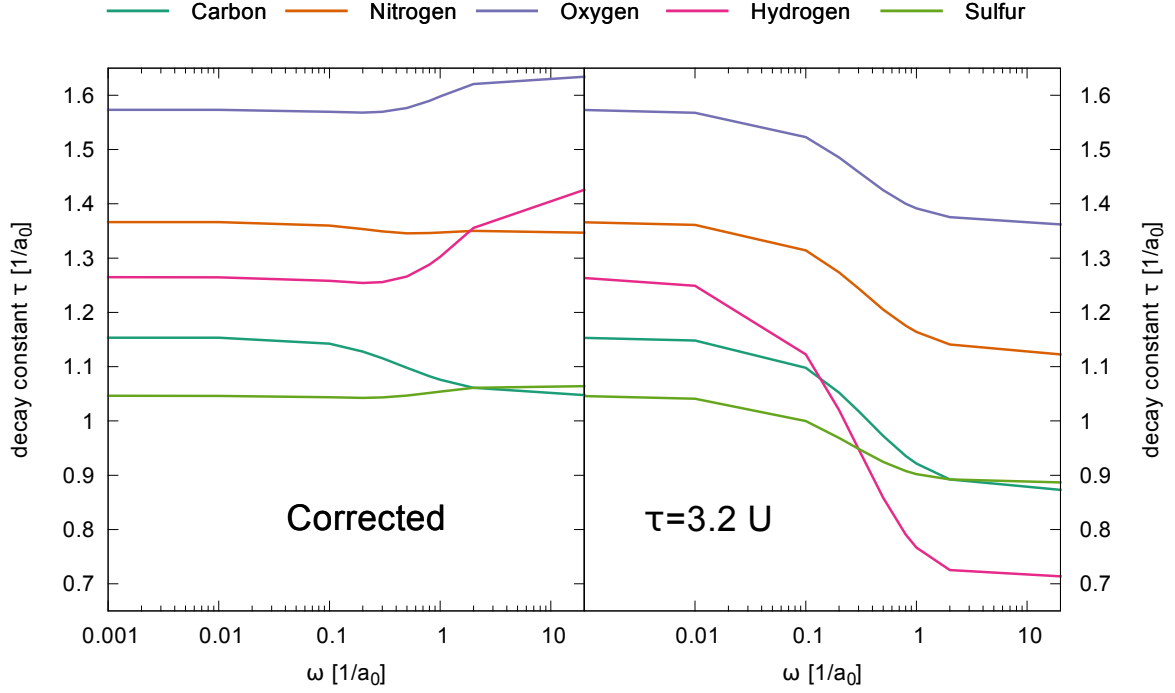


Figure 3.7: Decay constants obtained from condition  $\tau = 3.2U$  (right panel) and corrected decay constants  $\tau$  as used in LC-DFTB method, obtained by solving the Eq. 3.52 (left panel) for elements H,C,N,O,S as a function of range-separation parameter  $\omega$ .

Hubbard derivative as obtained from the LC-DFTB calculation on a single atom as a function of the decay constant  $\tau_A$  and the range-separation parameter  $\omega$

$$\begin{aligned}
 U^{\text{LC-DFTB}} &= \gamma_{AA}^{\text{fr}} - \frac{1}{2} \frac{1}{2l+1} \gamma_{AA}^{\text{lr}} \\
 &= \frac{5}{16} \tau \left[ 1 - \frac{1}{2(2l+1)} \left( 1 - \frac{\tau^8 + 3\tau^6\omega^2 - \tau^4\omega^4 + 0.2\omega^6\tau^2 - 3.2\tau^7\omega}{(\tau^2 - \omega^2)^4} \right) \right], \quad (3.52)
 \end{aligned}$$

where  $l$  is the angular momentum of the HOMO orbital of the considered atom. Back to the Fig. 3.6, we confirm the analytical formula Eq. 3.52 (orange line), where we use the same decay constants as for the numerical LC-DFTB calculation (gray dots).

With this analytical result we can now enforce the condition Eq. 3.39. For this, the decay constants have to be redefined. The Hubbard parameter  $U_A^{\text{LC-DFT}}$  from the atomic LC-DFT calculation for a given atom  $A$  and range-separation parameter  $\omega$  fixes via Eq. 3.52 (where  $U^{\text{LC-DFTB}} = U_A^{\text{LC-DFT}}$ ) the decay constant  $\tau_A$ . Solving this equation for  $\tau$  gives the corrected decay constants which are then used for the LC-DFTB parametrization. A possible numerical algorithm for the solution of the equation Eq. 3.52 is outlined in Appendix C. The decay constants and corrected decay constants are plotted in Fig. 3.7 for different elements as a function of range-separation parameter  $\omega$ .

### 3.8.2 Comment on the decay constant

We assumed that the decay constant for full-range and long-range  $\gamma$ -integrals for a given atomic species should be the same. The full-range  $\gamma$ -integral stems from the Hartree potential and the linearized xc-potential. However, it is approximated as the integral over the Coulomb interaction only. This is the same situation as in standard DFTB, where by imposing the condition  $\tau = 3.2U$ , the effective charge distribution  $F_A \sim e^{-\tau r}$  is in fact determined also by the xc-potential of the atomic DFT. With other words the effect of the xc-potential is included into the decay constant. Assuming the decay constant for long-range and full-range integrals to be equal and imposing the condition Eq. 3.39 to obtain the decay constant, we introduce the effects of the xc-potential into both full-range and long-range integrals. This does not correspond to the pure long-range corrected functional. Thus one can argue, that the effective charge distributions  $F_A$  for the full-range integral and the long-range integral should be different. The introduction of a second decay constant requires, however, an additional condition. Alternative scheme for example could be to evaluate the long-range  $\gamma$ -integral directly from averaged basis functions

$$\gamma^{\text{lr}} = \int \frac{1}{(l+1)^2} \sum_{\mu \in A} |\phi_{\mu}(\mathbf{r})|^2 \frac{1 - e^{-\omega|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} \frac{1}{(l+1)^2} \sum_{\mu \in A} |\phi_{\mu}(\mathbf{r}')|^2 d\mathbf{r} d\mathbf{r}' \quad (3.53)$$

and then using the formula Eq. 3.52 the full-range decay constant can be obtained from

$$\frac{5}{16} \tau^{\text{fr}} = U^{\text{LC-DFT}} + \frac{1}{2(2l+1)} \gamma^{\text{lr}}. \quad (3.54)$$

This again will ensure that the atomic LC-DFTB will give the correct Hubbard derivative.

## 3.9 RUN-TIME HAMILTONIAN EVALUATION

We found that the bottleneck of the test run-time LC-DFTB calculations for the systems with number of atoms  $< 1000$  is the evaluation of the exchange matrix (Eq. 3.55 below). The naive four loop algorithm is inefficient and does not reflect the factual quadratic scaling of the LC-DFTB method with an increase of system size. We have identified two ways of achieving the quadratic scaling. On the one hand we can use the neighbor list-based algorithms. On the other hand we can use the four loop method in combination with a cutoff condition, which allows to decide whether the given Hamiltonian sub-block has to be evaluated or whether the evaluation should be avoided. In following we present both approaches.

### 3.9.1 Direct SCF and thresholding algorithm

We search for inspiration in the techniques, successfully used in *ab initio* software packages. The implementation of a hybrid DFT-HF SCF method, based on a finite auxiliary basis expansion requires the evaluation of large numbers of one- and two-electron integrals and an efficient way to assemble the Hamiltonian from the density matrix and the respective matrix elements. The

evaluation of the two-electron integrals is computationally very expensive. In order to use the computational resources in an optimal way there are essentially two ways of organizing the calculation. On the one hand, the one- and two-electron integrals do not change during the single point calculation and can be thus precomputed. This approach is used in the atomic DFT code from the DFTB parametrization toolkit, which has been mentioned already. For large systems (hundreds of atoms) and/or large basis sets (10 basis functions per atom), however, this approach is not practicable due to the high memory requirements. An alternative approach is to use cutoff conditions, which allow to decide whether a sub-block of a Hamiltonian is negligible or should be taken into account. Because of this, the integrals are evaluated only if needed. This technique is known as direct SCF [4, 64].

In the LC-DFTB method, the integrals and Hamiltonian matrix elements are precomputed or parametrized in analytical form. The computational challenge is thus only in the construction of the Hamiltonian from these parameters. It turns out that the new exact exchange term, due to the way it couples to the density matrix complicates the construction of the LC-DFTB Hamiltonian as compared to the standard DFTB. While in the latter efficient sparse-matrix techniques, applied to Hartree-like terms lead to excellent performance [6], the evaluation of the exchange matrix in the LC-DFTB method

$$K_{\mu\nu}^{\text{LC-DFTB}} = -\frac{1}{8} \sum_{AB} \underbrace{(\gamma_{CB} + \gamma_{AB} + \gamma_{CD} + \gamma_{AD})}_{\Gamma_{ABCD}} \sum_{\alpha \in A} \sum_{\beta \in B} S_{\mu\alpha} S_{\beta\nu} \Delta P_{\alpha\beta} \quad (3.55)$$

is the bottleneck of calculation as observed in test calculations.

We notice that an atomic sub-block (C,D) of the exchange matrix is proportional to a product of overlap integrals between the basis functions at different centers. Extending the system (add more atoms), the number of non-vanishing overlap matrix elements will grow linearly, although the overall number of matrix elements grows quadratically with basis size. In Fig. 3.8 the number of numerically non-vanishing overlap matrix elements ( $S_{\mu\nu} > 10^{-16}$ ) as a function of basis size for polyacene oligomer geometries (chemical structural formula is sketched as inset) with number of monomer units  $n$  ranging from  $n = 1$  to  $n = 150$  is depicted. The scaling is clearly linear for large systems. For this reason we expect that the exchange matrix construction in the LC-DFTB should scale quadratically with the number of basis functions. In fact the thresholding algorithm, which will be presented in following and the neighbor list-based algorithm, presented in section 3.9.2, use this finding, however in different ways.

The direct SCF approaches use the linearity of the Hamiltonian with respect to the density matrix  $P_{\alpha\beta}$

$$H(P^n) = H(P^{n-1} + \underbrace{P^n - P^{n-1}}_{\Delta P^n}) = H(P^{n-1}) + H(\Delta P^n), \quad (3.56)$$

where  $P^n$  is the density matrix at the  $n$ -th SCF cycle. In the calculation, which approaches convergence the elements of the difference density matrix  $\Delta P^n$  get smaller and hence the

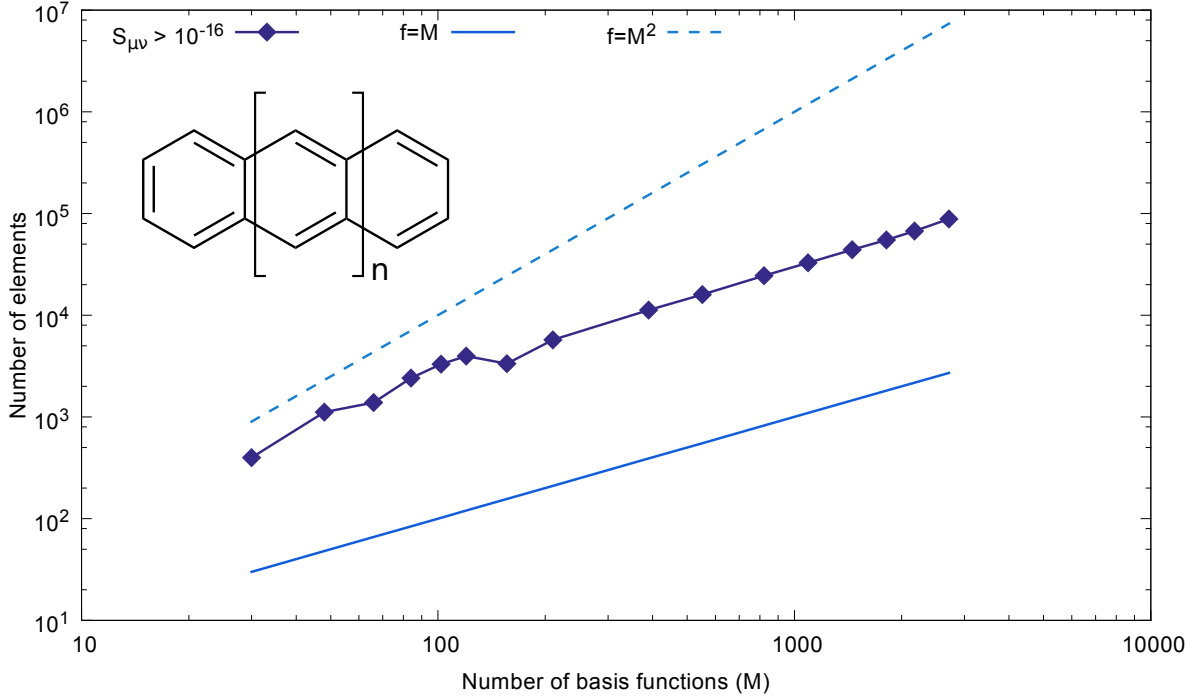


Figure 3.8: Number of non-vanishing overlap matrix elements as a function of basis size  $M$  for oligoacene geometries (inset) with number of monomer units  $n = 1, \dots, 150$  (diamonds). The functions  $f = M$  (solid line) and  $f = M^2$  (dashed line) serve as the guide to the eyes.

contribution to the Hamiltonian  $H(\Delta P^n)$  also gets smaller. At each SCF cycle the Hamiltonian sub-blocks are tested due to some thresholding condition. If the contribution of a sub-block of the Hamiltonian  $H(\Delta P^n)$  is estimated to be negligible it's evaluation is omitted. We adopt these ideas to the LC-DFTB method. First, we rewrite the LC-DFTB exchange matrix Eq. 3.55 in terms of quantities

$$Q_{AB}^{\mu\nu} = \sum_{\alpha \in A} \sum_{\beta \in B} S_{\mu\alpha} \Delta P_{\alpha\beta}^n S_{\beta\nu}, \quad (3.57)$$

which then yields ( $\mu \in C, \nu \in D$ )

$$K_{\mu\nu}^{\text{LC-DFTB}} = -\frac{1}{8} \sum_{AB} \Gamma_{ABCD} Q_{AB}^{\mu\nu}. \quad (3.58)$$

For each Hamiltonian sub-block  $(C, D)$  and atom pair  $(A, B)$ , we estimate the quantities  $Q_{AB}^{\mu\nu}$  as

$$Q_{AB}^{\mu\nu} \leq \sum_{\alpha \in A} \sum_{\beta \in B} |S_{\mu\alpha}| |\Delta P_{\alpha\beta}^n| |S_{\beta\nu}| \leq \tilde{s}_{BD} \cdot \tilde{s}_{AC} \cdot P_{\max}^n \sum_{\alpha \in A} \sum_{\beta \in B} \quad (3.59)$$

and decide whether to evaluate them according to the condition

$$\tilde{s}_{BD} \cdot \tilde{s}_{AC} \cdot P_{\max}^n \leq \epsilon_{\text{threshold}}, \quad (3.60)$$

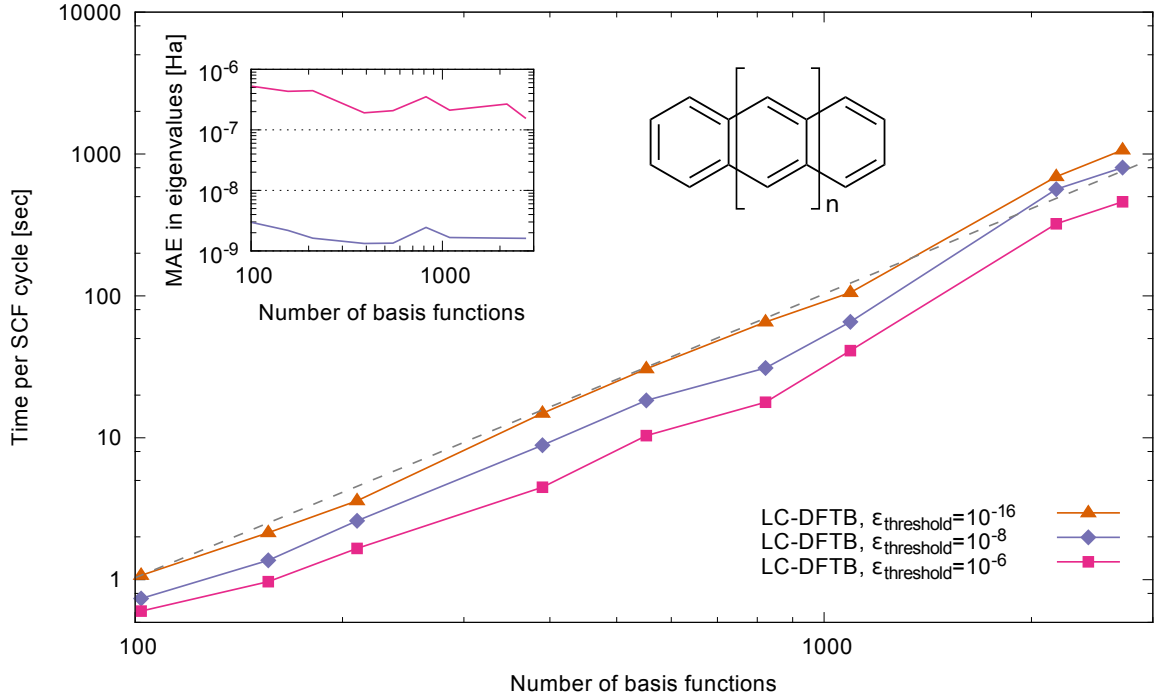


Figure 3.9: The average time per SCF cycle for the polyacene oligomers as a function of the number of basis functions for the threshold parameters  $\epsilon_{\text{threshold}} = 10^{-16}, 10^{-8}, 10^{-6}$ . The gray dashed line gives an extrapolation for quadratic scaling. The inset shows the mean absolute error in the eigenvalues in Hartree for  $\epsilon_{\text{threshold}} = 10^{-8}$  and  $\epsilon_{\text{threshold}} = 10^{-6}$ . The execution time was measured by the Linux *time* utility and the calculations have been performed on a single core of an Intel Core-i7 CPU.

where

$$\tilde{s}_{AB} = \max_{\alpha \in A, \beta \in B} (|S_{\alpha\beta}|), \quad P_{\max}^n = \max_{\alpha\beta} (|\Delta P_{\alpha\beta}^n|). \quad (3.61)$$

The quantities  $\tilde{s}_{AB}$  have to be evaluated only once at each geometry. The factor coming from the summation in Eq. 3.59 is absorbed in the threshold parameter  $\epsilon_{\text{threshold}}$ , which we obtain empirically by testing the error in eigenvalues for different values of  $\epsilon_{\text{threshold}}$ . The thresholding algorithm has been tested on a series of polyacene oligomers with number of monomer units ranging from  $n = 5$  to  $n = 150$ . In Fig. 3.9 the time per SCF cycle as a function of basis size is depicted. We use the threshold values  $\epsilon_{\text{threshold}} = 10^{-16}, 10^{-8}, 10^{-6}$ . In addition, we extrapolate the timing for the case  $\epsilon_{\text{threshold}} = 10^{-16}$ , assuming a quadratic dependence of the execution time per SCF cycle  $t(M)$  on the basis size  $M$

$$t(M) = cM^2, \quad (3.62)$$

where  $t(M_0)$  is the execution time for the smallest oligomer ( $M_0 = 102$ ). It can be clearly seen, that the algorithm shows quadratic scaling with the basis size. The cubic scaling due to the



diagonalization does not seem to show up for the systems tested here. However, it is expected for larger systems. Moreover, for  $\epsilon_{\text{threshold}} = 10^{-6}$  we find a reduction of computational cost of about 2 – 3 times by keeping the mean absolute error in the eigenvalues below  $10^{-6}$  Hartree (inset of Fig. 3.9). This suggests, that the value  $\epsilon_{\text{threshold}} = 10^{-6}$  can be considered an appropriate choice for practical calculations.

### 3.9.2 Neighbor list-based algorithm

The DFTB+ [6] code, which we use as a basis for the LC-DFTB extension, is implemented using the neighbor list technique. The concept of neighbor list can be introduced on the following simple example. Assume the quantity  $S_{ij} \in \mathbb{R}$ ,  $\forall (i, j) \in A = \{(k, l) | k, l \in [1, N]\}$  and  $S_{ij} = S_{ji}$ . This can be for instance the overlap integral matrix of orbitals with indices  $i, j$ . Further let's assume, that there exists a subset  $B \subset A$  of non-vanishing elements  $S_{ij}$ , such that  $B = \{(k, l) \in A | S_{kl} \neq 0\}$ . Our task is to evaluate the sum over all elements  $S_{ij}$ . The naive summation is obviously a  $N^2$  operation. This method does not take into account, that there are vanishing elements of  $S_{ij}$  which do not contribute to the sum. Especially if the number of vanishing elements is large this means a considerable computational overhead. The idea is to sum over the elements of the set  $B$  instead of  $A$ . Let us additionally define  $N(k) = \{l \in \mathbb{N} | (k, l) \in B, l \geq k\}$ ,  $\bar{N}(k) = \{(k, l) \in B | l > k\}$ . Then

$$\sum_{i=1}^N \sum_{j=1}^N S_{ij} = \sum_{i=1}^N \left[ \sum_{j < i} S_{ij} + \sum_{j > i} S_{ij} + S_{ii} \right] = 2 \sum_{i=1}^N \sum_{j > i} S_{ij} + \sum_{i=1}^N S_{ii} = 2 \sum_{i=1}^N \sum_{j \in \bar{N}(i)} S_{ij} + \sum_{i=1}^N S_{ii}, \quad (3.63)$$

where we used the property  $S_{ij} = S_{ji}$ . This reduces the scaling of the sum evaluation from quadratic to linear (although with possibly large prefactor). To see this, let us assume that the density of atoms (number of atoms per volume) in a typical physical system is  $D$ . Moreover, let us assume, that the overlap integrals for two arbitrary atoms vanish for some common cutoff distance  $R$ . Then, given an atom, the number of neighbors of this atom is constant, regardless of the system size, since the number of atoms inside the sphere around this atom with radius  $R$  will usually not exceed  $\frac{4}{3}\pi R^3 D$ .

From this consideration we can conclude that evaluation of quantities, involving the double sum over the overlap matrix will scale quadratically. This is the case for the evaluation of the exchange matrix. We want to apply the neighbor list technique to the exchange matrix evaluation problem. Since the overlap matrix is symmetric, the neighbor list is defined in the way, such that the summation as in example above is the triangle sum. While the evaluation of the Hartree term (second order term) in the DFTB Hamiltonian is straightforward, the HF exchange term has a different structure and a new formula has to be derived. Note, that the Hamiltonian evaluation is actually a summation. The idea is to run two sums over all basis elements and two sums over the neighbors of each of the variables in the full sums such that one index of the overlap matrix belongs to the full sum and the second to the neighbor sum. For this we have to rearrange the

four-loop sum over all basis elements in the following way

$$\begin{aligned} \sum_{\mu\nu} \sum_{\alpha\beta} \Delta P_{\alpha\beta} S_{\mu\alpha} S_{\beta\nu} \Gamma_{\mu\alpha\beta\nu}^{\text{lr}} &= \sum_{\mu \leq \alpha} \sum_{\nu \leq \beta} \Delta P_{\alpha\beta} S_{\mu\alpha} S_{\beta\nu} \Gamma_{\mu\alpha\beta\nu}^{\text{lr}} + \sum_{\mu \leq \alpha} \sum_{\beta < \nu} \Delta P_{\alpha\beta} S_{\mu\alpha} S_{\beta\nu} \Gamma_{\mu\alpha\beta\nu}^{\text{lr}} \\ &+ \sum_{\alpha < \mu} \sum_{\nu \leq \beta} \Delta P_{\alpha\beta} S_{\mu\alpha} S_{\beta\nu} \Gamma_{\mu\alpha\beta\nu}^{\text{lr}} + \sum_{\alpha < \mu} \sum_{\beta < \nu} \Delta P_{\alpha\beta} S_{\mu\alpha} S_{\beta\nu} \Gamma_{\mu\alpha\beta\nu}^{\text{lr}} \end{aligned} \quad (3.64)$$

$$\begin{aligned} &= \sum_{\mu \leq \alpha} \sum_{\nu \leq \beta} \Delta P_{\alpha\beta} S_{\mu\alpha} S_{\beta\nu} \Gamma_{\mu\alpha\beta\nu}^{\text{lr}} + \sum_{\mu \leq \alpha} \sum_{\nu < \beta} \Delta P_{\alpha\nu} S_{\mu\alpha} S_{\nu\beta} \Gamma_{\mu\alpha\nu\beta}^{\text{lr}} \\ &+ \sum_{\mu < \alpha} \sum_{\nu \leq \beta} \Delta P_{\mu\beta} S_{\alpha\mu} S_{\beta\nu} \Gamma_{\alpha\mu\beta\nu}^{\text{lr}} + \sum_{\mu < \alpha} \sum_{\nu < \beta} \Delta P_{\mu\nu} S_{\alpha\mu} S_{\nu\beta} \Gamma_{\alpha\mu\nu\beta}^{\text{lr}}, \end{aligned} \quad (3.65)$$

where  $\Gamma_{\mu\alpha\beta\nu}^{\text{lr}} = (\gamma_{\mu\beta}^{\text{lr}} + \gamma_{\alpha\beta}^{\text{lr}} + \gamma_{\mu\nu}^{\text{lr}} + \gamma_{\alpha\nu}^{\text{lr}})$ . Thus we can now express the four loop sum by the two-loop sum and two sums over the neighbors

$$\begin{aligned} \sum_{\mu\nu} \sum_{\alpha\beta} \Delta P_{\alpha\beta} S_{\mu\alpha} S_{\beta\nu} \Gamma_{\mu\alpha\beta\nu}^{\text{lr}} &= \sum_{\mu\nu} \left[ \sum_{\alpha \in N(\mu)} \sum_{\beta \in N(\nu)} \left( \Delta P_{\alpha\beta} S_{\mu\alpha} S_{\beta\nu} \Gamma_{\mu\alpha\beta\nu}^{\text{lr}} \rightarrow H_{\mu\nu} \right) \right. \\ &+ \sum_{\alpha \in N(\mu)} \sum_{\beta \in N(\nu)} \left( \Delta P_{\alpha\nu} S_{\mu\alpha} S_{\nu\beta} \Gamma_{\mu\alpha\nu\beta}^{\text{lr}} \rightarrow H_{\mu\beta} \right) \\ &+ \sum_{\alpha \in N(\mu)} \sum_{\beta \in N(\nu)} \left( \Delta P_{\mu\beta} S_{\alpha\mu} S_{\beta\nu} \Gamma_{\alpha\mu\beta\nu}^{\text{lr}} \rightarrow H_{\alpha\nu} \right) \\ &\left. + \sum_{\alpha \in N(\mu)} \sum_{\beta \in N(\nu)} \left( \Delta P_{\mu\nu} S_{\alpha\mu} S_{\nu\beta} \Gamma_{\alpha\mu\nu\beta}^{\text{lr}} \rightarrow H_{\alpha\beta} \right) \right]. \end{aligned} \quad (3.66)$$

Here, the operation  $\rightarrow H_{\mu\nu}$  means that each summand from the sub-sum over  $\alpha, \beta$  in that line is added to the contents of the Hamiltonian matrix element  $H_{\mu\nu}$ . We test the scaling of this algorithm on a set of polyacene oligomers as it was done for the case of thresholding algorithm. We confirm again the quadratic scaling. The average time per SCF cycle is plotted in Fig. 3.10 (triangles) and compared to the thresholding algorithm with  $\epsilon_{\text{threshold}} = 10^{-16}$  (diamonds) and  $\epsilon_{\text{threshold}} = 10^{-6}$  (squares). We find comparable performance of the neighbor list-based algorithm and thresholding algorithm with  $\epsilon_{\text{threshold}} = 10^{-16}$ , while for  $\epsilon_{\text{threshold}} = 10^{-6}$  the latter clearly outperforms the neighbor list-based algorithm. As has been shown, the thresholding algorithm uses in addition to the sparsity of the overlap matrix also the fact, that if the calculation approaches convergence the difference between the density matrices at subsequent iterations is getting smaller. It should be noted, that the combination of both approaches is in principle possible.

### 3.10 EVALUATION OF ENERGY GRADIENTS

In the DFT and DFTB the total energy of a  $N$ -electron system is treated within the Born-Oppenheimer approximation. The electronic motion has been separated from the motion of nuclei, which are assumed to be fixed. Thus the total energy depends parametrically on the set of nuclear coordinates. With the SCF procedure the total energy is minimized for a given geometric configuration. This procedure is called single point calculation. Of course the minimum

of the total energy with respect to a given geometric configuration is in general not the global minimum. To obtain the geometric configuration, which globally minimizes the total energy it is thus necessary to perform the two-step minimization process. The total energy is minimized in the space of the geometric configurations. For each geometric configuration the total energy is minimized in a single point calculation. To perform such a minimization process the matrix of second derivatives of the total energy with respect to all nuclear coordinates (Hessian) is required. There are, however, methods which do not require the explicit evaluation of the Hessian. It is guessed from the matrix of first derivatives, which are called the energy gradients or forces. The steepest descent and conjugate gradient methods are the examples of such methods. In the DFTB+ code [6], which we use as the basis for the extension of DFTB to LC-DFTB the steepest descent and conjugate gradient algorithms are already implemented. Thus we will discuss only the evaluation of energy gradients.

In following we derive the formulas for the energy gradients due to the additional long-range HF exchange term and describe the algorithm for the efficient evaluation of these contributions. The force component  $F_k$  due to the change in the  $k$ -th nuclear coordinate reads

$$F_k = \frac{dL}{dR_k} = \sum_{\mu i} \frac{\partial L}{\partial c_{\mu i}} \frac{\partial c_{\mu i}}{\partial R_k} + \frac{\partial L}{\partial R_k}, \quad (3.67)$$

where the functional  $L = L(c_{\mu i}(R_k), R_k)$  is defined as

$$L = E^{\text{tot}} - \sum_i n_i \epsilon_i \sum_{\mu\nu} (c_{\mu i} S_{\mu\nu} c_{\nu i} - 1). \quad (3.68)$$

Since on SCF convergence the condition

$$\frac{\partial L}{\partial c_{\mu i}} = 0, \quad \forall c_{\mu i} \quad (3.69)$$

is fulfilled, only explicit dependence of the total energy  $E^{\text{tot}}$  on nuclear coordinates should be considered. This means, that the coefficients  $c_{\mu i}$  and thus the density matrix don't have to be differentiated (same argument as for standard DFTB [41]). All total energy contributions in the LC-DFTB can be treated in the same way as in the standard DFTB [41] with exception of long-range HF exchange term. The contribution to forces, due to this new term will be derived in following. First, we reformulate the expression for the long-range HF exchange energy by resummation

$$E^{\text{lr}} = -\frac{1}{16} \sum_{CD} \sum_{AB} \sum_{\alpha \in A} \sum_{\beta \in B} \sum_{\mu \in C} \sum_{\nu \in D} \Delta P_{\mu\nu} \Delta P_{\alpha\beta} S_{\mu\alpha} S_{\beta\nu} (\gamma_{CB}^{\text{lr}} + \gamma_{CD}^{\text{lr}} + \gamma_{AB}^{\text{lr}} + \gamma_{AD}^{\text{lr}}) \quad (3.70)$$

$$= -\frac{1}{8} \sum_{CD} \sum_{AB} \sum_{\alpha \in A} \sum_{\beta \in B} \sum_{\mu \in C} \sum_{\nu \in D} \gamma_{CD}^{\text{lr}} (\Delta P_{\alpha\beta} \Delta P_{\mu\nu} + \Delta P_{\mu\beta} \Delta P_{\alpha\nu}) S_{\mu\alpha} S_{\nu\beta}. \quad (3.71)$$

We perform now the derivative with respect to the coordinate  $R_K = (R_x, R_y, R_z)_K$  of the atom  $K$ . It is clear, that only those overlap matrix elements and  $\gamma$ -integrals contribute to the derivative,

which depend on the atom  $K$  and some other atom  $A \neq K$ . The task is thus to sort out the terms which vanish on doing the derivative. We introduce the differential operator  $D_K$ , such that  $D_K E = F_K = \nabla_K E$ , then

$$F_K = -\frac{1}{8} \sum_{CD} \sum_{AB} \sum_{\alpha \in A} \sum_{\beta \in B} \sum_{\mu \in C} \sum_{\nu \in D} (\Delta P_{\mu\nu} \Delta P_{\alpha\beta} + \Delta P_{\mu\beta} \Delta P_{\alpha\nu}) D_K (S_{\mu\alpha} S_{\beta\nu} \gamma_{CD}^{\text{lr}}) \quad (3.72)$$

$$\begin{aligned} &= -\frac{1}{8} \sum_{CD} \sum_{AB} \sum_{\alpha \in A} \sum_{\beta \in B} \sum_{\mu \in C} \sum_{\nu \in D} (\Delta P_{\mu\nu} \Delta P_{\alpha\beta} + \Delta P_{\mu\beta} \Delta P_{\alpha\nu}) \times \\ &\times \left[ S_{\beta\nu} \gamma_{CD}^{\text{lr}} \begin{cases} D_K S_{\mu\alpha} & K = C \text{ or } K = A \text{ and } C \neq A \\ 0 & \text{else} \end{cases} \right. \\ &+ S_{\mu\alpha} \gamma_{CD}^{\text{lr}} \begin{cases} D_K S_{\beta\nu} & K = B \text{ or } K = D \text{ and } B \neq D \\ 0 & \text{else} \end{cases} \\ &\left. + S_{\mu\alpha} S_{\beta\nu} \begin{cases} D_K \gamma_{CD}^{\text{lr}} & K = C \text{ or } K = D \text{ and } C \neq D \\ 0 & \text{else} \end{cases} \right] \quad (3.73) \end{aligned}$$

$$\begin{aligned} &= -\frac{1}{8} \sum_C \sum_{A \neq C} \sum_{BD} \sum_{\alpha \in A} \sum_{\beta \in B} \sum_{\mu \in C} \sum_{\nu \in D} \gamma_{CD}^{\text{lr}} S_{\beta\nu} (D_K S_{\mu\alpha}) [\delta_{KC} + \delta_{KA}] (\Delta P_{\mu\nu} \Delta P_{\alpha\beta} + \Delta P_{\mu\beta} \Delta P_{\alpha\nu}) \\ &- \frac{1}{8} \sum_D \sum_{B \neq D} \sum_{AC} \sum_{\alpha \in A} \sum_{\beta \in B} \sum_{\mu \in C} \sum_{\nu \in D} \gamma_{CD}^{\text{lr}} S_{\mu\alpha} (D_K S_{\beta\nu}) [\delta_{KB} + \delta_{KD}] (\Delta P_{\mu\nu} \Delta P_{\alpha\beta} + \Delta P_{\mu\beta} \Delta P_{\alpha\nu}) \\ &- \frac{1}{8} \sum_D \sum_{C \neq D} \sum_{AB} \sum_{\alpha \in A} \sum_{\beta \in B} \sum_{\mu \in C} \sum_{\nu \in D} S_{\beta\nu} S_{\mu\alpha} (D_K \gamma_{CD}^{\text{lr}}) [\delta_{KC} + \delta_{KD}] (\Delta P_{\mu\nu} \Delta P_{\alpha\beta} + \Delta P_{\mu\beta} \Delta P_{\alpha\nu}) \quad (3.74) \end{aligned}$$

$$\begin{aligned} &= -\frac{1}{4} \sum_{C \neq K} \sum_{AB} \sum_{\alpha \in A} \sum_{\beta \in B} \sum_{\mu \in C} \sum_{\kappa \in K} \left( S_{\alpha\beta} (D_K S_{\kappa\mu}) (\Delta P_{\kappa\beta} \Delta P_{\mu\alpha} + \Delta P_{\kappa\alpha} \Delta P_{\mu\beta}) (\gamma_{KB}^{\text{lr}} + \gamma_{CB}^{\text{lr}}) \right. \\ &\left. + (D_K \gamma_{KC}^{\text{lr}}) S_{\kappa\beta} S_{\alpha\mu} (\Delta P_{\kappa\mu} \Delta P_{\beta\alpha} + \Delta P_{\kappa\alpha} \Delta P_{\beta\mu}) \right). \quad (3.75) \end{aligned}$$

This expression is written in the form, which is convenient to formulate the neighbor list-based algorithm. The implementation of the LC-DFTB in the DFTB+ code uses the neighbor list algorithm for the evaluation of the contributions to the force due to the long-range HF exchange part. The sketch of the algorithm is provided in appendix E.

We test the implementation of the algorithm on series of oligoacenes with number of monomer units ranging from  $n = 5$  to  $n = 150$ . The evaluation time for the single force matrix (energy gradients for all atoms) as a function of basis size is plotted in Fig. 3.10. Additionally we plot the time per SCF cycle for the neighbor list-based algorithm for the Hamiltonian evaluation (section 3.9.2) and the thresholding algorithm (section 3.9.1). We observe the quadratic scaling for the force matrix evaluation. We find that it is slower than the average SCF iteration from the neighbor list-based algorithm by the factor 1–2. Further tests show that the convergence behavior of a geometry optimization and the accuracy of forces of the present LC-DFTB implementation are comparable to that of the standard DFTB.

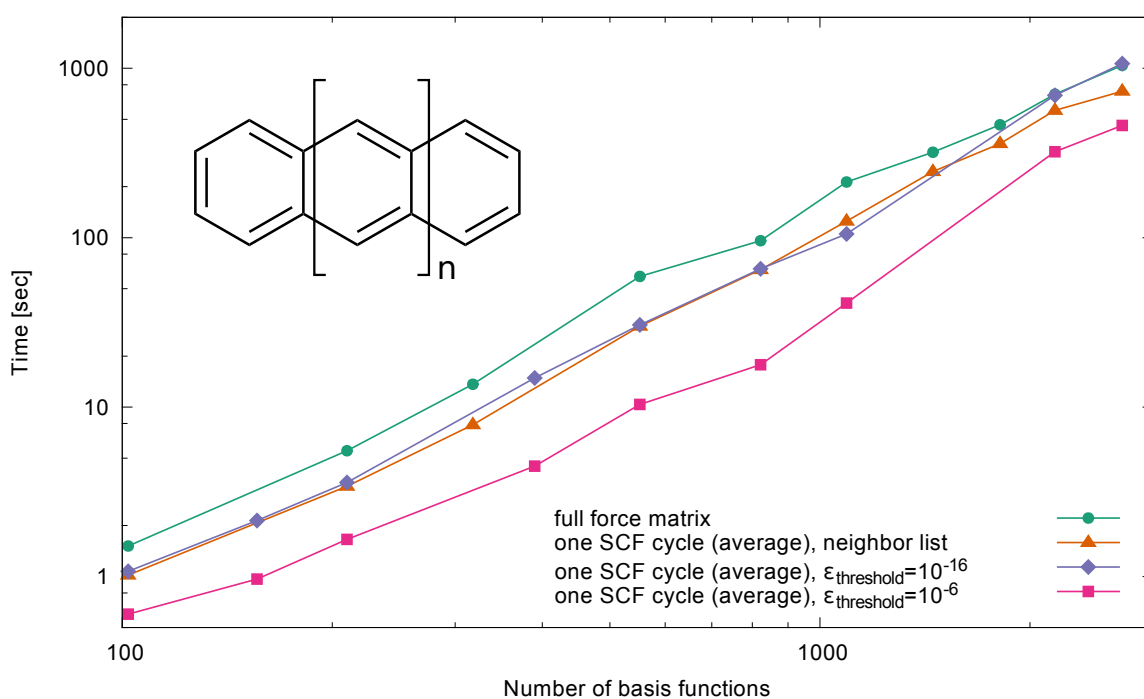


Figure 3.10: The force calculation time compared to the average time per SCF iteration for neighbor list-based algorithm and thresholding algorithm with  $\epsilon_{\text{threshold}} = 10^{-16}$  and  $\epsilon_{\text{threshold}} = 10^{-6}$ . The test geometries are that of the polyacene oligomers (inset) with number of monomer units  $n = 5$  to  $n = 150$ .



## EVALUATION OF TWO-ELECTRON INTEGRALS

Evaluation of two-electron integrals is one of the most problematic parts of practical DFT or Hartree-Fock calculations. It is no surprise, that historically the algorithms for *ab initio* calculations have been designed according to the available integral evaluation methods. The wide use of the Gaussian-type orbitals (GTO) in the majority of implementations is due to the existence of efficient analytical procedures for the evaluation of two-electron integrals over the GTO. In contrast to the physical Slater-type orbitals (STO), GTO are thus a mathematical compromise. Tremendous work have been done on optimization and design of efficient GTO basis sets, which now together with plane-wave based methods dominate in the practical all-electron calculations. However, GTO basis sets, suitable for publication quality calculations, go beyond minimal basis and at least double zeta size is usually recommended. Since STO have correct physical behavior, the minimal STO basis outperforms the minimal GTO. For this reason minimal STO basis sets have been the first choice in the early semi-empirical methods, like CNDO or INDO [153, 154]. The parametrization tools of standard DFTB method and in this way the DFTB itself relies on the STO basis as well. Since the extension to the long-range corrected functionals requires the evaluation of two-electron integrals over Yukawa interaction, the efficient integration routines should be available. This chapter is dedicated to the evaluation of two-electron integrals. In section 4.1 we motivate the choice of the particular integration scheme for the extension of the LC-DFTB method. The general description of the integration method is given in section 4.2. Specific routines for one-center integrals, which are required in the atomic DFT code, which is a part of the parametrization toolkit are discussed in section 4.3. The evaluation method for the two-center integrals, required for the modification of the two-center code from the parametrization toolkit is presented in section 4.4.

#### 4.1 CHOICE OF THE INTEGRATOR

The right choice of an integration routine is crucial to the method. As has been stated before, the standard DFTB parametrization relies on the atomic DFT code and the two-center code, which generates the zeroth-order Hamiltonian and overlap matrix elements. The extension to the long-range corrected functionals requires the evaluation of the non-local long-range HF exchange term within the two-center code and the evaluation of one-center exchange integrals over Yukawa interaction for the atomic DFT code.

We choose the integration method, proposed by Becke et. al [13, 18], which fits very well to our purposes. First, it is fully numerical. This goes inline with the essence of the DFTB method, which relies on the set of converged pseudo-atomic orbitals as a basis, independent of their representation. This allows for more flexibility in the choice of the basis. The numerical orbitals tabulated on some grid, or finite basis set expansion as the representation method are treated on equal footing. Furthermore, the Becke integration method allows to use essentially the same procedure for the integrals over Coulomb and Yukawa interactions. Only minor changes are necessary. Finally, the method is robust, easy to implement and test. Thus it provides an all in one solution for the parametrization machinery of the LC-DFTB method, described in this thesis.

1

#### 4.2 DESCRIPTION OF THE SCHEME AND EXTENSION TO YUKAWA INTERACTION

In the section 2.4 the total energy expression and Hamiltonian of a general hybrid Hartree-Fock-DFT theory has been expressed in terms of matrices, which elements are the one- and two-electron integrals ( Eqn. 2.23, 2.24 and 2.25 ). The problem of the evaluation of two-electron integrals over Slater-type orbitals or numerical orbitals  $\phi_\mu$  for the Coulomb  $\omega = 0$  and Yukawa  $\omega \neq 0$  interaction arises

$$(\mu\nu|\alpha\beta)^{\text{lr}} = \int \int \phi_\mu(\mathbf{r})\phi_\nu(\mathbf{r}) \frac{\exp(-\omega|\mathbf{r}-\mathbf{r}'|)}{|\mathbf{r}-\mathbf{r}'|} \phi_\alpha(\mathbf{r}')\phi_\beta(\mathbf{r}') d\mathbf{r}d\mathbf{r}'. \quad (4.1)$$

This is the six dimensional integral with orbitals  $\phi_\mu$ , which in general are located at four different centers. The Becke method is formulated for the integral over the general charge distributions  $\rho(\mathbf{r}), \sigma(\mathbf{r})$ , rather than atom-centered orbitals

$$I = \int \int \rho(\mathbf{r})\sigma(\mathbf{r}')g^\omega(\mathbf{r},\mathbf{r}') d\mathbf{r}d\mathbf{r}' = \int \rho(\mathbf{r}) \left[ \int \sigma(\mathbf{r}')g^\omega(\mathbf{r},\mathbf{r}') d\mathbf{r}' \right] d\mathbf{r}, \quad (4.2)$$

where  $g^\omega(\mathbf{r},\mathbf{r}')$  is some interaction operator in space representation, which in our case is Coulomb or Yukawa operator. The inner integral of Eq. 4.2 can be seen as a potential  $V(\mathbf{r})$ , generated by the charge distribution  $\sigma(\mathbf{r})$  and the interaction  $g^\omega(\mathbf{r},\mathbf{r}')$

$$V(\mathbf{r}) = \int \sigma(\mathbf{r}')g^\omega(\mathbf{r},\mathbf{r}') d\mathbf{r}'. \quad (4.3)$$

---

<sup>1</sup> The choice of the particular parametrization components, however, is not of crucial importance. The same results can be in principle achieved with more elaborate analytical integrators as for example those, described in ref. [177].



The problem of the direct evaluation of this integral can be transformed into an equivalent problem of solving a boundary value problem with some differential operator  $\mathcal{D}$  and the boundary condition  $\lim_{|\mathbf{r}|\rightarrow\infty} V(\mathbf{r}) = 0$ , such that

$$\mathcal{D}V(\mathbf{r}) = -4\pi\sigma(\mathbf{r}). \quad (4.4)$$

The interaction operator  $g^\omega(\mathbf{r}, \mathbf{r}')$  in Eq. 4.3 is the Greens function of the problem and satisfies

$$\mathcal{D}g^\omega(\mathbf{r}, \mathbf{r}') = -4\pi\delta(\mathbf{r} - \mathbf{r}'). \quad (4.5)$$

For the known interaction operator (Greens function) with a given boundary condition (potential vanishes at infinity) we need to find a corresponding differential operator  $\mathcal{D}$ . For the case of Coulomb interaction operator  $g^0(\mathbf{r}, \mathbf{r}') = |\mathbf{r} - \mathbf{r}'|^{-1}$  the Poisson equation

$$\nabla^2 V(\mathbf{r}) = -4\pi\sigma(\mathbf{r}) \quad (4.6)$$

for the potential vanishing at infinity has to be solved. The original paper deals with this case. The extension to the integral over the Yukawa interaction, using this reasoning is then straightforward.<sup>2</sup> The Yukawa interaction operator  $g^\omega(\mathbf{r}, \mathbf{r}') = \exp(-\omega|\mathbf{r} - \mathbf{r}'|)|\mathbf{r} - \mathbf{r}'|^{-1}$  is the Greens function of the modified Helmholtz equation

$$(\nabla^2 - \omega^2)V(\mathbf{r}) = -4\pi\sigma(\mathbf{r}) \quad (4.7)$$

with potential vanishing at infinity.

The numerical solution of Eq. 4.7 for a given charge distribution (integrand)  $\sigma(\mathbf{r})$  should be performed in the real space numerically. For this reason we have to discretize the equation on a grid, which is “good enough” to obtain the solution with desired precision and with as small amount of operations as possible. We need thus to find an optimal distribution of grid points, such that the integrand and the potential are represented in the best possible way with as less grid points as possible. The typical molecular charge distribution and the corresponding potential are atom-like and their main features are mainly changing around the nuclei. They decay fast and monotonous to zero if far away from nuclei. Thus it is common to use a set of atom-centered grids, with an optimal grid point distribution (quadrature). To avoid multiple counting of the integrand contributions due to the use of overlapping grids, the integration space has to be partitioned into the atom-centered cells by some set of partitioning functions  $f_A(\mathbf{r})$

$$\rho(\mathbf{r}) = \sum_A^{\text{atoms}} f_A(\mathbf{r})\rho(\mathbf{r}) = \sum_A^{\text{atoms}} \rho_A(\mathbf{r}), \quad \sum_A^{\text{atoms}} f_A(\mathbf{r}) = 1, \quad \forall \mathbf{r} \in \mathbb{R}^3. \quad (4.8)$$

<sup>2</sup>The range separation with error-function leads to the interaction operator, which could not be easily identified as a Greens function of some boundary value problem. This is the reason for the choice of the range separation of Yukawa-type, which is also common in the literature.

In the choice of partitioning functions we follow the suggestion of ref. [14]. This reduces the integral over the general charge distributions, which can extend over many centers, to a sum of two-center integrals

$$I = \sum_{AB}^{\text{atoms}} I_{AB} = \sum_{AB}^{\text{atoms}} \int \rho_B(\mathbf{r}) \int \sigma_A(\mathbf{r}') g^\omega(\mathbf{r}, \mathbf{r}') d\mathbf{r}' d\mathbf{r} = \sum_{AB}^{\text{atoms}} \int \rho_B(\mathbf{r}) V_A(\mathbf{r}) d\mathbf{r}. \quad (4.9)$$

The inner integral is evaluated on a given one-center grid. The outer integral is then an overlap integral  $\int \rho_B V_A d\mathbf{r}$ , evaluated on the two-center grid over the potential  $V_A$  and the remaining charge distribution  $\rho_B$ .

The evaluation of the inner integral on the one-center grid is done by a solution of the Poisson or Helmholtz equation. The potential and the corresponding charge distribution are expanded into spherical harmonics  $Y_{lm}(\Omega)$  on a given center  $A$

$$V_A(\mathbf{r}) = \sum_{lm} \frac{1}{r} \left( \int V_A(r, \Omega') Y_{lm}(\Omega') d\Omega' \right) Y_{lm}(\Omega) = \sum_{lm} \frac{1}{r} V_{A,lm}(r) Y_{lm}(\Omega) \quad (4.10)$$

$$\sigma_A(\mathbf{r}) = \sum_{lm} \left( \int \sigma_{A,lm}(r, \Omega') Y_{lm}(\Omega') d\Omega' \right) Y_{lm}(\Omega) = \sum_{lm} \sigma_{A,lm}(r) Y_{lm}(\Omega), \quad (4.11)$$

where  $r = |\mathbf{r}|$  and the integration over solid angle ( $d\Omega'$ ) is done using Lebedev grids [108]. Inserting these expansions into Eq. 4.6 or Eq. 4.7 and comparing the components of the expansion (spherical harmonics constitute a basis) on the left-hand side and the right-hand side, we obtain the radial part of the differential equation

$$\left[ \frac{d^2}{dr^2} - \left( \frac{l(l+1)}{r^2} + \omega^2 \right) \right] V_{A,lm}(r) = -4\pi r \sigma_{A,lm}(r) \quad (4.12)$$

with following boundary conditions. In the case  $\omega \neq 0$  (Yukawa interaction),  $V_{A,lm}(0) = 0$ ,  $V_{A,lm}(r \rightarrow \infty) = 0$ . In the case  $\omega = 0$  (Coulomb interaction),  $V_{A,lm}(0) = 0$ ,  $V_{A,lm}(r \rightarrow \infty) = 0$ , if  $l > 0$  and  $V_{A,lm}(r \rightarrow \infty) = \sqrt{4\pi} q_A$ , if  $l = 0$ . Here  $q_A = \int \rho_A(\mathbf{r}) d\mathbf{r}$  is the total charge on a single atom  $A$ . Thus we need to solve at the moment infinite number of two-point boundary value problems. However, the expansions in Eqn. 4.10 and 4.11 can be truncated and the recommended cutoff angular momenta  $l_{\max}$  are connected to the size of Lebedev grids for the angular integrations [14]. For the discretization of Eq. 4.12 the variable  $0 \leq r < \infty$  is mapped to the interval  $-1 \leq x \leq 1$  using  $r = r_m(1+x)/(1-x)$ , where  $r_m$  is the atomic size parameter, which can be used to further optimize the scheme. Then the new variable  $x$  is discretized due to the Gauss-Chebyshev quadrature points  $x_i = \cos(\pi z_i)$  and  $z_i = i/(N+1)$  for  $N$  radial points. The so obtained radial grid points  $r_i = r_m(1 + \cos(\pi z_i))/(1 - \cos(\pi z_i))$  are non-equidistantly distributed. To use the method of finite differences we transform the equation Eq. 4.12 from the real domain  $r_i$  to the computational domain  $z_i$  with equidistant point distribution. This allows to use straightforward finite differences methods for the solution of the boundary value problem and simple cubic spline interpolation for

the evaluation of the outer integral [14]. Given the mapping  $r \rightarrow z(r)$

$$r = r_m \frac{1 + \cos(\pi z)}{1 - \cos(\pi z)} \quad (4.13)$$

$$z = \frac{1}{\pi} \arccos \left( \frac{r - r_m}{r + r_m} \right), \quad (4.14)$$

we use the chain and product rules of differential calculus and obtain the transformed equation

$$\left[ \frac{d^2}{dz^2} + \beta(z) \frac{d}{dz} + \gamma_l^\omega(z) \right] V_{A,lm}(z) + \delta(z) \sigma_{A,lm}(z) = 0, \quad (4.15)$$

where the prefactors are defined as

$$\beta(z) = \left( \frac{\pi}{\sin(\pi z)} + \frac{\pi}{2} \frac{\sin(\pi z)}{\sin^2\left(\frac{\pi}{2}z\right)} \right) \quad (4.16)$$

$$\gamma_l^\omega(z) = - \left( \frac{4\pi^2 l(l+1)}{\sin^2(\pi z)} + \frac{\omega^2 \pi^2 r_m^2 \sin^2(\pi z)}{4 \sin^8\left(\frac{\pi}{2}z\right)} \right) \quad (4.17)$$

$$\delta(z) = 4\pi^3 r_m^3 \frac{\cos^4\left(\frac{\pi}{2}z\right)}{\sin^8\left(\frac{\pi}{2}z\right)}. \quad (4.18)$$

With these definitions we discretize the equation using 7-point finite difference scheme [19]

$$y'_i = \frac{1}{12h} (-y_{i-3} + 6y_{i-2} - 18y_{i-1} + 10y_i + 3y_{i+1}) \quad (4.19)$$

$$y'_i = \frac{1}{12h} (-3y_{i-1} - 10y_i + 18y_{i+1} - 6y_{i+2} + y_{i+3}) \quad (4.20)$$

$$y'_i = \frac{1}{60h} (3y_{i-2} - 30y_{i-1} - 20y_i + 60y_{i+1} - 15y_{i+2} + 2y_{i+3}) \quad (4.21)$$

$$y'_i = \frac{1}{60h} (-2y_{i-3} + 15y_{i-2} - 60y_{i-1} + 20y_i + 30y_{i+1} - 3y_{i+2}) \quad (4.22)$$

$$y'_i = \frac{1}{60h} (-y_{i-3} + 9y_{i-2} - 45y_{i-1} + 45y_{i+1} - 9y_{i+2} + y_{i+3}) \quad (4.23)$$

$$y''_i = \frac{1}{12h^2} (11y_{i-1} - 20y_i + 6y_{i+1} + 4y_{i+2} - y_{i+3}) \quad (4.24)$$

$$y''_i = \frac{1}{12h^2} (-y_{i-3} + 4y_{i-2} + 6y_{i-1} - 20y_i + 11y_{i+1}) \quad (4.25)$$

$$y''_i = \frac{1}{12h^2} (-y_{i-2} + 16y_{i-1} - 30y_i + 16y_{i+1} - y_{i+2}) \quad (4.26)$$

$$y''_i = \frac{1}{180h^2} (2y_{i-3} - 27y_{i-2} + 270y_{i-1} - 490y_i + 270y_{i+1} - 27y_{i+2} + 2y_{i+3}) \quad (4.27)$$

and obtain a 7-band matrix. Note, that different forms of first and second derivatives are necessary to connect to the boundary. The resulting system of linear equations can be solved with standard numerical libraries (for example LAPACK [5]). In this way we obtain the  $lm$ -components  $V_{A,lm}(r_i)$  on the discrete grid. The full potential is then constructed from  $V_{A,lm}(r_i)$  using Eq. 4.10.

To carry out the remaining outer integral the values of the just obtained potential  $V_A$  on the center  $B$  are required. The angular part is known analytically through the spherical

$F_A$	$F_B$	$R_{AB} [a_0]$	Harris	$40 \times 110 \times 8$	$60 \times 194 \times 11$	$80 \times 302 \times 14$
$\zeta=2.0, 2p_z$	$\zeta=1.0, 1s$	0.6	0.1988 5344	0.1988 5194	0.1988 5347	0.1988 5344
		1.1	0.2312 3038	0.2312 3080	0.2312 3039	0.2312 3038
		1.6	0.1915 2647	0.1915 1797	0.1915 2648	0.1915 2647
		2.6	0.0699 8224	0.0699 8185	0.0699 8223	0.0699 8224
$\zeta=2.5, 2p_z$	$\zeta=1.1, 2p_z$	0.6	0.1726 3404	0.1726 3410	0.1726 3405	0.1726 3404
		1.1	0.0666 4252	0.0666 4275	0.0666 4253	0.0666 4252
		1.6	0.0605 0881	0.0605 0820	0.0605 0885	0.0605 0881
		2.6	0.0696 4352	0.0696 4271	0.0696 4349	0.0696 4352

Table 4.1: Selected integrals (all in Hartree), obtained with the prototype integrator for different grid size (denoted by  $(N_{\text{rad}} \times N_{\text{ang}} \times l_{\text{max}})$ ) and compared to the results of Harris [63].

harmonics, and the radial part is interpolated using cubic spline algorithms. With this, the numerical evaluation of the two-center overlap integral is straightforward and we are done with the solution.

We demonstrate the accuracy of the prototype integrator for two-center integrals of the form

$$I = \int F_A(\mathbf{r} - \mathbf{R}_A) F_B(\mathbf{r} - \mathbf{R}_B) \int \frac{F_A(\mathbf{r}' - \mathbf{R}_A) F_B(\mathbf{r}' - \mathbf{R}_B)}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' d\mathbf{r} \quad (4.28)$$

$$F_A(\mathbf{r}) = \frac{(2\zeta)^{n+1/2}}{\sqrt{(2n)!}} |\mathbf{r}|^{n-1} e^{-\zeta|\mathbf{r}|} Y_{lm}(\Omega), \quad (4.29)$$

where  $n, l, m$  are main, azimuthal and magnetic quantum numbers respectively, by comparison to the analytical integration routines, implemented for the MAPLE package [129] by Harris [63]. We use the grids recommended in ref. [14], which we denote as  $(N_{\text{rad}} \times N_{\text{ang}} \times l_{\text{max}})$ . Here  $N_{\text{rad}}$  is the size of the atomic radial grid,  $N_{\text{ang}}$  is the size of Lebedev grid and  $l_{\text{max}}$  is the cutoff angular momentum. In Tab. 4.1 we present the results for two selected integrals evaluated at four distinct interatomic distances. The convergence to the analytical result is clearly seen for increasing grid size. For the integrals over Yukawa interaction explicit analytical results are in general not available. The aforementioned  $\gamma^{\text{lr}}$ -integrals have been accurately reproduced by the integrator.

### 4.3 SCHEME FOR ONE-CENTER INTEGRALS

The parameters for the (LC-)DFTB are generated with respect to the pseudo-atomic basis set. This basis set consists of Kohn-Sham orbitals of some particular atomic DFT calculation with confinement potential (section 3.4).

The LC-DFTB requires the eigenvalues and orbitals from the atomic LC-DFT. To obtain these quantities it is necessary to extend the atomic DFT code to include the long-range corrected functionals. The modification of the local DFT functional to the screened analogon is rather simple as has been described in section 3.1. On contrast, the efficient modification of the integration routine for the evaluation of the exchange integrals over the descreened Yukawa interaction is

more sophisticated and will be described in following. The atomic DFT code uses Slater-type orbitals and resembles essentially the techniques from refs. [165, 166]. The implementation requires the complete evaluation of the one- and two-electron matrices before the SCF procedure begins. In order to include the long-range HF exchange term, we need to evaluate the exchange supermatrix  $K_{\lambda pq, \mu rs}^\omega$  for Yukawa interaction

$$K_{\lambda pq, \mu rs}^\omega = \frac{1}{d_\lambda d_\mu} \sum_{\alpha\beta} \int \int \chi_{p\lambda\alpha}(\mathbf{r}) \chi_{r\mu\beta}(\mathbf{r}') \frac{\exp(-\omega|\mathbf{r}-\mathbf{r}'|)}{|\mathbf{r}-\mathbf{r}'|} \chi_{s\mu\beta}(\mathbf{r}) \chi_{q\lambda\alpha}(\mathbf{r}') d\mathbf{r} d\mathbf{r}', \quad (4.30)$$

where  $d_\mu = (2\mu + 1)$  is the angular momentum degeneracy and we adopt the notation from the references [165, 166] for convenience.<sup>3</sup> The molecular orbitals are expanded in terms of atom-centered basis functions as  $\phi_k(\mathbf{r}) = \phi_{i\lambda\alpha}(\mathbf{r}) = \sum_p \chi_{p\lambda\alpha}(\mathbf{r}) c_{\lambda pi}$ . The molecular orbital index  $k$  has been split into the main quantum number  $i$ , angular momentum  $\lambda$  and magnetic quantum number  $\alpha$ . The basis functions  $\chi_{p\lambda\alpha}(\mathbf{r})$  are defined as

$$\chi_{p\lambda\alpha}(\mathbf{r}) = R_{p\lambda}(r) Y_{\lambda\alpha}(\Omega) \quad (4.31)$$

$$R_{p\lambda}(r) = \frac{2\alpha_{\lambda p}^{n_{\lambda p} + \frac{1}{2}}}{\sqrt{2n_{\lambda p}}} r^{n_{\lambda p} - 1} e^{-\alpha_{\lambda p} r}, \quad (4.32)$$

where  $n_{\lambda p} = n_p + \lambda$ ,  $r = |\mathbf{r}|$  and  $\Omega$  is the solid angle. We insert the definitions of the basis functions in Eq. 4.30 and obtain

$$K_{\lambda pq, \mu rs}^\omega = \frac{1}{d_\lambda d_\mu} \sum_{\alpha\beta} \int R_{p\lambda}(r) R_{s\mu}(r) Y_{\lambda\alpha}(\Omega) Y_{\mu\beta}(\Omega) \times \underbrace{\left[ \int \frac{\exp(-\omega|\mathbf{r}-\mathbf{r}'|)}{|\mathbf{r}-\mathbf{r}'|} R_{q\lambda}(r') R_{r\mu}(r') Y_{\lambda\alpha}(\Omega') Y_{\mu\beta}(\Omega') d\mathbf{r}' \right]}_{V_{q\lambda r\mu, \alpha\beta}^\omega(\mathbf{r})} d\mathbf{r}. \quad (4.33)$$

Note that for an atomic problem there is only one atom-centered grid and there is no need to do space partitioning. The inner integral is obtained by the solution of a set of one-dimensional Helmholtz equations

$$\left[ \frac{d^2}{dr^2} - \left( \frac{l(l+1)}{r^2} + \omega^2 \right) \right] V_{q\lambda r\mu, \alpha\beta}^{lm, \omega}(r) = -4\pi r \rho_{q\lambda r\mu, \alpha\beta}^{lm}(r) \quad (4.34)$$

as already described. The density, projected on the real spherical harmonics in the case of one-center inner integral has the simple form

$$\rho_{q\lambda r\mu, \alpha\beta}^{lm}(r) = \int R_{q\lambda}(r) R_{r\mu}(r) Y_{\lambda\alpha}(\Omega) Y_{\mu\beta}(\Omega) Y_{lm}(\Omega) d\Omega = R_{q\lambda}(r) R_{r\mu}(r) G(\lambda\alpha|\mu\beta|lm), \quad (4.35)$$

where  $G(\lambda\alpha|\mu\beta|lm) = \int Y_{lm}(\Omega) Y_{\lambda\alpha}(\Omega) Y_{\mu\beta}(\Omega) d\Omega$  are the real Gaunt coefficients. We insert this expression into Eq. 4.34 and obtain

$$\left[ \frac{d^2}{dr^2} - \left( \frac{l(l+1)}{r^2} + \omega^2 \right) \right] V_{q\lambda r\mu, \alpha\beta}^{lm, \omega}(r) = -4\pi r R_{q\lambda}(r) R_{r\mu}(r) G(\lambda\alpha|\mu\beta|lm). \quad (4.36)$$

<sup>3</sup>the atomic DFT code uses the same nomenclature.

The dependence of the solution on magnetic quantum numbers  $m, \alpha, \beta$  is only through the Gaunt coefficient on the right hand side of Eq. 4.36. We rescale the solution  $V_{q\lambda r\mu, \alpha\beta}^{l,m,\omega}(r) = V_{q\lambda r\mu}^{l,\omega}(r)G(\lambda\alpha|\mu\beta|lm)$ , and solve the equations

$$\left[ \frac{d^2}{dr^2} - \left( \frac{l(l+1)}{r^2} + \omega^2 \right) \right] V_{q\lambda r\mu}^{l,\omega}(r) = -4\pi r R_{q\lambda}(r) R_{r\mu}(r). \quad (4.37)$$

This means, that we don't need to solve the Helmholtz equations for all  $(l, m)$ , but only for  $l \leq l_{\max}$ . The inner integral of Eq. 4.33 takes the form

$$V_{r\mu q\lambda, \alpha\beta}^\omega(r, \Omega) = \sum_{lm} r^{-1} V_{r\mu q\lambda}^{l,\omega}(r) G(\mu\beta|\lambda\alpha|lm) Y_{lm}(\Omega). \quad (4.38)$$

Inserting this into Eq. 4.33 and performing the angular integration we arrive at

$$K_{\lambda pq, \mu rs}^\omega = \frac{1}{d_\lambda d_\mu} \int r dr R_{p\lambda}(r) R_{s\mu}(r) \sum_l V_{r\mu q\lambda}^{l,\omega}(r) \sum_{\alpha\beta m} G^2(\mu\beta|\lambda\alpha|lm) \quad (4.39)$$

$$= \frac{1}{d_\lambda d_\mu} \int r dr R_{p\lambda}(r) R_{s\mu}(r) \sum_l V_{r\mu q\lambda}^{l,\omega}(r) G_{\mu\lambda l}. \quad (4.40)$$

Each such integral requires the solution of  $(\lambda + \mu)/2$  Helmholtz equations Eq. 4.37 (due to the selection rules for Gaunt coefficients [73]), construction of the inner integral Eq. 4.38 and one radial integration Eq. 4.40.

We notice, that since the left-hand side of the Helmholtz equation Eq. 4.37 is determined by  $\omega$  and  $l$  only and does not depend on the integrand, the finite difference matrix and its LU-decomposition can be precomputed. If we assume that a matrix  $A$  can be represented by a product of lower triangular and upper triangular matrices  $A = LU$ , the linear problem can be written as

$$Ax = (LU)x = L(Ux) = b. \quad (4.41)$$

Thus the solution is performed by successive solution of two linear problems  $Ly = b$  and  $Ux = y$ , which are easily obtained by forward/backward substitution if matrices are triangular. Thus, provided the LU-decomposition of the finite difference matrix for a given  $\omega$  and each  $0 < l \leq l_{\max}$  is known, the only operations, which have to be done are the forward/backward substitutions for a given integrand. Additionally, the integrands can be precomputed in order to reduce the number of the exponential function evaluations.

To test the extension of the atomic DFT code to the new long-range corrected functional BNL-Y, defined in section 3.1, we compute the total energies of beryllium, carbon, oxygen and nitrogen atoms as a function of the range-separation parameter  $\omega$ . We compare our results to the reference calculation with NWCHEM package. The results are presented in Fig. 4.1. The dots denote the result of the atomic DFT code with BNL-Y functional. The basis consists of 12 STO functions per angular momentum. The solid lines denote the result, obtained on the BNL/aug-cc-pVDZ theory level, calculated with NWCHEM. The atomic configuration in the NWCHEM calculation

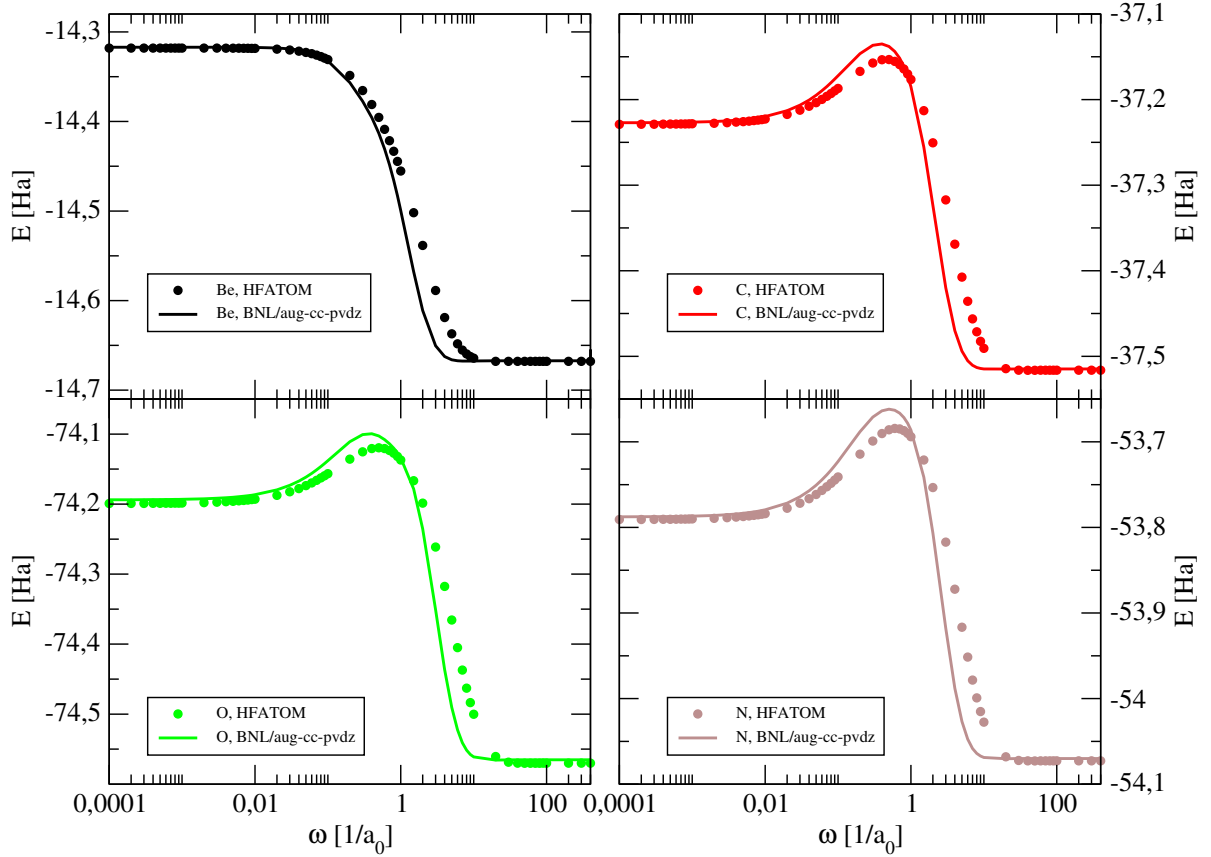


Figure 4.1: Total energies for the BNL-Y functional (section 3.1) of selected pseudo-atoms as a function of the range-separation parameter  $\omega$ , calculated with the atomic DFT code (HFATOM) as compared to BNL/aug-cc-pVDZ level of theory, calculated with NWCHEM quantum chemistry package. Note, that the functional in HFATOM is implemented with Yukawa-type range separation, while the functional in the NWCHEM uses the error-function.

is chosen in a way that the atom is spherically symmetric. For example for carbon atom the spin orbitals of p-shell are fractionally occupied with 1/3 electrons. We observe an expectable agreement in the DFT limit ( $\omega \rightarrow 0$ ) and in the HF+c limit ( $\omega \rightarrow \infty$ ). In the intermediate region the curves deviate, although they are qualitatively similar. The deviation stems from the different screening functions, used in both codes. NWCHEM functional uses the error-function and the atomic DFT code the Yukawa-type range separation.

#### 4.4 SCHEME FOR TWO-CENTER INTEGRALS

The two-center code evaluates the zeroth-order Hamiltonian and overlap matrix elements from the basis functions  $\phi_\mu$ , atomic densities  $\rho_A$  and atomic potentials  $V_A$ , which all are provided by

the atomic DFT code. The Hamiltonian in standard DFTB is obtained according to

$$\langle \phi_\mu | H[\rho_A + \rho_B] | \phi_\nu \rangle = \int \phi_\mu(\mathbf{r}) \left( -\frac{1}{2} \nabla^2 + V_A(\mathbf{r}) + V_B(\mathbf{r}) + v_{\text{xc}}[\rho_A + \rho_B](\mathbf{r}) \right) \phi_\nu(\mathbf{r}) d\mathbf{r}, \quad (4.42)$$

where the potential  $V_A(\mathbf{r})$  contains the contributions from Hartree and nuclear potentials of the atom  $A$ . Note, that this is essentially a two-center overlap integral. The xc-potential  $v_{\text{xc}}$ , which is non-linear in density has to be evaluated separately on a two-center grid. So far the integrals are simple overlap integrals on two-centers. If we include the non-local HF exchange term to the Hamiltonian, the sum of exchange integrals (compare Eq. 3.12)

$$\begin{aligned} I_x^\omega &= -\frac{1}{2} \sum_{\alpha\beta} P_{\alpha\beta}^0 (\mu\alpha|\alpha\nu)^\text{lr} = -\frac{1}{2} \sum_A \sum_{\alpha \in A} n_\alpha (\mu\alpha|\alpha\nu)^\text{lr} \\ &= -\frac{1}{2} \sum_A \sum_{\alpha \in A} n_\alpha \int \phi_\mu(\mathbf{r}) \phi_\alpha(\mathbf{r}) \frac{1 - e^{-\omega|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} \phi_\alpha(\mathbf{r}') \phi_\nu(\mathbf{r}') d\mathbf{r} d\mathbf{r}' \end{aligned} \quad (4.43)$$

has to be evaluated, where  $n_\alpha$  are the orbital occupations and  $P_{\alpha\beta}^0 = n_\alpha \delta_{\alpha\beta}$  is the reference density matrix. We note, that in general orbitals  $\mu, \alpha, \nu$  are centered on different atoms. This means, that the integral above is a three-center integral. Doing the two-center approximation, we require  $\mu \in A, \nu \in B \Rightarrow \alpha \in A$  or  $\alpha \in B$ . This allows us to rewrite the integral in the way, that inner integration is performed over the same center, i.e. we perform the one-center integration (in the same way as for atomic DFT). The basis functions  $\phi_\mu$  are the converged numerical orbitals, which have been calculated by the atomic DFT code and have the form

$$\phi_\mu(\mathbf{r}) = R_{n_\mu l_\mu}(r) Y_{l_\mu m_\mu}(\Omega), \quad (4.44)$$

where  $n_\mu$  is the main quantum number,  $l_\mu$  the angular momentum and  $m_\mu$  the magnetic quantum number. The radial part  $R_{n_\mu l_\mu}(r)$  is tabulated numerically, or can be reconstructed from the converged molecular orbital coefficients  $c_{\lambda p i}$  from an atomic DFT calculation. We assume that the radial part is tabulated. Let us further assume without loss of generality, that  $\alpha \in B$ , then

$$I_x^\omega = -\frac{1}{2} \int \phi_\mu(\mathbf{r}) \sum_\alpha \phi_\alpha(\mathbf{r}) \left[ \int \frac{1 - e^{-\omega|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} \phi_\alpha(\mathbf{r}') \phi_\nu(\mathbf{r}') d\mathbf{r}' \right] d\mathbf{r} \quad (4.45)$$

$$\begin{aligned} &= -\frac{1}{2} \int R_{n_\mu l_\mu}(r) Y_{l_\mu m_\mu}(\Omega) \sum_{n_\alpha l_\alpha m_\alpha} R_{n_\alpha l_\alpha}(r) Y_{l_\alpha m_\alpha}(\Omega) \times \\ &\times \left[ \int \frac{1 - e^{-\omega|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|} R_{n_\alpha l_\alpha}(r') Y_{l_\alpha m_\alpha}(\Omega') R_{n_\nu l_\nu}(r') Y_{l_\nu m_\nu}(\Omega') d\mathbf{r}' \right] d\mathbf{r}. \end{aligned} \quad (4.46)$$

The integral in the square brackets is the one-center integral over the descreened Yukawa interaction  $(1 - \exp(-\omega r))/r$ . The value of this integral is provided by solving the radial Poisson equations  $(1/r)$  and the Helmholtz equations  $(\exp(-\omega r))/r$  and subtracting the resulting  $l$ -components of the potentials (quantities  $V_{r\mu q \lambda}^{l, \omega}(r)$  in Eq. 4.38). We denote these components now as  $U_l^{n_\alpha l_\alpha n_\nu l_\nu, \omega}(r)$ .



The integral reads

$$I_x^\omega = -\frac{1}{2} \int R_{n_\mu l_\mu}(r) Y_{\lambda_\mu m_\mu}(\Omega) \sum_{n_\alpha l_\alpha} R_{n_\alpha l_\alpha}(r) \left[ \sum_l r^{-1} U_l^{n_\alpha l_\alpha n_\nu l_\nu, \omega}(r) \times \right. \\ \left. \times \sum_{m_\alpha m} G(l_\alpha m_\alpha | l_\nu m_\nu | l m) Y_{lm}(\Omega) Y_{l_\alpha m_\alpha}(\Omega) \right] d\mathbf{r} \quad (4.47)$$

$$= -\frac{1}{2} \int R_{n_\mu l_\mu}(r) Y_{\lambda_\mu m_\mu}(\Omega) \sum_{n_\alpha l_\alpha} R_{n_\alpha l_\alpha}(r) \left[ \sum_l r^{-1} U_l^{n_\alpha l_\alpha n_\nu l_\nu, \omega}(r) \sum_{l' m'} Y_{l' m'}(\Omega) \times \right. \\ \left. \times \sum_{m_\alpha m} G(l_\alpha m_\alpha | l_\nu m_\nu | l m) G(l m | l_\alpha m_\alpha | l' m') \right] d\mathbf{r} \quad (4.48)$$

$$= -\frac{1}{2} \int R_{n_\mu l_\mu}(r) Y_{\lambda_\mu m_\mu}(\Omega) \sum_{n_\alpha l_\alpha} R_{n_\alpha l_\alpha}(r) \times \\ \times \left[ \sum_l r^{-1} U_l^{n_\alpha l_\alpha n_\nu l_\nu, \omega}(r) \sum_{l' m'} Y_{l' m'}(\Omega) T_{l_\alpha l}^{l_\nu m_\nu} \delta_{l_\nu l'} \delta_{m_\nu m'} \right] d\mathbf{r} \quad (4.49)$$

$$= -\frac{1}{2} \int R_{n_\mu l_\mu}(r) Y_{\lambda_\mu m_\mu}(\Omega) \sum_{n_\alpha l_\alpha} R_{n_\alpha l_\alpha}(r) \left[ \sum_l r^{-1} U_l^{n_\alpha l_\alpha n_\nu l_\nu, \omega}(r) Y_{l_\nu m_\nu}(\Omega) T_{l_\alpha l}^{l_\nu m_\nu} \right] d\mathbf{r}. \quad (4.50)$$

Here we expanded the product of spherical harmonics, located at the same center  $Y_{lm}(\Omega) Y_{l_\alpha m_\alpha}(\Omega) = \sum_{l' m'} G(l m | l_\alpha m_\alpha | l' m') Y_{l' m'}(\Omega)$  and evaluated the sum over the product of Gaunt coefficients according to

$$\sum_{m, m_\alpha} G(l_\alpha m_\alpha | l m | l_\nu m_\nu) G(l_\alpha m_\alpha | l m | l' m') = \quad (4.51)$$

$$= \sum_{m, m_\alpha} \int Y_{l_\alpha m_\alpha}(\Omega) Y_{lm}(\Omega) Y_{l_\nu m_\nu}(\Omega) Y_{l_\alpha m_\alpha}(\Omega') Y_{lm}(\Omega') Y_{l' m'}(\Omega') d\Omega d\Omega' \quad (4.52)$$

$$= \int \sum_m Y_{lm}(\Omega) Y_{lm}(\Omega') \sum_{m_\alpha} Y_{l_\alpha m_\alpha}(\Omega) Y_{l_\alpha m_\alpha}(\Omega') Y_{l_\nu m_\nu}(\Omega) Y_{l' m'}(\Omega') d\Omega d\Omega' \quad (4.53)$$

$$= \int \frac{2l+1}{4\pi} P_l(\hat{\mathbf{x}} \cdot \hat{\mathbf{x}}') \frac{2l_\alpha+1}{4\pi} P_{l_\alpha}(\hat{\mathbf{x}} \cdot \hat{\mathbf{x}}') Y_{l_\nu m_\nu}(\Omega) Y_{l' m'}(\Omega') d\Omega d\Omega' \quad (4.54)$$

$$= \int \frac{2l+1}{4\pi} \frac{2l_\alpha+1}{4\pi} \sum_{L=|l_\alpha-l|}^{l_\alpha+l} C_{l_\alpha, l}^L P_L(\hat{\mathbf{x}} \cdot \hat{\mathbf{x}}') Y_{l_\nu m_\nu}(\Omega) Y_{l' m'}(\Omega') d\Omega d\Omega' \quad (4.55)$$

$$= \frac{2l+1}{4\pi} \frac{2l_\alpha+1}{4\pi} \sum_{L=|l_\alpha-l|}^{l_\alpha+l} C_{l_\alpha, l}^L \frac{4\pi}{2L+1} \sum_{M=-L}^L \int Y_{LM}(\Omega) Y_{LM}(\Omega') Y_{l_\nu m_\nu}(\Omega) Y_{l' m'}(\Omega') d\Omega d\Omega' \quad (4.56)$$

$$= \frac{2l+1}{4\pi} \frac{2l_\alpha+1}{4\pi} \sum_{L=|l_\alpha-l|}^{l_\alpha+l} C_{l_\alpha, l}^L \frac{4\pi}{2L+1} \sum_{M=-L}^L \delta_{L l_\nu} \delta_{M m_\nu} \delta_{L l'} \delta_{M m'} \quad (4.57)$$

$$= \frac{2l+1}{4\pi} \frac{2l_\alpha+1}{2l_\nu+1} C_{l_\alpha, l}^{l_\nu} \delta_{l_\nu l'} \delta_{m_\nu m'} = T_{l_\alpha l}^{l_\nu m_\nu} \delta_{l_\nu l'} \delta_{m_\nu m'}. \quad (4.58)$$

The expansion coefficients  $C_{l_\alpha, l}^{l_\nu}$  are given by  $3j$ -Wigner symbols

$$C_{l_\alpha, l}^\lambda = \begin{pmatrix} l_\sigma & l & \lambda \\ 0 & 0 & 0 \end{pmatrix}^2 (2\lambda+1). \quad (4.59)$$

The factors  $T_{l_\alpha l}^{l_\nu m_\nu}$  contain Gaunt coefficients, which can be precomputed and efficiently tabulated [152, 159], evaluated using recursive algorithms [173] or directly evaluated on Lebedev grid as the angular integral over products of spherical harmonics. In our implementation we use the algorithm by Pinchon et al. [152] which precomputes the Gaunt coefficients up to a given maximal angular momenta and efficiently stores them.

The integral Eq. 4.50 is a two-center overlap integral. We choose the coordinate system such that the  $z$ -axis is along the line, which connects the two centers. We switch to the spherical coordinates. The angular dependence of the integrand is only due to the spherical harmonics  $Y_{l_\mu m_\mu}(\Omega)$  and  $Y_{l_\nu m_\nu}(\Omega)$ , which are located at different centers. The choice of the coordinate system allows to perform the  $\phi$ -integration analytically. The real spherical harmonics are defined as

$$Y_{\ell m} = \begin{cases} \frac{1}{\sqrt{2}} (Y_\ell^m + (-1)^m Y_\ell^{-m}) = \sqrt{2} N_{(\ell, m)} P_\ell^m(\cos \theta) \cos m\varphi & \text{if } m > 0 \\ Y_\ell^0 & \text{if } m = 0 \\ \frac{1}{i\sqrt{2}} (Y_\ell^{-m} - (-1)^m Y_\ell^m) = \sqrt{2} N_{(\ell, |m|)} P_\ell^{|m|}(\cos \theta) \sin |m|\varphi & \text{if } m < 0. \end{cases} \quad (4.60)$$

Let  $a, b \geq 0$ , then we recall

$$\int_0^{2\pi} d\phi \sin(a\phi) \sin(b\phi) = \frac{1}{2} \int_0^{2\pi} d\phi (\cos[(a-b)\phi] - \cos[(a+b)\phi]) = \begin{cases} \pi & \text{for } a = b \\ 0 & \text{otherwise} \end{cases} \quad (4.61)$$

$$\int_0^{2\pi} d\phi \cos(a\phi) \sin(b\phi) = \frac{1}{2} \int_0^{2\pi} d\phi (\sin[(a+b)\phi] - \sin[(a-b)\phi]) = 0 \quad (4.62)$$

$$\int_0^{2\pi} d\phi \cos(a\phi) \cos(b\phi) = \frac{1}{2} \int_0^{2\pi} d\phi (\cos[(a+b)\phi] + \cos[(a-b)\phi]) = \begin{cases} \pi & \text{for } a = b \\ 0 & \text{otherwise.} \end{cases} \quad (4.63)$$

Thus for the product of two spherical harmonics we obtain

$$\int_0^{2\pi} Y_{l_1 m_1}(\theta, \phi) Y_{l_2 m_2}(\theta', \phi) d\phi = \bar{Y}_{l_1 m_1}(\theta) \bar{Y}_{l_2 m_2}(\theta') \times \begin{cases} 2\pi & m_1 = m_2 = 0 \\ \pi & m_1 = m_2 \neq 0 \\ 0 & \text{else,} \end{cases} \quad (4.64)$$

where the  $\bar{Y}_{l_1 m_1}(\theta) = \sqrt{2} N_{(\ell, m)} P_\ell^{|m|}(\cos \theta)$ . This gives the final expression for the integral

$$I_x^\omega = -\frac{1}{2} \int R_{n_\mu l_\mu}(r) \bar{Y}_{l_\mu m_\mu}(\theta) \sum_{n_\alpha l_\alpha} R_{n_\alpha l_\alpha}(r) \sum_l \frac{U_l^{n_\alpha l_\alpha n_\nu l_\nu, \omega}(r)}{r} \bar{Y}_{l_\nu m_\nu}(\theta) T_{l_\alpha l}^{l_\nu m_\mu} r^2 \sin \theta dr d\theta \times \begin{cases} 2\pi & m_\mu = 0 \\ \pi & m_\mu \neq 0 \end{cases} \quad (4.65)$$

We notice, that the integral vanishes unless  $m_\mu = m_\nu$ . It can be numerically evaluated together with other contributions in Eq. 4.42 on two atom-centered grids.

## PARAMETRIZATION OF REPULSIVE POTENTIALS FOR HYDROCARBONS

In section 3.2 the exact expression of the repulsive energy in the LC-DFTB method has been defined. It can be in principle directly evaluated within the two-center approximation using the representation in Eq. 3.13. However, the usual way of obtaining the repulsive energy for a given system is to approximate it by a sum of fast decaying pair-potentials

$$E_{\text{rep}} \approx \sum_{AB}^{\text{atoms}} V_{AB}(R_{AB}). \quad (5.1)$$

The sum in this expression runs over all atoms. The potentials for each pair of atomic species (type of the atom) are determined by a fit to a reference theory, which is usually the DFT. At the moment successful parametrizations of the repulsive potential for the standard DFTB method for a variety of elements exists [42, 52, 59, 104, 139, 182].

In following we describe the fit procedure and apply it to the LC-DFTB method for the carbon and hydrogen species. It should be noted, that the applicability of this approach to the LC-DFTB has been already verified in ref. [137].

### 5.1 WHY IS THE NEW PARAMETRIZATION NEEDED?

The repulsive energy in LC-DFTB method consists not only of nuclei-nuclei repulsion, but contains contributions from the electronic energy as can be seen in Eq. 5.2. The new term due to the long-range HF exchange lowers the energy if the range-separation parameter is increased. At the same time the exact repulsive energy

$$E_{\text{rep}} = E_{\text{xc,sr}}^{\omega}[\rho_0] - \sum_{\mu\nu} P_{\mu\nu}^0 v_{\mu\nu}^{\text{xc}}[P^0] - \frac{1}{2} \sum_{\mu\nu\alpha\beta} P_{\mu\nu}^0 P_{\alpha\beta}^0 (\mu\nu|\alpha\beta) + \frac{1}{4} \sum_{\mu\nu\alpha\beta} P_{\mu\nu}^0 P_{\alpha\beta}^0 (\mu\alpha|\beta\nu)^{\text{lr}} + E_{NN}, \quad (5.2)$$

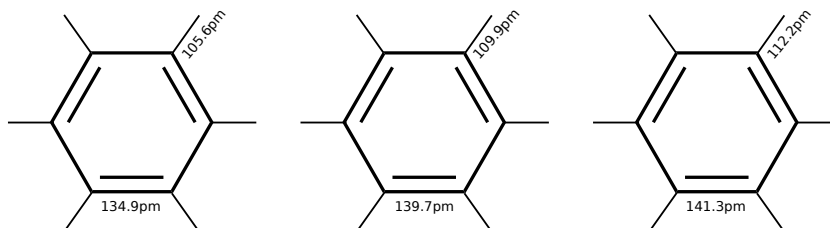


Figure 5.1: The schematic geometry of benzene molecule. Marked are the C-C and C-H bond lengths, as obtained from LC-DFTB with mio-1-1 repulsive (left), standard DFTB (center) and DFT on the BNL/6-311G\* level of theory (right). We clearly see a considerable shrinking effect due to the too weak repulsive potentials of the mio-1-1 parameter set if combined with the LC-DFTB.

includes the corresponding repulsive counter-term, evaluated at the zeroth-order density matrix. This suggests, that for the new theory the approximate repulsive pair-potentials Eq. 5.1 should be reparametrized. Specifically, they should be made more repulsive than the respective potentials from the standard DFTB method. The expectation is that the standard DFTB repulsive potential, which has been fitted to the hybrid B3LYP functional, will be too weak and the geometries will shrink. The question is how large is this effect and whether the existing repulsive potentials can still be used?

We optimized the geometries for a set of hydrocarbons (the list can be found in Tab. D.1) with the LC-DFTB method and standard mio-1-1 repulsive potential. The bond lengths have been compared to the values, obtained from the standard DFTB, the local PBE/6-311G\*, the long-range corrected BNL/6-311G\* and to experimental data. We found a systematic underestimation of bond lengths by the LC-DFTB with mio-1-1 repulsive potentials. The mean absolute error ranges from 5.0 pm (with respect to experiment), to 7.1 pm (with respect to the BNL/6-311G\*). As an example, in Fig 5.1 we show the schematic geometries of benzene molecule, where we mark the carbon-carbon and carbon-hydrogen bond lengths, as obtained by the LC-DFTB method (left), DFTB (middle) and BNL/6-311G\* (right). The LC-DFTB underestimates the carbon-carbon (carbon-hydrogen) bond length by 6.4 (6.6) pm with respect to the BNL/6-311G\* and by 4.8 (4.3) pm with respect to the standard DFTB. Thus we conclude that the repulsive potential should be adjusted (made more repulsive) for the new electronic energy.

## 5.2 REPULSIVE POTENTIAL FIT PROCEDURE

In following we describe the procedure of the repulsive potential parametrization as it is usually performed in the standard DFTB. The total LC-DFTB energy Eq. 3.36 can be written in terms of the total electronic energy  $E_{\text{LC-DFTB}}^{\text{elec}}$  and the aforementioned sum of pair-potentials  $V_{AB}(R)$

$$E_{\text{LC-DFTB}}^{\text{total}} = E_{\text{LC-DFTB}}^{\text{elec}} + \sum_{AB}^{\text{atoms}} V_{AB}(R_{AB}). \quad (5.3)$$

We want to fit the pair-potentials in such a way, that the total LC-DFTB energy  $E_{\text{LC-DFTB}}^{\text{total}}$  approximately reproduces a Born-Oppenheimer potential energy surface of some reference systems at a specific theory level. We then assume, that such pair-potentials are transferable and can be applied to all other systems. The reasonable way to do this is to find a path in the space of geometric configurations of some suitable physical system, along which only the pair-potential which we are interested in, or a multiple of it, is changed together with the electronic energy. Practically this can be achieved by partitioning the system into fragments, and stretching them along the appropriate direction. Once the repulsive pair-potential  $V(R)$  is isolated due to this procedure, we fit the total LC-DFTB energy to the reference of our choice, usually a DFT total energy. The prescription can be formulated as follows

$$E_{\text{DFT}}(R) = E_{\text{LC-DFTB}}^{\text{elec}}(R) + nV(R) + C, \quad (5.4)$$

where  $R$  is the stretch coordinate (distance between the fragments, which are pulled apart),  $n$  the number of the repulsive potentials of the same type, which are simultaneously varied upon stretching (for example of C-H bonds in  $\text{CH}_4$  molecule, as described in section 5.3) and  $C$  is some constant. From  $\lim_{R \rightarrow \infty} V(R) = 0$  it follows that  $C = \lim_{R \rightarrow \infty} (E_{\text{DFT}}(R) - E_{\text{LC-DFTB}}^{\text{elec}}(R))$ . The constant  $C$  contains contributions due to the pair-potentials, which do not depend on the stretch coordinate, and other contributions from the difference of the total DFT and the electronic LC-DFTB energies. The direct evaluation of this constant is usually not possible, since for example not all pair-potentials are available during the parametrization. For this reason the fit procedure is carried out in two steps. First, we vary the stretching parameter  $R$  in some suitable interval and calculate the difference of the total DFT energy and the electronic part of the LC-DFTB energy

$$\Delta E(R) = \frac{1}{n} \left[ E_{\text{DFT}}(R) - E_{\text{LC-DFTB}}^{\text{elec}}(R) \right]. \quad (5.5)$$

Then, the difference is shifted, such that it decays to zero for  $R \rightarrow \infty$ . Practically this happens at some finite cutoff distance. This is necessary to avoid the superposition of repulsive potentials at long-range. The shifted energy difference curve is the repulsive potential, which we are interested in. To represent the potential, the spline interpolation is performed and the resulting spline is attached to the corresponding Slater-Koster file in a usual way.

At this point it is worth mentioning, that the total (LC-)DFTB energy, obtained in such a way, is not even nearly the same as the total energy from the first principles methods. However, the absolute value of the total energy is usually of no interest. The energy differences from DFTB are reasonable, as can be seen from benchmarks for atomization energies and reaction energies [35, 41–44].

It should be also noted, that the parametrization for a set of  $N$  atomic species is a hard optimization task. Recently the automatic parametrization became possible [21]. Thus it is tempting to apply these tools to the LC-DFTB Hamiltonian.

Molecule	This		mio-1-1		n	Bond type
	min [Å]	max [Å]	min [Å]	max [Å]		
H <sub>2</sub>	0.54	0.99	0.55	0.97	1	H–H
CH <sub>4</sub>	0.76	1.57	n/a	n/a	4	C–H
C <sub>2</sub> H <sub>2</sub>	0.88	1.24	0.90	1.26	1	C≡C
C <sub>2</sub> H <sub>4</sub>	1.24	1.42	1.26	1.43	1	C=C
C <sub>2</sub> H <sub>6</sub>	1.42	1.99	1.43	3.02	1	C–C

Table 5.1: Interval limits [min,max], used for the parametrization of the repulsive potentials for carbon and hydrogen species for the LC-DFTB theory with  $\omega = 0.3a_0^{-1}$  (this). Additionally the corresponding parameters for the mio-1-1 parametrization are given.

### 5.3 PARAMETRIZATION AND TEST

The parametrization of LC-DFTB repulsive potentials ( $\omega = 0.3a_0^{-1}$ ) for hydrogen and carbon species has been performed using B3LYP/6-311G\* level of theory for the reference fit. This choice and the selection of reference molecules, parallels that of the standard DFTB mio-1-1 set parametrization [42]. The hydrogen-hydrogen repulsive potential is fitted to the H<sub>2</sub> molecule, which is stretched along its symmetry axis. For the hydrogen-carbon repulsive potential the model system is the methane molecule CH<sub>4</sub>. It contains four symmetric C-H bonds, which are then simultaneously stretched. Note, that it is necessary to take the factor  $n = 4$  in Eq. 5.5 into account. Finally, the repulsive potential for the carbon-carbon interaction is fitted to the three molecules with different type of carbon-carbon bond in different intervals. These intervals are chosen such that the corresponding typical bond length is included into the interval. The acetylene (C<sub>2</sub>H<sub>2</sub>) accounts for the triple bond ( $\approx 120$  pm), ethylene (C<sub>2</sub>H<sub>4</sub>) for double bond ( $\approx 132$  pm) and ethane (C<sub>2</sub>H<sub>6</sub>) for the single bond ( $\approx 150$  pm). In the inset of Fig. 5.2 the shifted energy differences as a function of interatomic distance for all three situations are plotted. By smooth connection (shift upwards) of these pieces, the resulting carbon-carbon repulsive potential curve emerges. All reference geometries have been optimized on the B3LYP/6-311G\* level. The stretching was applied to the considered bond only. No additional geometry relaxation (e.g. on methyl groups of C<sub>2</sub>H<sub>6</sub>) have been done.

We summarize the interval limits, in which the respective bond type has been parametrized in Tab. 5.1. Compared to the original mio-1-1 set, the upper limit for the single bond interval of the carbon-carbon interaction is significantly different and is 1.99 Å instead of 3.02 Å. The parametrization with the mio-1-1 value of 3.02 Å (we denote it as S<sub>1</sub>) resulted in very large C-C bond lengths ( up to 170 pm for single C-C bonds in *trans*-polyacetylene chains). Moreover, the single bond on the reference structure (ethane) was not reproduced. The error was roughly 1 pm, which is too large. Shifting the LC-DFTB curve S<sub>1</sub> down we found the limit of 1.99 Å to give the best results in terms of bond lengths of the reference systems. In this case the whole repulsive curve is shifted down by 0.071167 Hartree. This shifted curve is cut at its intersection point

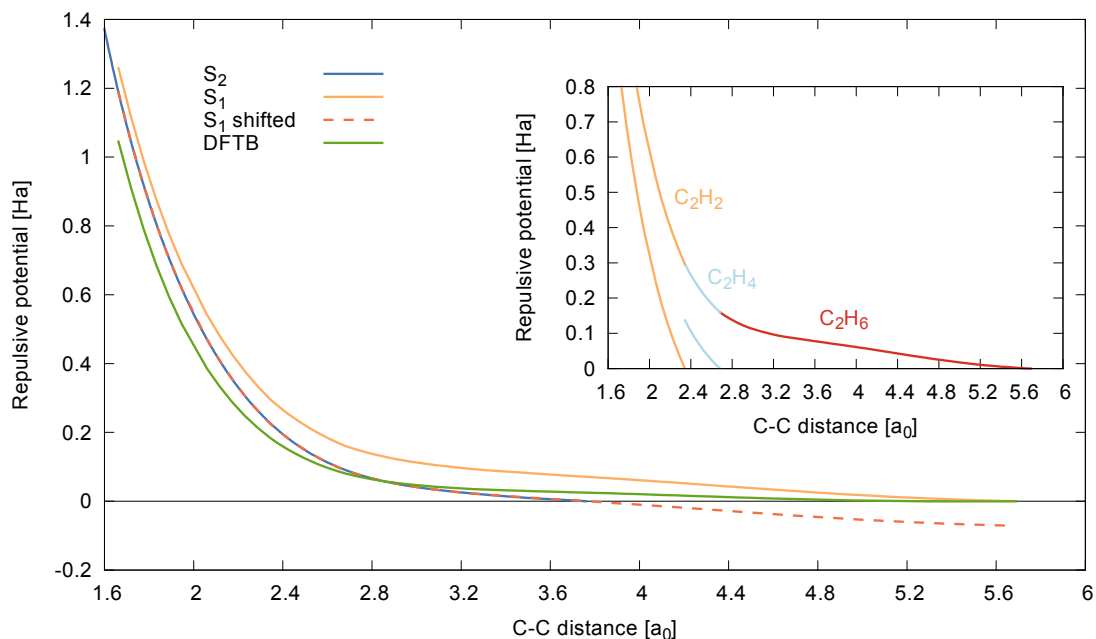


Figure 5.2: The C-C repulsive potentials  $S_1$  and  $S_2$  as a function of interatomic distance (see main text). The spline  $S_2$  has been obtained by shifting the  $S_1$  and cutting it at the point, where it crosses the x-axis. The C-C repulsive for the standard DFTB is plotted for comparison. In the inset the separate repulsives for distinct bond types and the combined repulsive are depicted.

with the x-axis and is denoted as  $S_2$ . We plot both curves  $S_1$  and  $S_2$  in the main part of Fig. 5.2. Additionally we plot the combined C-C repulsive potential, which we obtain for the standard DFTB theory. For this case we used the same intervals as for  $S_1$  parametrization (the upper limit of single bond set to  $3.02 \text{ \AA}$  as in mio-1-1 set). This curve is expected to be close to that of the original mio-1-1 fit. The DFTB repulsive potential seems to decay sharper than the  $S_1$  and  $S_2$  curves. We observe also, that it is in the whole range weaker than the  $S_1$  repulsive potential. The shifted curve ( $S_2$ ) is still more repulsive in the interval between 0 and  $2.89 a_0$  ( $\approx 150 \text{ pm}$ ) than the DFTB potential. In the region of approx  $2.89 a_0$  to the  $3.76 a_0$  (the cutoff value) it is, however, less repulsive than the DFTB potential. The general impression is, that the  $S_1$  curve does not decay fast enough. This leads to the repulsive potential, which is too strong. However, as will be shown in following the shifted  $S_2$  curve constitutes a reasonable parametrization.

The bond lengths and angles for the reference systems, as obtained from the geometry relaxation with the LC-DFTB with the new repulsive potential parametrization, are summarized in the tail of Tab. D.1. In addition, we provide the values from experiment [65, 115], standard DFTB and first principles B3LYP/6-311G\* and BNL/6-311G\* levels of theory. The new parametrization (curve  $S_2$ ) reproduces the bond lengths and angles of hydrogen molecule and methane exactly. For the acetylene, ethylene and ethane we found deviations in C-H bonds (maximal  $2.2 \text{ pm}$ ) and angles (maximal  $0.2^\circ$ ). The carbon-carbon bond length for acetylene and ethane is reproduced

	Exp.	LC-DFTB	DFTB	B3LYP/6-311G*	BNL/6-311G*
Exp.	-	0.016	0.010	0.009	0.021
LC-DFTB	1.5	-	0.016	0.018	0.012
DFTB	1.3	0.5	-	0.010	0.023
B3LYP/6-311G*	1.5	0.4	0.5	-	0.026
BNL/6-311G*	1.5	0.5	0.4	0.3	-

Table 5.2: Mean absolute error for bond lengths in [ $\text{\AA}$ ] (upper triangle) and angles in [deg] (lower triangle) for the hydrocarbon benchmark set, obtained from different theories and experiment. Experimental values are the average of refs. [65, 115].

exactly, while for the ethane molecule the carbon bond length is overestimated by 0.1 pm. It can be in principle cured by sampling the energy differences as a function of interatomic distance on a finer grid and further shifting the potential curve. For this first fit we accept this error.

The vibrational frequencies for all reference molecules, but hydrogen <sup>1</sup> as obtained by LC-DFTB, DFTB, B3LYP/6-311G\* and BNL/6-311G\* are summarized in Tab. D.2. We find the mean absolute error (MAE) for the LC-DFTB of  $108 \text{ cm}^{-1}$  and the mean signed error (MSE) of  $106 \text{ cm}^{-1}$  with respect to the experimental data [115]. Especially large relative errors of 10-17% are found for C-C stretch modes. Similar behavior has been observed also for standard DFTB [41]. Compared to the LC-DFTB, the MAE in vibrational frequencies from DFTB ( $51 \text{ cm}^{-1}$ ), B3LYP/6-311G\* ( $64 \text{ cm}^{-1}$ ) and BNL/6-311G\* ( $45 \text{ cm}^{-1}$ ) are a factor of 2 smaller.

Finally, we benchmark the bond lengths and angles on a set of selected hydrocarbons. The detailed results for LC-DFTB, standard DFTB, B3LYP/6-311G\* and BNL/6-311G\* as well as corresponding experimental values [65, 115] are summarized in Tab. D.1 of the appendix D. The mean absolute errors for bond lengths (upper triangle) and angles (lower triangle) of this set for all theories are presented in Tab. 5.2. Note, that dihedral angle of biphenyl molecule was not included into statistics. With respect to the experiment all theories show the MAE in bond angles below  $2^\circ$ . The bond lengths for LC-DFTB with respect to the experiment show MAE of 1.6 pm. This is slightly worse than the results of DFTB and B3LYP/6-311G\*, but better than the BNL/6-311G\* result. Inspection of the data in Tab. D.1 reveals, that the LC-DFTB shows the same tendency of overestimation of the bond lengths as the BNL/6-311G\*. From the qualitative point, only the LC-DFTB result for the isobutane molecule was conspicuous. In contrast to the DFTB and first principles methods, it shows the same HCH angles on methyl groups. To conclude, the overall performance of the LC-DFTB with respect to the geometries is worse than that of DFTB and B3LYP/6-311G\*, but shows similar errors as the long-range corrected BNL/6-311G\* theory. Since the MAE in bond lengths and angles of the benchmark set for the LC-DFTB with the new parametrization are below 2 pm and  $2^\circ$  respectively, we consider this fit as useful.

<sup>1</sup> For the hydrogen molecule the LC-DFTB method gives the value  $4394.8 \text{ cm}^{-1}$ , which compared to the experimental value  $4161.2 \text{ cm}^{-1}$  [37] gives an absolute error of  $233.6 \text{ cm}^{-1}$  and relative error of 5.6%.



## SUMMARY

**B**ased on the work of Niehaus and Della Sala [137], we extended the standard DFTB method to the important class of long-range corrected hybrid exchange-correlation functionals. The extension required the modification of the atomic DFT code and the two-center code to include the long-range corrected functionals. Both programs constitute an important part of the parametrization toolkit of the DFTB method. A long-range corrected functional with the specific range-separation function essentially consists of the screened local DFT exchange, standard DFT correlation functional and the long-range Hartree-Fock exchange term (section 3.1). The introduction of screening factor in the local DFT functional is, apart from possible numerical issues due to the particular implementation, unproblematic. At the same time, efficient evaluation of the long-range Hartree-Fock exchange contribution in the implementations, based on the Slater-type orbitals, is not straightforward.

The chapter 4 was dedicated to this problem. We extended the numerical integration method for two-electron integrals over Coulomb interaction [13, 18] to the case of Yukawa interaction. Further, we adopted this algorithm for the atomic DFT code, which requires the evaluation of one-center integrals and for the two-center code, which requires two-center integrals. It should be mentioned, that the actual DFTB calculation requires no run-time integral evaluation. However, the computational performance of the parametrization tools is crucial if the optimization of parameters, such as confinement radii (section 3.4) is necessary. Our problem-specific extensions do not considerably increase the execution times of atomic DFT and two-center codes. With the extended parametrization tools for the specific long-range corrected functional, defined in section 3.1, we generated the parameters for the electronic part of the Hamiltonian, which includes the elements C,N,O,H and S. The parametrization of the DFTB method, based on the long-range corrected functional (section 3.1) is called LC-DFTB.

The zeroth-order LC-DFTB, which by design has the same computational performance as the standard DFTB method turned out to be not useful. Specifically, we did not observe the correction of the HOMO eigenvalue towards higher absolute values, which would indicate the reduction of the delocalization problem and is characteristic for the long-range corrected DFT. It turns out, that the choice of the reference density matrix as the superposition of atomic density matrices, which perfectly works in the standard zeroth-order DFTB, is insufficient if the long-range Hartree-Fock exchange term is present. We found, that this problem is also present in the first-principles methods if the initial guess is taken to be the superposition of atomic densities and only one diagonalization is performed. For this reason we conclude that the self-consistent solution is indispensable.

In order to ensure the compatibility with the standard DFTB, which should emerge as the limit  $\omega \rightarrow 0$  from the LC-DFTB method, the approximations, which are usually used in the so called self-consistent charge extension of the zeroth-order DFTB, have been applied also to the LC-DFTB. In contrast to the standard DFTB, the self-consistency in the LC-DFTB is achieved with respect to the density matrix and not to the Mulliken charges. The two-electron integrals are approximated by two-parameter  $\gamma$ -integrals. In the new method there are two kinds of  $\gamma$ -integrals. The full-range integral, known from the standard DFTB accounts for Hartree potential and the linearized exchange-correlation potential. The long-range integral is new and accounts for the long-range HF exchange term. We were able to derive an analytical formula for the long-range  $\gamma$ -integral in the way it is done for standard DFTB (section 3.7 and appendix A). The parameters for the  $\gamma$ -integrals (decay constants  $\tau$ ) are fixed according to the condition that the Hubbard derivatives  $U$ , obtained from a reference atomic LC-DFT calculation and that of an atomic LC-DFTB calculation should be equal. Because of the exact exchange term the usual relation  $\tau = 3.2U$  is not valid anymore (section 3.8). The correction has been proposed, which requires the solution of Eq. 3.52 for  $\tau$ . A possible numerical solution algorithm is described in appendix C.

The inclusion of the long-range HF exchange term into the Hamiltonian and energy expressions of the LC-DFTB method results in the increased computational requirements. We proposed the algorithms for the evaluation of the HF exchange part, which scale quadratically with the basis size. The thresholding algorithm (section 3.9.1) is the adaptation of the direct SCF techniques, known from the first principles approaches. It omits the evaluation of the contributions to the Hamiltonian which are not significant according to given cutoff criteria. The reduction in the computational cost by the factor of 2-3 is achieved, while keeping the mean absolute error in eigenvalues below  $10^{-6}$  Ha. Such an error is acceptable for practical calculations. The neighbor list-based algorithm (section 3.9.2) has been formulated, since the implementation of the DFTB+ code, which will be extended to include the LC-DFTB method, is entirely based on the neighbor list concept.

Finally, we extended the algorithms for the evaluation of energy gradients with respect

---

to the nuclear positions (section 3.10). This allows to perform the geometry relaxation in the LC-DFTB method and is important for molecular dynamics simulations. The adjustment of the repulsive potentials for the case of hydrogen and carbon has been described in chapter 5. The LC-DFTB with this parametrization of the repulsive potentials performed well with respect to the geometries of selected hydrocarbons. The MAE error in bond lengths of 0.016 Å with respect to the experiment can be considered as acceptable. The vibrational frequencies of the reference molecules showed MAE of  $\approx 108 \text{ cm}^{-1}$  which is larger than the DFTB error by factor of 2. This attempt reveals the importance of automatic repulsive parametrization techniques, since the parameters of the repulsive fit procedure, used for the standard DFTB seem to be not the optimal choice for the LC-DFTB method.

The predictive power of the new method will be assessed in the second part II of this thesis on a series of typical problems.



**Part II**

**Applications**



## QUASI-PARTICLE ENERGIES FROM LC-DFTB

The long-range corrected functionals have been shown to reduce the delocalization problem (chapter 2). The key quantity in this respect is the linearity condition, which implies that the HOMO eigenvalue from a LC-DFT calculation should be equal up to a sign to the ionization potential of the system, which is also defined as the difference of the total ground state energy of the neutral system and the total energy of its cation. This condition can be seen as a measure for the reduction of the delocalization problem. For the LC-DFT a series of studies in this context have been done [34, 91, 99, 100, 103, 160]. The general finding is that the LC-DFT provides the HOMO eigenvalues, which are quite close to the experimental values. Moreover, the HOMO-LUMO gaps approximate the experimental fundamental gaps with remarkable accuracy. For this reason the assessment of the eigenvalue spectrum of the LC-DFTB is important for its characterization. We want to compare the eigenvalue spectrum, especially the HOMO and LUMO eigenvalues as calculated by the LC-DFTB method to the experimental data (where it is available), standard DFTB method, DFT with local, global hybrid and long-range corrected xc-functionals. For the benchmark we choose a set of 35 organic molecules, which contain the elements H,C,N,O,S. For all these molecules the experimental gas-phase ion energetics data (specifically the ionization potentials) is available at the NIST database [115], which we take as the reference. The set contains among others the photovoltaically relevant molecules, studied in refs. [20, 160] with the GW method and the optimally tuned LC-DFT. In addition, the important series of polyacene oligomers from benzene to hexacene is included. The structural formulas of these molecules are summarized in Fig. 7.1 We begin with ionization potentials in section 7.1. The HOMO-LUMO gaps for this benchmark set are discussed in section 7.2. Inspired by numerous studies on the photoemission spectra from the LC-DFT, we discuss the LC-DFTB eigenvalue spectrum for pentacene and the perylene-3,4,9,10-tetracarboxylic-dianhydride (PTCDA) molecule in section

Molecule	Exp.	PBE/cc-pVTZ	B3LYP/cc-pVTZ	BNL/3-21G	BNL/cc-pVDZ	BNL/cc-pVTZ	LC-DFTB	DFTB
Fluorene	7.91	5.40	7.02	7.68	7.81	7.97	8.08	5.89
PTCDA	8.20	6.30	6.88	7.84	8.05	8.32	8.46	6.41
C <sub>60</sub>	7.60	5.80	6.35	8.01	7.83	8.02	7.80	5.85
H <sub>2</sub> P	6.90	4.97	5.48	6.31	6.62	6.85	6.79	5.16
H <sub>2</sub> TPP	6.42	4.64	5.07	5.83	6.10	6.30	6.32	4.91
H <sub>2</sub> Pc	6.41	4.98	5.27	6.17	6.15	6.32	6.12	4.95
Benzothiazole	8.80	5.94	6.66	8.36	8.42	8.62	8.46	6.24
Benzothiadiazole	9.00	6.10	6.82	8.44	8.55	8.79	8.66	6.50
Thiadiazole	10.10	6.83	7.69	9.70	9.56	9.76	8.75	6.47
Thiophene	8.86	5.79	6.58	8.54	8.58	8.79	8.76	6.24
2-Thiophene	7.75	5.00	5.66	7.40	7.43	7.62	7.85	5.52
3-Thiophene	7.43	4.70	5.29	6.92	6.94	7.12	7.40	5.23
5-Thiophene	7.11	4.47	4.99	6.53	6.54	6.71	6.79	5.00
1-Acene	9.24	6.29	7.04	8.89	9.04	9.21	9.23	6.69
2-Acene	8.14	5.44	6.09	7.70	7.87	8.04	8.22	5.98
3-Acene	7.44	4.94	5.50	6.95	7.12	7.29	7.55	5.52
4-Acene	6.97	4.61	5.12	6.44	6.62	6.78	7.10	5.22
5-Acene	6.63	4.38	4.85	6.08	6.26	6.42	6.79	5.02
6-Acene	6.40	4.23	4.66	5.82	6.00	6.15	6.57	4.88
Perylene	6.96	4.70	5.20	6.56	6.71	6.86	7.17	5.32
Coronene	7.29	5.21	5.74	7.11	7.24	7.39	7.63	5.71
NTCDA	9.67	6.84	8.04	9.20	9.42	9.73	9.33	6.79
Methane	12.61	9.43	10.76	12.62	12.74	12.90	11.79	9.14
Pyridine	9.26	5.77	7.09	8.23	8.78	9.06	8.39	6.21
Cyclopropene	9.67	5.94	6.86	8.74	8.86	9.03	8.74	6.29
Ketene	9.62	5.86	6.90	8.20	8.63	8.98	8.46	6.41
Dimethylether	10.03	5.63	7.06	7.98	8.64	9.02	8.24	5.89
Dimethylsulfide	8.69	5.05	6.06	7.92	7.90	8.13	7.99	5.69
Formaldehyde	10.88	6.17	7.56	8.61	9.02	9.41	8.61	6.36
Pyrene	7.43	5.06	5.61	7.00	7.17	7.33	7.56	5.59
1,3-Butadiene	9.07	5.79	6.55	8.31	8.54	8.73	8.97	6.44
Propene	9.73	6.14	7.04	8.94	9.14	9.33	9.40	6.78
Pyridazine	8.74	5.24	6.61	7.55	8.19	8.53	7.89	5.72
Pyrimidine	9.33	5.73	7.04	8.27	8.74	9.05	8.28	6.10
Pyrazine	9.28	5.71	7.01	8.10	8.62	8.94	7.96	5.84

Table 7.1: The negative of the HOMO eigenvalue for different theories compared to the experimental ionization potential. All energies in eV.

7.3. Finally, the computational performance with respect to the high level theories is compared in section 7.5. The results of this chapter have been published in ref. [119].

## 7.1 IONIZATION POTENTIALS

We compare the negative of the HOMO eigenvalues from the LC-DFTB, standard DFTB and first principles theories with local xc-functional (PBE/cc-pVTZ) [146], the hybrid functional B3LYP [15, 109, 185, 197] (B3LYP/cc-pVTZ) and the long-range corrected BNL functional with small basis set (BNL/3-21G), medium basis set (BNL/cc-pVTZ) and large basis set (BNL/cc-pVDZ). The latter basis set, the correlation-consistent polarized valence-only triple zeta basis set [39] is usually recommended for practical high quality calculations, whereas its double zeta variant cc-pVDZ is often used for preoptimization. The smaller basis set 3-21G and cc-pVDZ are included to track the basis set effects. The BNL functional is similar to that, which



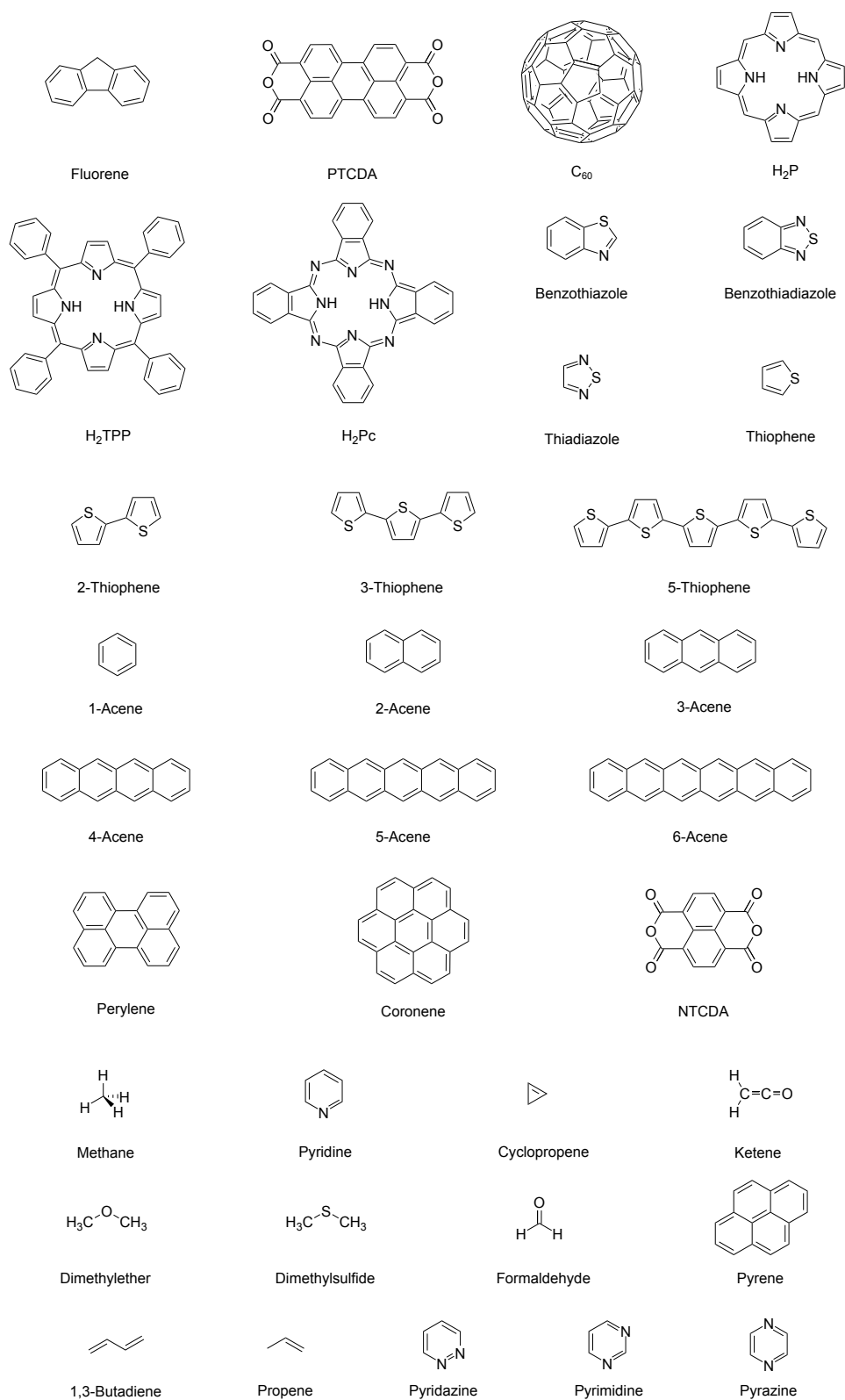


Figure 7.1: The structural formulas of the benchmark set of 35 organic molecules.

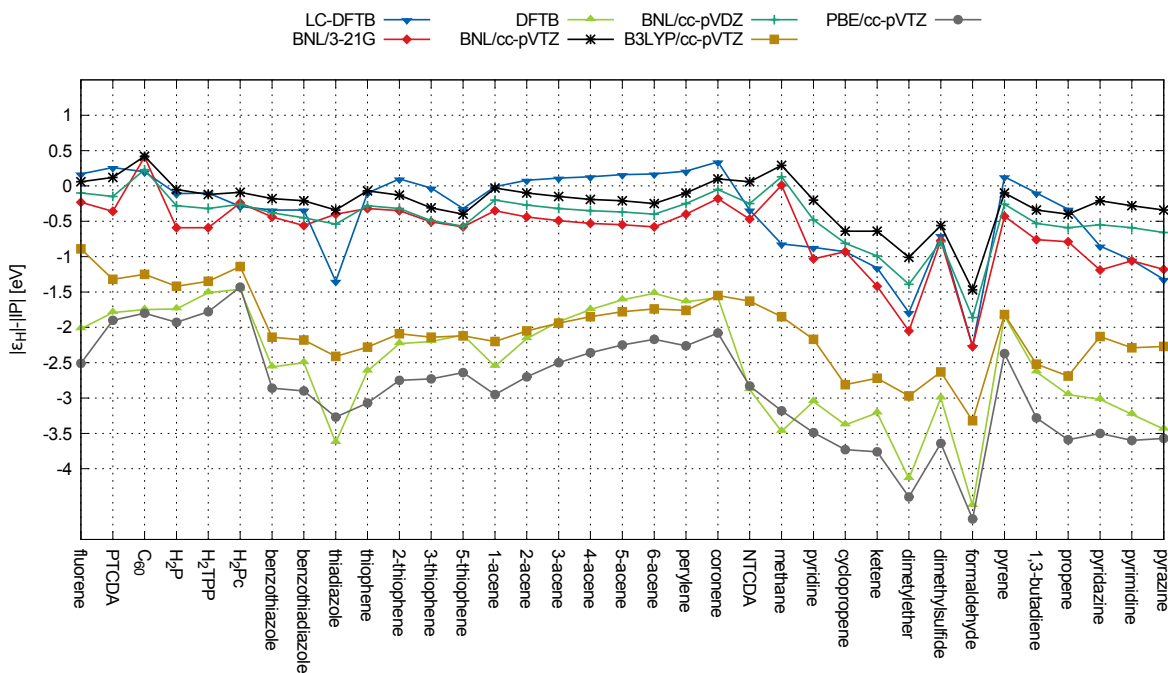


Figure 7.2: The deviation of the HOMO eigenvalue from the experimental ionization potential for the 35 molecule benchmark set for different theories.

has been used for the parametrization of LC-DFTB (section 3.1). However, it uses the error-function as the range-separation function and a fraction of 90% of LYP correlation energy [109]  $E_{xc}^{BNL} = E_{sr,x}^{LDA} + 0.9E_c^{LYP} + E_{x,lr}^{\omega, HF}$  [116].

The geometries for all calculations in this benchmark, unless stated otherwise, have been optimized on the DFTB level with the mio-1-1 parameter set [42, 139]. This procedure is somewhat arbitrary, although quite common. It is based on the assumption, that the ground state geometry from other theory level will not impose a considerable deviation of the electronic ground state as compared to the native ground state geometry. Because of this, in principle, one could choose the first-principles geometries as well. The particular motivation for the choice of DFTB as geometry optimization level for this study is the computational efficiency of the method on the one hand and it's ability to give useful geometries on the other. We additionally note, that the DFTB geometry is not native to the LC-DFTB. For this reason this choice of geometry does not privilege the LC-DFTB method. All first-principles calculations have been performed using NWCHEM package, with it's default settings, unless convergence issues have been observed. In such case the symmetry option has been turned off.

The deviation  $\Delta = |\epsilon_{HOMO}| - |IP_{exp}|$  is presented in Fig 7.2 for all theories. The numerical values for all compounds are summarized in Tab. 7.1. We observe a better agreement of the LC-DFTB with experiment, compared to the standard DFTB, PBE and B3LYP. The mean absolute

error for LC-DFTB is 0.50 eV compared to BNL/3-21G (0.67 eV), BNL/cc-pVDZ (0.47 eV), and BNL/cc-pVTZ (0.29 eV). The deviations for the local and hybrid functionals on the other hand are much larger: B3LYP/cc-pVTZ (2.04 eV), PBE/cc-pVTZ (2.87 eV), DFTB (2.50 eV).

Despite a quite satisfying overall picture provided by the LC-DFTB method, there are some remarkable outliers. We observe deviations of more than 0.5 eV for small compounds, which contain nitrogen (pyrazine, pyrimidine, pyridazine, pyridine), oxygen (ketene, dimethylether, formaldehyde), sulfur (thiadiazole and dimethylsulfide) as well as simple hydrocarbons (methane, cyclopropene). It should, however, be noted, that BNL/cc-pVTZ also shows errors larger 0.5 eV for compounds cyclopropene, ketene, dimethylether, dimethylsulfide, formaldehyde. Excluding problematic cases, where the deviation of LC-DFTB results from experiment is larger than 0.5 eV, we obtain for the remaining 25 compounds the MAE of 0.18 eV, compared to 0.17 eV for BNL/cc-pVTZ for the same subset.

For the problematic cases, where the deviation from both experiment and theory is large, the deviations seem to be an effect of the minimal basis set, which is used in the LC-DFTB method. The basis set of the LC-DFTB method has not been optimized with respect to the compression radii and the standard compression of mio-1-1 set has been used for all parametrizations. In addition, the minimal basis sets in general might not provide enough variational flexibility to describe some chemical environments. The results for the BNL functional at the 3-21G, cc-pVDZ and cc-pVTZ level support this assumption. We find, however, one exception. The BNL/3-21G and BNL/cc-pVTZ results for thiadiazole are nearly the same and underestimate the experimental IP by roughly 0.4 eV, while the LC-DFTB deviates strongly ( $\approx 1.4$  eV) from the experiment. In general one can conclude, that the results of the LC-DFTB method are comparable to LC-DFT with small double-zeta basis (3-21G). The new scheme clearly outperforms first principles DFT calculations based on the PBE and B3LYP functionals as well as the standard DFTB method for the description of ionization potentials. The results of this section suggest that the LC-DFTB reduces the delocalization problem to a considerable extent.

## 7.2 FUNDAMENTAL GAPS

We investigate now the HOMO-LUMO gap from the LC-DFTB. This quantity if obtained from a single ground state calculation is an approximation to the fundamental gap of the system. The fundamental gap can be in principle obtained exactly from two ground state calculations with  $N$  and  $N + 1$  electrons. We note, however, that the present implementation of the LC-DFTB method is valid only for closed-shell systems. Thus a meaningful result for an anion can not be obtained and the extension to the spin-unrestricted formalism is necessary. Nevertheless, the assessment of the LC-DFTB HOMO-LUMO gap from a ground state calculation on neutral closed-shell molecules is an important first step for the characterization of the ability of the LC-DFTB method to predict reliable fundamental gaps.

While the experimental ionization potentials are usually available, the experimental data

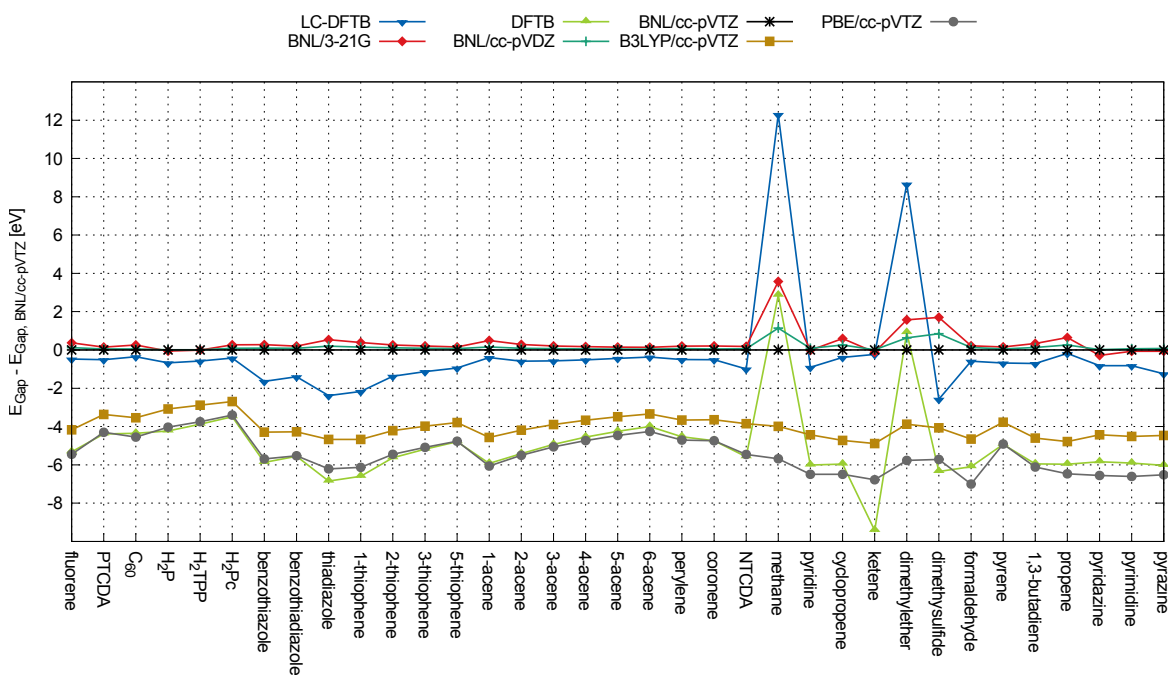


Figure 7.3: The deviation of the HOMO-LUMO gap from the BNL/cc-pVTZ reference for the 35 molecule benchmark set for different theories.

for the electron affinities is sparse. The theoretical reference method of choice in this case is usually a GW variant. It is, however, usually much more expensive than a hybrid DFT. For this particular reason we use the BNL/cc-pVTZ results as reference. They are the most accurate in the benchmark. Moreover, it has been mentioned in the literature, that the long-range corrected functionals (like BNL) usually provide HOMO-LUMO gaps, which are good estimates to the experimental fundamental gaps.

We present the deviations from the reference of the HOMO-LUMO gaps obtained from LC-DFTB, DFTB, PBE/cc-pVTZ, B3LYP/cc-pVTZ, BNL/3-21G and BNL/cc-pVDZ theory levels in Fig. 7.3 and summarize the numerical values in Tab. 7.2. The MAE deviation for LC-DFTB is 1.36 eV, compared to DFTB (5.06 eV), BNL/3-21G (0.41 eV), BNL/cc-pVDZ (0.15 eV), PBE/cc-pVTZ (5.29 eV), and B3LYP/cc-pVTZ (3.92 eV). For LC-DFTB, we find remarkably large deviations for the case of methane and dimethylether of 12.30 eV and 8.66 eV respectively. On the other hand, if we compare the DFTB result to the PBE, we find a remarkable agreement again up to methane, ketene and dimethylether. The general agreement of the DFTB with PBE HOMO-LUMO gap is the manifestation of the fact, that the DFTB has been optimized with respect to PBE. However, the three mentioned compounds seem to be a problem already at the DFTB level. The qualitatively similar, although much smaller deviations of 3.58 eV and 1.57 eV for methane and dimethylether can be also seen for BNL/3-21G. In view of these observations we assign these

Molecule	BNL/3-21G	BNL/cc-pVDZ	BNL/cc-pVTZ	B3LYP/cc-pVTZ	PBE/cc-pVTZ	LC-DFTB	DFTB
Fluorene	9.43	9.18	9.06	4.89	3.61	8.59	3.74
PTCDA	5.74	5.64	5.59	2.22	1.29	5.08	1.20
C <sub>60</sub>	6.41	6.20	6.15	2.61	1.60	5.80	1.80
H <sub>2</sub> P	5.85	5.92	5.91	2.83	1.88	5.24	1.68
H <sub>2</sub> TPP	5.31	5.33	5.32	2.44	1.57	4.74	1.44
H <sub>2</sub> Pc	5.07	4.89	4.80	2.11	1.40	4.38	1.31
Benzothiazole	9.83	9.65	9.55	5.26	3.86	7.91	3.68
Benzothiadiazole	8.24	8.13	8.04	3.77	2.51	6.64	2.49
Thiadiazole	10.64	10.30	10.10	5.43	3.89	7.72	3.25
Thiophene	11.01	10.77	10.62	5.95	4.49	8.45	4.02
2-Thiophene	8.55	8.40	8.29	4.08	2.84	6.92	2.66
3-Thiophene	7.49	7.37	7.28	3.30	2.19	6.15	2.10
5-Thiophene	6.58	6.49	6.42	2.63	1.65	5.48	1.63
1-Acene	11.69	11.36	11.20	6.63	5.14	10.80	5.28
2-Acene	9.19	8.99	8.90	4.71	3.40	8.31	3.49
3-Acene	7.58	7.44	7.37	3.48	2.31	6.80	2.44
4-Acene	6.50	6.38	6.33	2.66	1.60	5.81	1.78
5-Acene	5.73	5.62	5.58	2.09	1.12	5.14	1.33
6-Acene	5.16	5.06	5.02	1.68	0.76	4.65	1.02
Perylene	6.75	6.62	6.55	2.89	1.84	6.05	2.02
Coronene	7.83	7.67	7.62	3.98	2.88	7.11	2.88
NTCDA	7.23	7.11	7.05	3.20	1.59	6.06	1.47
Methane	19.42	16.98	15.84	11.85	10.16	28.14	18.70
Pyridine	10.47	10.58	10.49	6.06	4.00	9.58	4.48
Cyclopropene	11.93	11.59	11.33	6.61	4.84	10.94	5.37
Ketene	10.50	10.63	10.62	5.73	3.84	10.39	1.20
Dimethylether	13.36	12.42	11.79	7.91	6.02	20.45	12.70
Dimethylsulfide	12.49	11.64	10.79	6.72	5.07	8.22	4.44
Formaldehyde	11.01	10.92	10.79	6.14	3.79	10.20	4.69
Pyrene	7.69	7.58	7.53	3.76	2.62	6.85	2.60
1,3-Butadiene	10.44	10.23	10.10	5.50	3.99	9.39	4.15
Propene	12.45	12.06	11.80	7.02	5.33	11.62	5.83
Pyridazine	9.03	9.33	9.31	4.88	2.76	8.49	3.47
Pyrimidine	10.00	10.13	10.07	5.55	3.46	9.25	4.16
Pyrazine	9.66	9.81	9.73	5.27	3.21	8.49	3.70

Table 7.2: HOMO-LUMO gap for the molecules in the benchmark set. All values are in eV.

failures to the minimal basis set. We also note that for the dimethylsulfide compound, which is essentially dimethylether with oxygen being replaced by sulfur, the error is much smaller. A possible reason for this observation is the fact that the sulfur in present parametrization contains additional polarization functions with d-symmetry (compare section 3.4.2). Inclusion of polarization functions for oxygen and nitrogen could possibly reduce the problem.

### 7.3 PHOTOEMISSION SPECTRUM

It has been reported, that the single-particle spectra from theoretical methods, such as GW and hybrid DFT compare well to the experimental photoionization spectra, which are usually obtained by photoelectron spectroscopy. At the same time the single-particle spectra from the local DFT show qualitatively wrong picture of the spectra [38, 123, 178]. The underbinding of the

occupied orbitals beyond the HOMO has a systematical character and has been attributed to the delocalization problem. The long-range corrected (and CAM) functionals at least partially cure these deficiencies and their spectra become qualitatively comparable to the experimental data [38, 99, 161].

In view of this we investigate the occupied eigenvalue spectrum of two molecules, which have been already studied in this context with first principles approaches. Pentacene (5-acene) and perylene-3,4,9,10-tetracarboxylic-dianhydride (PTCDA) are both  $\pi$ -conjugated molecules, which are interesting in photovoltaic applications and molecular electronics. We calculate the eigenvalue spectrum of both molecules with the standard DFTB, the LC-DFTB and with the first principles PBE/cc-pVTZ and BNL/cc-pVTZ methods. The spectra are presented in Fig. 7.4. We use a simple Gaussian broadening profile with the full width at half minimum of 0.1 eV to mimic the experimental resolution and broadening. All spectra have been rigidly shifted such that the HOMO position is at 0 eV. To ensure the comparability with earlier studies, we use the geometries, optimized at the B3LYP/cc-pVTZ level of theory.

The experimental photoemission spectrum of the PTCDA molecule is characterized by the gap of 1.5 eV between the first and the second photoemission peaks. The second peak is observed at the energies between -1.5 eV and -2.1 eV, relative to the first peak, which corresponds to the HOMO [170]. The GW method, the DFT with long-range corrected functionals or CAM functionals usually manage to reproduce these features qualitatively well. This is also confirmed by our BNL/cc-pVTZ calculation. The eigenvalue spectrum is plotted in part a) of Fig. 7.4 and is comparable to the standard LC- $\omega$ PBE result of ref. [99]. As has been already discussed in refs. [99, 161] the local DFT usually fails to reproduce the photoemission spectrum even qualitatively. This can be also observed in our PBE/cc-pVTZ calculation (compare part c) of Fig. 7.4). Specifically, a degenerate  $\sigma$ -state is found right in the middle of the gap between the HOMO and HOMO-1.<sup>1</sup> There is no reason to expect the better performance for the DFTB, which is essentially the approximate DFT with PBE functional. Comparison of part c) and d) of Fig. 7.4 shows, that the DFTB spectrum strongly differs from the PBE spectrum as well.

We keep this finding in mind and take a look at the LC-DFTB spectrum, part b) of Fig. 7.4. As in the case of DFTB, we observe a considerable difference to the BNL/cc-pVTZ spectrum. We find four  $\sigma$ -orbitals, which are mostly located at the anhydride groups of the PTCDA molecule, right in the middle of the gap between HOMO and HOMO-1 states. These orbitals fit into the level ordering scheme of the DFTB. We visualize the level ordering by the colored dashed lines, connecting the peaks and the pictogram with the spatial distribution of the respective orbitals. It seems therefore, that the LC-DFTB just shifts the levels of the DFTB if the range-dependent long-range HF exchange term is added. However, from the calculations on other systems, we find that this shift is usually non-uniform and the level ordering is not preserved (see for example the discussion of the pentacene below). For long-range corrected functionals the transition from

<sup>1</sup>if not stated otherwise, the state numbering is with respect to BNL/cc-pVTZ spectrum.

$\omega = 0$  (DFT limit) to finite  $\omega$  values is generally accompanied by smooth non-uniform shifts of single-particle levels. The occurrence of the level reordering is thus expected. Obviously the BNL/cc-pVTZ theory demonstrates the correct reordering of the single-particle levels with respect to the PBE/cc-pVTZ. At the same time the LC-DFTB fails to reproduce such reordering compared to the DFTB for the case of PTCDA. We attribute this failure to the typical DFTB approximations employed in the LC-DFTB method, such as exploitation of the minimal basis set, the two-center approximation and the integral approximations. This is supported by the fact, that already the DFTB results in Fig. 7.4 d) show strongly underbound  $\sigma$ -orbitals.

However, the level ordering issues have been also observed in the first principles LC-DFT. In general LC-DFT with different short-range functionals and different choice of the range-separation parameter can exhibit distinct level ordering. Körzdörfer et. al [99] employed the concept of orbital many-electron self-interaction error (OMSIE) to gain further understanding. They found, that even if the frontier orbitals are well described by the LC-DFT method, there are states, usually of different symmetry (e.g.  $\sigma$ -orbitals), which show a considerable OMSIE. In particular they discussed the spectrum of the tuned LC- $\omega$ PBE functional. Within the tuning procedure the range-separation parameter is chosen such that the HOMO eigenvalue is equal to the difference of the total energy of the neutral molecule and the total energy of its cation. This procedure allows to non-empirically enforce the condition that the HOMO eigenvalue equals to the ionization potential. For this functional the second peak of the PTCDA spectrum is composed of the degenerate  $\sigma$ -states (which correspond to the HOMO-1/HOMO-2 in the LC-DFTB spectrum in this work). Analysis of the orbital self-interaction error for the tuned LC- $\omega$ PBE showed small OMSIE for the orbitals with  $\pi$ -symmetry and large OMSIE for  $\sigma$ -orbitals. The LC- $\omega$ PBE with standard value of the range-separation parameter on contrast exhibited large OMSIE for  $\pi$ -orbitals, while the  $\sigma$ -orbitals were mostly free of self-interaction error. This suggests, that the level ordering failure of LC-DFTB could also be connected to the residual self-interaction of the  $\sigma$ -orbitals, localized at the anhydride groups. Indeed we observe, that the energy levels relative to HOMO, which correspond to the  $\pi$ -orbitals are quite well represented by the LC-DFTB.

We observe the problematic level ordering also for pentacene spectrum, which is presented in the right panel of Fig. 7.4. In DFT, PBE and BNL calculations the levels up to the HOMO-4 show the same order, while the LC-DFTB spectrum exhibits two  $\sigma$ -orbitals at its HOMO-3 and HOMO-5 positions. These orbitals are additionally indicated by the red and blue dashed lines respectively. Comparing the part g) and h) of Fig. 7.4 we see that already in the DFTB spectrum the position of these orbitals is different with respect to PBE. This indicates the tendency of the approximate theories to underbind these orbitals. In this case the influence of the DFTB approximations is more evident. The preceding analysis leads to the conclusion, that the level ordering issue of the LC-DFTB might be caused by the insufficient description of the  $\sigma$ -states, which is characterized by a considerable remaining orbital self-interaction error on the one hand and the applied DFTB approximations on the other. It is worth mentioning, that the LC-DFTB exhibits significant level

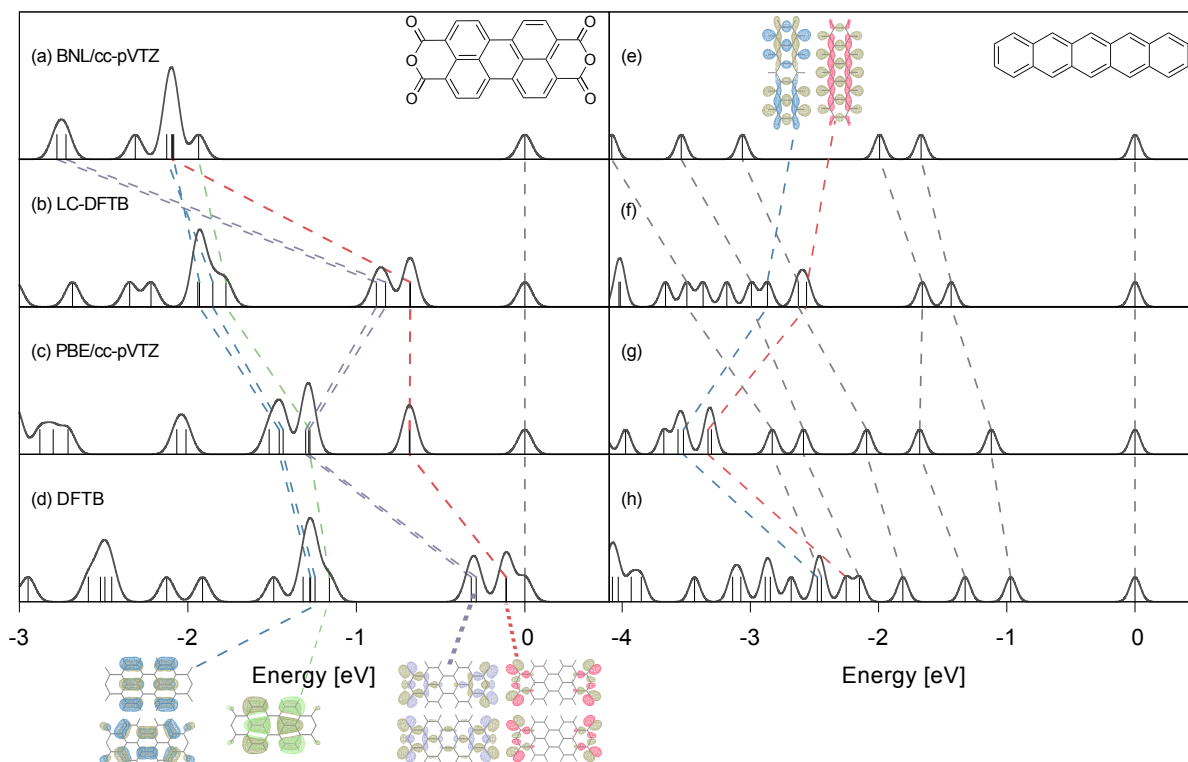


Figure 7.4: Eigenvalue spectrum for pentacene (right panel) and PTCDA (left panel) molecules from LC-DFTB compared to DFTB, BNL/cc-pVTZ and PBE/cc-pVTZ results. The HOMO levels have been shifted to the zero energy for all methods. The Gaussian broadening with FWHM of 0.1 eV was applied.

shifts as compared to the DFTB, which correlates to the behavior of the LC-DFT with respect to the local DFT. Moreover, the description of the levels with  $\pi$ -symmetry such as HOMO, HOMO-1, HOMO-2 is quite accurate. Nevertheless, we conclude that the exploitation of the LC-DFTB method as a tool for a full characterization of photoemission spectra is too ambitious.

#### 7.4 COMMENT ON THE CHOICE OF THE RANGE-SEPARATION PARAMETER

The value of the range-separation parameter, which we use by default in this thesis has been chosen to be  $\omega = 0.3 a_0^{-1}$ . We found this value to give reasonable results for the prediction of ionization potentials. We note, however, that we did not optimize this value with respect to the experimental ionization potentials. The values, close to this are typically used in the standard long-range corrected and CAM functionals. For example in the software package NWCHEM [195] the value is  $\omega = 0.33$  for CAM-B3LYP,  $\omega = 0.33$  for LC-BLYP,  $\omega = 0.3$  for LC-PBE and  $\omega = 0.33$  for BNL functionals. These functionals, however, are all based on the range separation with the error-function. The complementary error-function ( $\text{erfc}(x) = 1 - \text{erf}(x)$ ), which defines the short-range, decays sharper than the exponential function of the Yukawa-type range separation as can



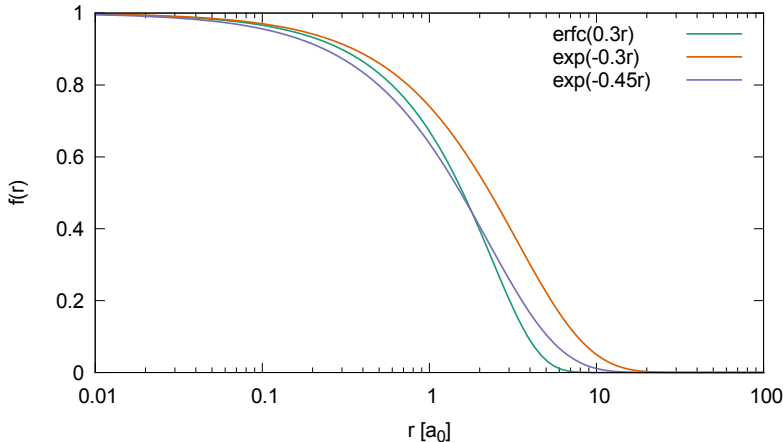


Figure 7.5: Complementary error-function  $\text{erfc}(\omega r)$  with decay constant  $\omega = 0.3a_0^{-1}$  compared to the exponential function  $\exp(-\omega r)$  with  $\omega = 0.3a_0^{-1}$  and  $\omega = 0.45a_0^{-1}$ . The erfc-function decays sharper than the exponential function for the same decay constant.

be seen in Fig. 7.5 (compare also total energies of the atoms as a function of range-separation parameter in Fig. 4.1). Thus it is expected, that the optimal values of the range-separation parameter for the case of Yukawa range separation should be higher than the corresponding values in the error-function-based functionals. In fact, this has been confirmed by Akinaga and Ten-no [2]. The optimization of the parameters in their work was carried out with respect to the atomization energies of the G2-1 benchmark set with the cc-pVTZ basis for different functionals. It is, however, not clear, whether these parameter values can be directly transferred to the LC-DFTB. From our calculations we found, that values ranging from  $\omega = 0.5a_0^{-1}$  to  $\omega = 0.75a_0^{-1}$ , suggested in their paper for pure long-range corrected functionals with Yukawa range separation, result in systematic overestimation of the ionization potentials, which are calculated from the HOMO eigenvalues and fundamental gaps, obtained from the HOMO-LUMO eigenvalue difference. However, for the CAMY-B3LYP functional (CAM-B3LYP [200] with Yukawa range separation), which contains a constant fraction of exact exchange, the optimal value was found to be  $\omega = 0.34a_0^{-1}$ . As has been outlined in section 3.8.2, the long-range  $\gamma$ -integral might effectively include the effect of xc-functional through the atomic decay constant  $\tau$ . Thus it is not granted, that the xc-functional in the LC-DFTB method is a pure long-range corrected functional.

The optimization of all standard range-separated functionals is usually carried out with respect to thermochemical data and atomization energies. Different ways of optimization and choice of underlying local xc-functionals result in a different optimal parameter. All these methodics rely on the total energy differences of the methods. A similar, consistent procedure for LC-DFTB would therefore require a high quality fit of repulsive potentials, which is not available at the moment. For this reason the optimization of the range-separation parameter has not yet been done. Nevertheless, we show the mean absolute error (MAE) and mean signed error (MSE) for the HOMO eigenvalue from LC-DFTB compared to experimental ionization potential as a function of range-separation parameter  $\omega$  for compounds used in the molecular benchmark set in Fig. 7.6. This should give a feeling for a potential predictive power of the LC-DFTB method with respect to

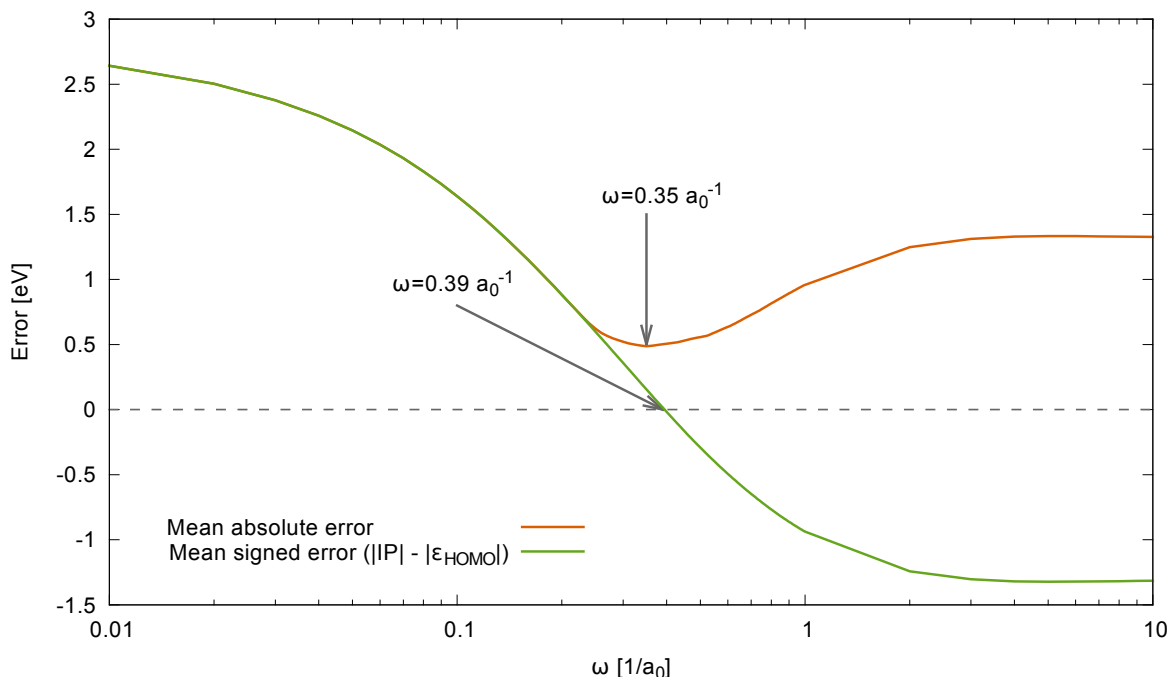


Figure 7.6: The mean absolute error and mean signed error of the HOMO eigenvalue from LC-DFTB compared to the experimental IP for the compounds of the benchmark set as a function of range-separation parameter  $\omega$ .

the experimental ionization potentials. We sample the values on a grid with step  $\Delta\omega = 0.01a_0^{-1}$  for  $0.01a_0^{-1} \leq \omega \leq 1.0a_0^{-1}$  and  $\Delta\omega = 1.0a_0^{-1}$  for  $\omega > 1.0a_0^{-1}$ . The MAE shows minimum at  $\omega = 0.35a_0^{-1}$ , while the mean signed error is minimal for  $\omega = 0.39a_0^{-1}$ . We also observe the aforementioned fact of overestimation of the IP by the LC-DFTB for values in the range  $\omega = 0.5a_0^{-1}$  to  $\omega = 0.75a_0^{-1}$ . For the values  $\omega = 0.3a_0^{-1}$  and  $\omega = 0.5a_0^{-1}$  the MAE are 0.48 eV and 0.55 eV respectively which are similar. The larger values lead to increasing MAE, for example for  $\omega = 0.8a_0^{-1}$  the MAE grows to 0.82 eV. We conclude, that the choice  $\omega = 0.3a_0^{-1}$  is reasonable, although not optimal for the IP prediction.

## 7.5 EXECUTION TIMES

The predictive power is not the only interesting characteristic of a computational method. The development of the LC-DFTB method in addition to the existing LC-DFT methods have been mainly motivated by the requirement of performing large-scale calculations which are too expensive for the first-principles LC-DFT implementations.

In section 3.9 it has been shown, that the Hamiltonian construction in the present LC-DFTB implementation exhibits quadratic scaling with respect to the basis size. For the tested systems (< 1000 atoms) this step is the bottleneck of the entire calculation, which lowers the computational

Molecule	BNL/3-21G	BNL/cc-pVDZ	BNL/cc-pVTZ	LC-DFTB <sup>1</sup>	LC-DFTB <sup>2</sup>	DFTB
5-Acene	325	2396	59701	11	6	1
Perylene	541	4447	113414	12	6	1
H <sub>2</sub> P	932	7486	23825*	16	8	1
Coronene	2116	15983	260677	14	6	1
6-Acene	507	3516	79993	15	8	2
5-Thiophene	1303	11581	144735	16	11	2
PTCDA	3748	20477	521424	22	10	2
H <sub>2</sub> Pc	3034	27231	366838*	39	14	2
H <sub>2</sub> TPP	5735	43304	744967*	72	22	4
C <sub>60</sub>	7221	65393	655789*	121	23	2

Table 7.3: Wall time [sec] of a single point calculation of LC-DFTB, DFTB and long-range corrected DFT with different basis sets for molecules with more than 30 atoms from the benchmark set, considered in this chapter. The asterisk denotes the parallel jobs on 12 CPUs, LC-DFTB<sup>1</sup> was performed with  $\epsilon_{\text{threshold}} = 10^{-16}$  and LC-DFTB<sup>2</sup> with  $\epsilon_{\text{threshold}} = 10^{-6}$ .

efficiency of the LC-DFTB method with respect to the standard DFTB. Here we present the absolute timings for selected molecules from LC-DFTB, DFTB and BNL with 3-21G, cc-pVDZ and cc-pVTZ basis sets. The results for molecules with more than 30 atoms from the benchmark set are presented in Tab. 7.3. The LC-DFTB calculations have been performed with the thresholding algorithm and  $\epsilon_{\text{threshold}} = 10^{-16}$  (denoted by LC-DFTB<sup>1</sup>) and  $\epsilon_{\text{threshold}} = 10^{-6}$  (denoted by LC-DFTB<sup>2</sup>) on a single core of an Intel Core-i7 CPU. We note, that for the threshold parameter  $\epsilon_{\text{threshold}} = 10^{-6}$  the eigenvalues deviate from the exact calculation by less than  $10^{-5}$  eV (compare inset of Fig. 3.9). Thus we consider this value as meaningful for practical calculations. The DFTB calculations have been performed on the same machine as LC-DFTB. The execution time in both cases has been measured by the Linux *time* utility. The DFT calculations have been carried out using the NWCHEM package in serial and parallel versions. The serial version was executed on Intel Xeon 2.8GHz machines. The parallel calculations have been distributed over the 12 CPUs on a cluster and are denoted by asterisk.

The first-principles methods require, as expected, considerable computational time. For example 12 CPU parallel single point calculation of the buckminsterfullerene on the BNL/cc-pVTZ level of theory required 7.6 days. This should be compared to 23 seconds for the LC-DFTB<sup>2</sup> on a *single* core. Even the calculation with a small basis set (BNL/3-21G) turns out to be at least 30 times slower than the LC-DFTB<sup>1</sup> and 50 times slower than LC-DFTB<sup>2</sup> for smaller molecules. For larger systems, the quadratic scaling<sup>2</sup> of the LC-DFTB results in the increase of the gap in computational times between the LC-DFTB and first-principles calculations. The standard DFTB is usually an order of magnitude faster than LC-DFTB<sup>2</sup>.

<sup>2</sup>the systems, for which the cubic scaling due to the diagonalization in the LC-DFTB method would dominate, are hardly accessible in a usual way by the first-principles calculations.

## 7.6 SUMMARY

We presented benchmark ground state calculations for a set of 35 organic molecules with the LC-DFTB method. We investigated the frontier orbitals (HOMO and LUMO) of the LC-DFTB, standard DFTB and first principles approaches which included the local PBE/cc-pVTZ, hybrid B3LYP/cc-pVTZ and long-range corrected BNL/cc-pVTZ, BNL/cc-pVDZ and BNL/3-21G level of theory.

The negative of the HOMO eigenvalues, which for the exact exchange-correlation functional should equal the ionization potential has been compared to the experimental data. The mean absolute deviation of 0.50 eV for LC-DFTB method is comparable to the result of the long-range corrected BNL/3-21G, which is 0.67 eV. At the same time the deviations for standard DFTB (2.50 eV) and the local PBE/cc-pVTZ (2.87 eV) are much larger. We consider this finding as clear indication of the reduction of delocalization problem in the LC-DFTB method. Indirectly it suggests, that the LC-DFTB method tends to satisfy the linearity condition (section 2.1). However, to confirm this directly (as for example in the calculation, corresponding to the results of Fig. 2.2) the extension to the spin-unrestricted formalism is necessary.

The fundamental gaps from the HOMO-LUMO difference have been compared to the BNL/cc-pVTZ level of theory since the experimental values of electron affinities are in general not available. With mean absolute error of 1.36 eV the LC-DFTB shows significantly better agreement with the reference than the local PBE/cc-pVTZ (5.29 eV), hybrid B3LYP/cc-pVTZ (3.92 eV) and standard DFTB (5.06 eV). At the same time the first principles BNL/3-21G shows the mean absolute error of 0.41 eV, which is more than three times smaller than the LC-DFTB error. We mention also occurrence of remarkable outliers, such as methane and dimethylether for which we find huge deviations of 12.30 eV and 8.66 eV for LC-DFTB. We attribute this behavior to the minimal basis set, used in the LC-DFTB.

We compared the full eigenvalue spectrum of LC-DFTB to the photoemission spectra of PTCDA and pentacene molecule, which are well studied in this context. The single particle energies beyond the HOMO show at first glance rather unsatisfactory results. Near inspection shows, however, that the  $\sigma$ -orbitals, generally problematic for DFTB are notoriously underbound by LC-DFTB as well. At the same time  $\pi$ -orbitals possess energies, similar to the first-principles calculations.

The computational cost of the current implementation of the LC-DFTB method is usually 10-20 times larger than that of standard DFTB. At the same time the BNL/3-21G calculations are still order of magnitude slower even for the small-sized systems, considered in this benchmark. The better scaling of the LC-DFTB method with the basis size will make the difference in the computational time for larger systems more significant.

The usual application scope of the approximate methods such as DFTB are extended systems. The  $\pi$ -conjugated polymers constitute an important class of such materials, which are important for applications such as field effect transistors [40, 51, 61], light-emitting diodes [12, 127, 151] and photovoltaic cells [23, 75]. Disappointing is the fact, that the local DFT and in some cases global hybrid DFT fail to reproduce the physics of these systems even qualitatively. Typical is also the fact that the failures of local DFT are usually stronger emphasized if the system size increases. Even in the case of a single extended molecule, where the intermolecular or surface-molecule van-der-Waals interactions, which are the weak point of local DFT, are excluded, the problems are not negligible. Specifically, local DFT fails to reproduce the correct bond length alternation (BLA), which is the difference between the lengths of single and double bonds. It is usually underestimated by the local DFT, while the Hartree-Fock theory systematically overestimates it [25, 71, 82–84]. The BLA is connected to other electronic properties of the conjugated molecular chains, such as band gap [24], linear and non-linear response to electric fields [22, 121], Raman spectra [122, 190] and the description of the localized charge defect formation. While for example the response to the electric field in the local DFT is drastically overestimated, the description of polaronic states is even qualitatively wrong. The local DFT predicts no polaron formation [133, 134], which contradicts the results from the HF theory and hybrid B3LYP calculations [27, 204]. All these flaws are naturally inherited by the DFTB method [44].

Körzdörfer et al. established the connection between BLA and self-interaction error [100], which indicates the decisive role of the delocalization problem in the erroneous description of the conjugated molecular chains. The long-range corrected functionals, which as we know successfully deal with the delocalization problem, provide an improved description of BLA [82, 85]. It has

n	This work				Ref. [85]			
	LC-DFTB	DFTB	HF	PBE	HF	PBE	B3LYP	LC- $\omega$ PBE
2	15.4	11.3	-	-	14.5	10.5	11.7	12.6
4	12.3	7.6	-	-	12.8	6.8	8.5	10.6
6	11.6	6.2	-	-	12.4	5.3	7.3	10.1
8	11.3	5.5	-	-	12.3	4.4	6.7	9.9
10	11.2	4.9	-	-	12.2	3.9	6.3	9.8
12	11.2	4.7	-	-	12.2	3.5	6.0	9.8
14	11.2	4.4	-	-	12.2	3.2	5.9	9.8
40	11.1	3.6	12.4	2.0	-	-	-	-

Table 8.1: Bond length alternation in the central monomer unit of a *trans*-polyacetylene chain,  $n$  denotes the number of  $C_2H_2$  units in the chain. All results are given in pm.

been further shown, that the response properties, such as static longitudinal polarizabilities are considerably better described as compared to the local DFT. We are interested thus in the performance of the LC-DFTB for this class of problems. We apply the LC-DFTB to polyacetylene (PA) oligomers in thermodynamically stable *trans* configuration. This is a simple, well-studied  $\pi$ -conjugated system [28, 117, 186]. Since the parametrization of repulsive potentials has been performed for hydrogen and carbon species only (chapter 5), we can only treat hydrocarbons at the moment. Small number of atoms per monomer unit makes this system further accessible for the first principles methods if longer chains are investigated. We begin with the bond length alternation of the PA chains in section 8.1. In section 8.2 we show, that the LC-DFTB predicts a formation of stable localized charge defects, known as bipolarons in the doped PA. The response to electric field, specifically the static longitudinal polarizability as obtained by LC-DFTB is discussed in section 8.3. Finally, in section 8.4, we briefly comment on the LC-DFTB calculations on proteins in zwitterionic conformation in the gas-phase. The contents of sections 8.3 and 8.4 are part of ref. [119].

## 8.1 BOND LENGTH ALTERNATION

The experimental value of the BLA for *trans*-PA has been found to be  $\Delta r = 8 \pm 3$  pm [49, 202]. In refs. [71, 84, 85] the BLA from the local DFT, DFT with hybrid functionals and wave function-based methods has been already discussed. The key finding is that the local DFT tends to underestimate the BLA by a factor of 2, while the Hartree-Fock theory overshoots the experimental value by a similar factor. The global hybrid functionals and post-Hartree-Fock methods (MP2, CCSD) give rather reasonable values for the BLA at least for short oligomers. The study in ref. [85] with basis set 6-31G(d) included among others the long-range corrected functionals LC-BLYP, LC- $\omega$ PBE and the CAM functional CAM-B3LYP [200]. They performed very well, providing BLA values ranging from 9 pm (CAM-B3LYP) to 10 pm (LC-BLYP), which are well within the

experimental error. We compare the results of ref. [85] for HF, PBE, B3LYP and LC- $\omega$ PBE to the results from the LC-DFTB with the repulsive potential from chapter 5 and the standard DFTB. The BLA values for *trans*-PA with different number  $n$  of monomers are summarized in Tab. 8.1.

For the LC-DFTB and standard DFTB methods we optimize the geometries for oligomers with number of monomer units ranging from  $n = 2$  to  $n = 14$ . In addition, we perform the geometry optimization of a  $n = 40$  oligomer on LC-DFTB, DFTB, HF/6-311G\* and PBE/6-311G\* level of theory (these geometries are then used in the next sections). The LC-DFTB predicts the BLA of  $n = 14$  oligomer to be 11.2 pm, while for the largest oligomer ( $n = 40$ ) we obtain 11.1 pm. Compared to the LC- $\omega$ PBE at  $n = 14$  this value is higher by 1.3 pm. We should note, that for LC- $\omega$ PBE the BLA value seems to saturate for  $n = 10 - 14$ , thus we do not expect a significant decrease in BLA for larger oligomers at this theory level. The BLA ( $n = 40$ ) for LC-DFTB, however, is still smaller than the HF value, which in our calculation ( $n = 40$ ) is 12.4 pm and in the calculation of ref. [85] 12.2 pm for  $n = 14$ . The discrepancy may come from different basis sets (we use larger 6-311G\*, while Jacquemin et al. use the 6-31G(d) set) or slightly distinct convergence criteria. Important is the fact, that the BLA, as obtained from LC-DFTB saturates for longer chains. This is in accord with LC-PBE and HF methods. The local PBE functional on contrast does not show such a saturation at the oligomers with  $n = 10 - 14$ , while the LC-PBE, LC-DFTB and HF already do. Furthermore, from our PBE calculation on the oligomer with  $n = 40$  we obtain the BLA value of only 2.0 pm. This clearly shows the tendency of the local DFT to equilibrate the single and double bond lengths, which increases with the oligomer chain length. We notice similar behavior for the standard DFTB method and the hybrid DFT with B3LYP functional.

We conclude, that the LC-DFTB shows qualitative improvement over the standard DFTB and DFT with local and global hybrid functionals. The observed systematic overestimation of the BLA by LC-DFTB can be possibly cured by providing a more careful repulsive potential parametrization. In such a case the LC-DFTB method can become a valuable tool for computationally efficient optimization of geometries in the systems, where standard DFTB and the first principles approaches fail.

## 8.2 EMERGING POLARON SIGNATURES

The next problematic case for the local DFT and DFTB, which is related to the BLA underestimation is the class of doped  $\pi$ -conjugated polymers. The doping in these systems leads to the conductance, in a similar way as in a semiconductor. The dopant acts as an electron acceptor or donor. In contrast to semiconductors, where the hole in the valence band (electron in the conduction band) lead to the transport, in the oligomer chains the charge transport is usually explained by a propagation of quasi-particles, which are called polarons, bipolarons or polaron pairs. The polaronic state is stabilized by the electron-phonon coupling and is characterized by a local charge accumulation and the corresponding lattice distortion around it. The state appears in the gap of the neutral species.

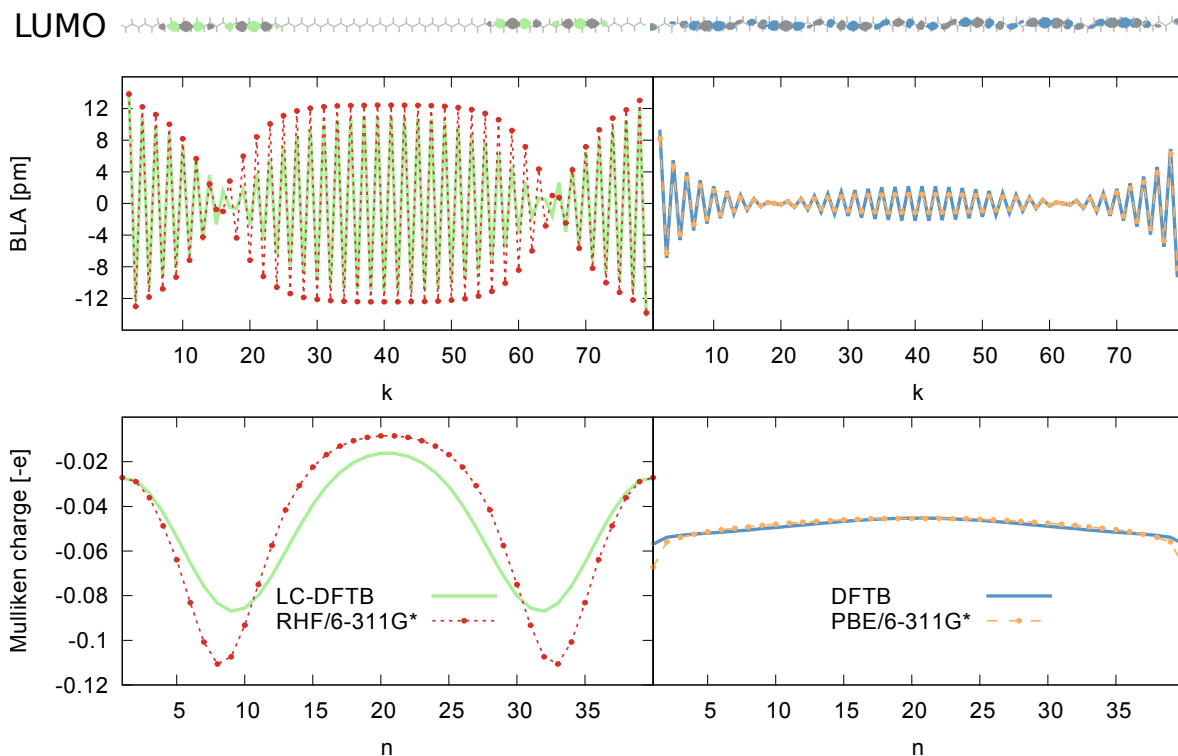


Figure 8.1: The BLA and Mulliken charge distribution of a 40-acetylene<sup>2+</sup> oligomer dication for different theories. *Bottom part:* excess Mulliken charge  $q_{\text{dication}} - q_{\text{neutral}}$  in units of electron charge [-e] for the traditional DFTB and local DFT at PBE/6-311G\* level (right panel), for the LC-DFTB and RHF/6-311G\* (left panel) as a function of the monomer position  $n$  on the oligomer chain. *Top part:* bond length alternation for all theories as a function of bond pair number  $k$ . *On the very top:* the spatial distribution of LUMO orbitals for LC-DFTB (left) and DFTB (right).

The local DFT predicts no polaron formation in polymer chains. This failure has been observed for different systems, such as polyacetylene, polythiophene and polyphenylene-vinylene. The geometry relaxation on charged systems leads to nearly uniform bond length alternation patterns and the excess charge is uniformly distributed over the whole chain. A single point calculation on a geometry, which exhibits the lattice distortion due to the polaron formation (for example resulting from a geometry relaxation on the restricted Hartree-Fock (RHF) level of theory) does not improve the result.

The traditional DFTB also exhibits significant problems with the description of localized polaronic states. Niehaus et al. [138] addressed this failure for polythiophene dications and polyphenylene-vinylene dications. They found that the self-consistent DFTB predicts no localized polaron state in the band gap. The analysis of the charge distribution and bond length alternation patterns also showed no signatures of polaronic state. On the other hand, the non-self-consistent (zeroth-order) DFTB was able to give qualitatively correct picture. This has been attributed to the fact, that by construction the zeroth-order DFTB tends to artificially localize the states.



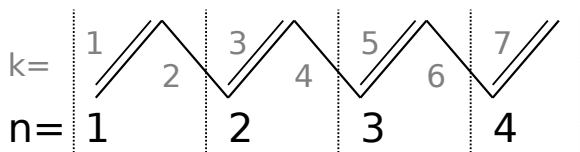


Figure 8.2: Sketch of the PA-geometry. The vertical dashed lines separate the  $C_2H_2$  monomer units, which are enumerated by the monomer position number in the chain  $n$ . The C-C bonds  $k$  are enumerated from left to right.

We apply the LC-DFTB method to a (p-)doped *trans*-PA in order to investigate its ability to predict a polaronic state in this system. The doping process causes the oxidation of the polymer molecule. This is modeled by reducing the number of electrons by 2 and we obtain a dication.<sup>1</sup> In this case two well-separated charges, which form a bipolaron emerge [26]. The degenerate defect state is located in the gap of the neutral species.

We optimize the geometry of a 40-acetylene<sup>2+</sup> oligomer dication on the LC-DFTB, DFTB and first principles PBE/6-311G\* and RHF/6-311G\* theory levels and investigate first the bond length alternation (BLA) patterns of the relaxed geometries. The BLA  $\Delta r_k = r_k - r_{k-1}$ ,  $2 \leq k \leq 79$ , where  $r_k$  is the bond length of the  $k$ -th C-C bond (compare Fig. 8.2), is presented in the top part of Fig 8.1. In the top left panel we show the RHF/6-311G\* and LC-DFTB result and in the right panel the result of the PBE/6-311G\* and DFTB methods. In all cases the pattern is irregular and shows minima of BLA at  $k = 16$  and  $k = 64$  (for LC-DFTB and RHF/6-311G\*) and at  $k = 18$  and  $k = 62$  (for DFTB and PBE/6-311G\*). This corresponds to the monomers  $n = 8$  and  $n = 32$  for LC-DFTB and RHF/6-311G\* ( $n = 9$  and  $n = 31$  for DFTB and PBE/6-311G\*). However, the absolute value of the BLA as well as its relative change is too small for the DFTB and PBE/6-311G\* methods. On contrast, LC-DFTB and RHF/6-311G\* show very pronounced changes in BLA pattern.

Nearly uniform BLA pattern for the methods based on the local xc-functionals results in the nearly uniform distribution of the excess Mulliken charge per monomer unit. For DFTB and PBE/6-311G\* it is plotted in the bottom right part of Fig. 8.1 as a function of the monomer position number  $n$  on the chain of the 40-acetylene<sup>2+</sup> (compare Fig. 8.2). This indicates the tendency to delocalize the charge density. The LC-DFTB and RHF/6-311G\* on contrast show two clearly defined minima at the monomers, where the lattice distortion occurs. Further inspection of molecular orbitals shows that the LUMO state of the dication, calculated on the LC-DFTB theory level, is localized at that positions (very top left of Fig. 8.1). The DFTB predicts LUMO which is delocalized over the whole chain (very top right of Fig. 8.1).

In Tab. 8.2 we summarize the values of the HOMO and LUMO eigenvalues of the neutral species and HOMO, LUMO and LUMO+1 eigenvalues of the dication. As can be seen, the LC-DFTB and RHF methods predict the appearance of the (nearly) degenerate LUMO and LUMO+1 states in the middle of the gap of the neutral species. This is not observed for the local theories. From other calculations (e.g. on polythiophene dications) we also see, that even without a proper geometry relaxation (with or without the lattice distortion due to the polarons) the LC-DFTB shows tendency to form a localized state, which is reflected in a local charge accumulation. Thus

<sup>1</sup>at the moment only closed-shell molecules can be reasonably described by the LC-DFTB.

Theory	Neutral			Dication		
	HOMO	LUMO	Gap	HOMO	LUMO	LUMO + 1
RHF/6-311G*	-6.39	0.90	7.29	-7.89	-3.46	-3.46
PBE/6-311G*	-3.98	-3.60	0.38	-6.39	-6.14	-5.93
DFTB	-4.68	-4.12	0.56	-7.06	-6.78	-6.63
LC-DFTB	-6.76	-1.84	4.92	-8.49	-5.85	-5.84

Table 8.2: Frontier orbitals of neutral 40-acetylene oligomer and its dication for different theories. All values are in eV.

we conclude, that the LC-DFTB is able to describe bipolaronic states in the conjugated polymers as opposed to the standard DFTB and DFT with local xc-functionals.

### 8.3 RESPONSE TO ELECTRIC FIELD

The problem of the disproportionate response of the local DFT to an applied electric field is well known in the literature. It has been attributed to the lack of a necessary non-local response term in the exchange-correlation functional [55, 92, 196]. It follows, that all local exchange-correlation functionals fail to produce the correct induced field, which counteracts an applied electric field. A major consequence of this fact is a wrong charge distribution, which exhibits a too strong separation of the induced charge. This leads to the strong overestimation of static polarizabilities. Remarkable is the fact, that this overestimation increases for larger systems. The failures get even more pronounced for the hyperpolarizability and second hyperpolarizability [31]. In this context it is worth mentioning, that this problem has also consequences for the important field of molecular electronics. The lack of the field-counteracting term, the delocalization of the density and the underestimated HOMO-LUMO gap lead to the erroneous description of transport in molecular systems [114].

It has been already shown in chapter 7, that the LC-DFTB essentially improves the description of the fundamental gap, compared to the DFTB and local DFT. This result suggests, that the delocalization problem is reduced. Indeed, in the discussion of the polarons in *trans*-PA, we observed the ability of the LC-DFTB to correctly describe the localized states. Recently Sekino et al. provided evidence for the ability of the LC-DFT to overcome the field response problem [176]. Due to the inclusion of the non-local range-dependent term in the LC-DFTB we expect the signatures of the field-counteracting term to show up in a similar way.

To prove this, we calculate static longitudinal polarizabilities of *trans*-polyacetylene chains with varying number of monomer units ( $C_2H_2$ ) and inspect the induced charge distribution along the chain using Mulliken population analysis. In the LC-DFTB and standard DFTB methods the

$\omega[a_0^{-1}]$	PA ( $n = 10$ )			PA ( $n = 40$ )	
	BNL/6-311G**	BNL/3-21G	LC-DFTB	BNL/3-21G	LC-DFTB
2.0	1288	1229	1156	9341	9624
0.8	1212	1138	1200	8258	10108
0.5	1193	1102	1256	7914	10929
0.3	1321	1215	1345	9572	12547
0.2	1513	1400	1429	12795	14403
0.1	1809	1696	1560	21002	18393
$10^{-2}$	2058	1938	1698	39665	28484
$10^{-3}$	2059	1939	1701	40318	29087

Table 8.3: Static longitudinal polarizability of PA ( $n = 10$ ) and PA ( $n = 40$ ) for different values of the range-separation parameter  $\omega$ . All values are in atomic units.

electric field  $\mathbf{F}$  is included via the additional term

$$E_{\text{field}} = - \sum_A \Delta q_A \mathbf{F} \cdot \mathbf{R}_A \quad (8.1)$$

in the total energy functional, where  $\Delta q_A = q_A - q_A^0$  is the difference Mulliken charge (Eq. 1.24) and  $\mathbf{R}_A$  are the atomic positions. Because DFTB and LC-DFTB use the minimal basis set, their performance in predicting accurate absolute values for the polarizabilities is in general poor. However, our calculations show, that even the LC-DFT with minimal basis (STO-3G) tends to correctly reduce the polarizability with respect to the local DFT in the same basis.

The polarizabilities, obtained from the LC-DFTB method are compared to the first principles long-range corrected DFT at the BNL/6-311G\*\* (larger basis) and BNL/3-21G (smaller basis) levels. The polarizabilities within the LC-DFT theory have been computed by solving the coupled-perturbed Kohn-Sham equations (CPKS) using the algorithms, implemented in the NWCHEM package. The geometries for the *trans*-polyacetylene (PA) with number of monomers  $n = 10$  and  $n = 40$  have been optimized at the B3LYP/6-311G\* level. For the LC-DFTB we obtain the polarizabilities using the finite field method. In this approach, the numerical derivative of the longitudinal component of the dipole moment  $\mu$  with respect to the perturbing electric field  $F$  is calculated with the center difference formula  $\alpha = (\mu(F) - \mu(-F))/2F$ , where the field strength was chosen to be  $F = 4 \cdot 10^{-4}$  [au].

In Tab. 8.3 the static longitudinal polarizability of both PA oligomers with  $n = 10$  and  $n = 40$  is presented for different values of the range-separation parameter  $\omega$ . The LC-DFTB and LC-DFT show similar qualitative behavior, although the quantitative differences are rather large. All three theories exhibit larger polarizabilities in the local DFT limit ( $\omega \rightarrow 0$ ) than in the opposite HF+c limit ( $\omega \rightarrow \infty$ ), where the xc-functional is composed of local DFT correlation functional only in addition to the full (unscreened) HF exchange. The rapid drop of the polarizability as the range-separation parameter is increased is characteristic for both the first-principles approach and the approximate LC-DFTB method (compare also main part of Fig. 8.3 for visual impression).

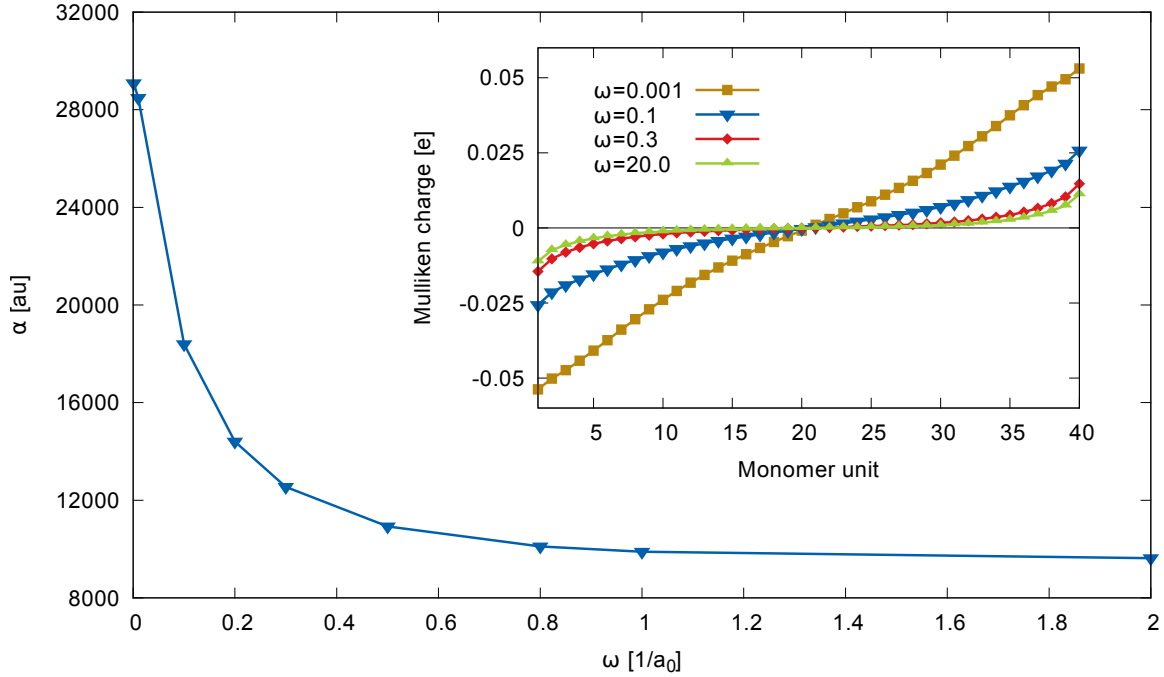


Figure 8.3: Static longitudinal polarizability of the PA ( $n = 40$ ) oligomer as a function of range-separation parameter  $\omega$  from LC-DFTB method. The induced Mulliken charge per monomer unit due to the applied electric field  $F = 10^{-3}$  [au] for different values of  $\omega$  [ $a_0^{-1}$ ] is shown in the inset.

In the case  $n = 10$  the ratio of HF+c to local DFT limits is 0.63 for both BNL/6-311G\*\* and BNL/3-21G and 0.68 for LC-DFTB method. It decreases to 0.23 for BNL/3-21G and 0.33 for LC-DFTB for the larger system with  $n = 40$  units. This indicates the aforementioned tendency of local functionals (DFT limit) to strongly overestimate the polarizability if the system size increases.

We visualize the charge distribution to obtain more complete picture. In the inset of Fig. 8.3 the induced Mulliken charge due to the applied electric field of magnitude  $F = 10^{-3}$  [au] along the longitudinal axis of the  $n = 40$  oligomer for different values of the range-separation parameter  $\omega$  is plotted. We observe almost linear charge distribution for the LC-DFTB in the limit  $\omega \rightarrow 0$ , which indicates the typical overpolarization of the local theories. We note, that the LC-DFTB in this limit is nearly identical to the DFTB, despite the fact that the xc-functional is slightly different. For this reason, we omit the discussion of DFTB for brevity. Increasing the parameter  $\omega$  and thus introducing the range-dependent non-local exchange term gives rise to an effective screening of the electric field. This leads to the correction of the polarizabilities towards more physical values. The polarizability for the LC-DFTB as a function of the range-separation parameter is plotted in the main part of Fig. 8.3 for comparison. It should be again noted, that the results, presented here are of qualitative value. The quantitatively correct polarizabilities require

	Chignolin			Trp-cage		
	HOMO	LUMO	Gap	HOMO	LUMO	Gap
LC-DFTB ( $\omega = 0.3 a_0^{-1}$ )	-3.15	0.63	3.78	-3.66	-1.18	2.48
LC-DFTB ( $\omega \rightarrow \infty$ )	-5.90	2.14	8.04	-6.14	1.32	7.46
LC-PBE/3-21G	-3.78	0.34	4.12	-4.46	-1.72	2.74
RHF/3-21G	-5.42	1.81	7.23	-6.12	-0.13	5.99
	Time [sec]			Time [sec]		
LC-DFTB ( $\omega = 0.3 a_0^{-1}$ )	73			882		
LC-DFTB ( $\omega \rightarrow \infty$ )	78			331		
LC-PBE/3-21G	8841			36844		
RHF/3-21G	1577			7534		

Table 8.4: Frontier orbital energies and fundamental gap (all in eV) of the chignolin and Trp-cage zwitterions in the gas-phase for different theories (top part). The overall execution time in [sec] (bottom part).

in general large basis sets. An alternative approach, which can be used in the approximate methods like LC-DFTB is the exploitation of the empirical correction methods [88, 89].

#### 8.4 PROTEINS IN GAS-PHASE

The properties of proteins are to a great extent determined by their folding structure. The experimental determination of the protein folding structure is usually performed in its native environment, where it exists in a solvated form. In recent years the experimental techniques have been developed, which allow to non-destructively extract the proteins from the solution and in combination with structural analysis to study the intramolecular interactions in the gas-phase [180]. Important methods in this respect are the electrospray ionization (ESI) [48], mass spectrometry and ion mobility measurements. Also, recent developments in the field of diffractive imaging with soft-X-ray free-electron lasers may permit the structural determination of single molecules at atomic resolutions in the gas-phase [32, 126]. On the side of computational methods the effective solvent models in combination with gas-phase calculations may provide a way to understand the protein folding mechanisms in different environments.

In this context the question whether peptides adopt the zwitterionic form, known from aqueous solution, also in the gas-phase is still not definitely answered [120, 156, 187]. The local DFT and thus the DFTB experience considerable difficulties in the description of the zwitterionic state, where the long-range charge-charge interactions are important. In a recent study on the model peptides chignolin [74] and Trp-cage [136] in their zwitterionic conformation in the gas-phase Nishimoto et al. [141] found that the DFTB self-consistency procedure failed to converge. The experimental results show, however, that the native zwitterionic configuration of Trp-cage remains stable in the gas-phase. This finding is supported by the force field-based molecular

dynamics calculations [94, 144].

The reason for this flawed behavior is the notorious underestimation of the HOMO-LUMO gap by the DFTB method, which is connected to the delocalization problem. The possible technical workaround is to rise the electronic temperature, which results in the converged calculation. However, this way the actual problem is not solved. In this context the application of the LC-DFTB method, which naturally reduces the delocalization problem, becomes interesting. We present the LC-DFTB single point calculations on the chignolin and Trp-cage structures and compare to the first-principles LC- $\omega$ PBE/3-21G and RHF/3-21G methods. For both proteins we use the geometrical structures, deposited in the protein data base (PDB). They are obtained by NMR measurements and constitute the native zwitterionic conformations of the proteins [74, 136]. For the gas-phase simulations we require the systems to be charge-neutral. To achieve this, we restore the basic and acidic side residuals to their neutral form by appropriate protonation or deprotonation. On the affected residuals the carboxy and amino groups have been locally relaxed at the DFTB level (with electronic temperature set to  $T=500$  K), while the rest of the atoms have been kept fixed. The charges of carboxy-terminus and amino-terminus have been preserved. A ribbon representation of both proteins is given in Fig. 8.4.

We performed the LC-DFT and RHF calculations with the parallel version of NWCHEM on 8 CPUs. The LC-DFTB calculations were done on a single core of an Intel Core i7 CPU. We summarize the eigenvalues of the frontier orbitals and the gap for both systems in Tab. 8.4. We again find a good agreement in the description of HOMO and LUMO levels by LC-DFTB as compared to the first-principles methods with small basis. We also show corresponding timings in the bottom part of the table. The LC-DFTB calculations for different values of the range-separation parameter  $\omega$  show, that the convergence issue do not show up for the typical values of  $\omega$  as can be seen in Fig. 8.4. Opening the gap in general improves the convergence.

These calculations show, that the LC-DFTB method is capable of correctly describe the proteins in their zwitterionic conformations in gas-phase. In this context it might be the method of choice as a basis for the fragment molecular orbital approach (FMO) [47, 135, 140], which allows to study biological systems with many thousands of atoms with quantum chemical methods.

## 8.5 SUMMARY

In this chapter we further demonstrated the signatures of the delocalization problem reduction in the LC-DFTB method. We performed the geometry optimization with the LC-DFTB method and the repulsive potentials from chapter 5 on *trans*-polyacetylene oligomers. The LC-DFTB shows the saturation of bond length alternation (BLA) for growing system size. Specifically, present parametrization predicts the BLA value of 11.1 pm for the molecule with 40 monomer units, which we consider here as the polymer limit. This result is slightly larger than the experimental value of 8.0 pm. Compared to the DFTB geometries which show the BLA of 3.6 pm and PBE/6-311G\* which predict the BLA of 2.0 pm the result of LC-DFTB is a clear improvement. In the case of

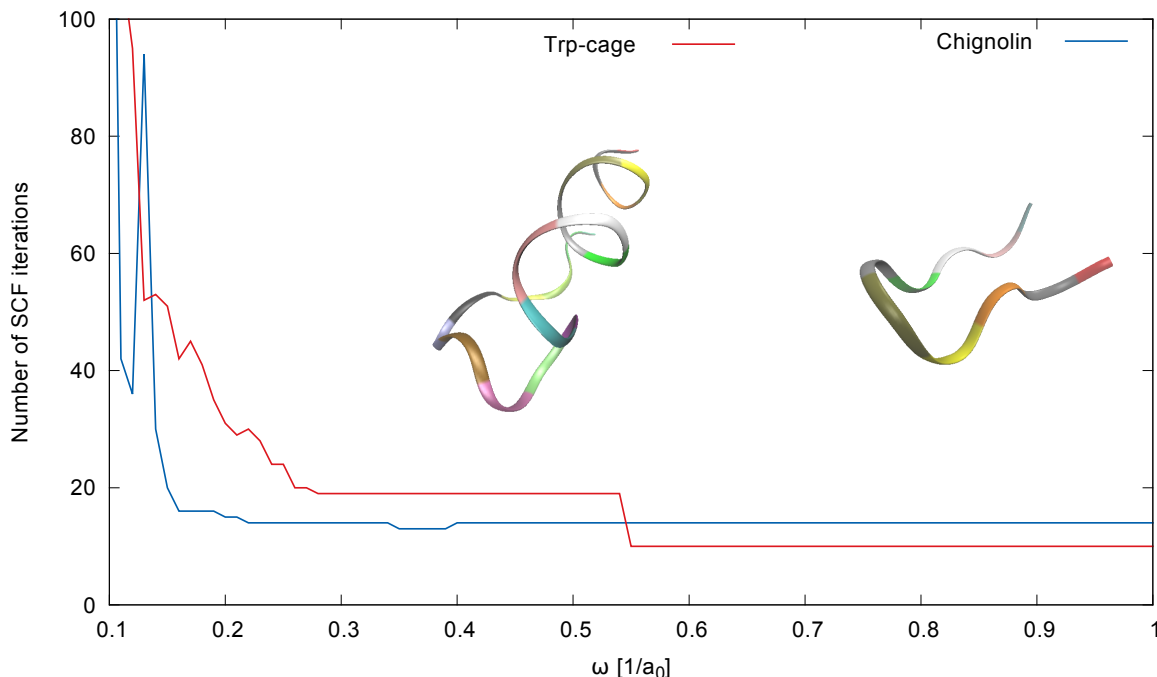


Figure 8.4: The SCF convergence plot of chignolin and Trp-cage proteins as a function of range-separation parameter  $\omega$  for the LC-DFTB method. The SCF does not converge for  $\omega < 0.1a_0^{-1}$ . Structures of the proteins are shown as insets. These plots have been generated from protein data base structures [74, 136] with the VMD software [79].

doped *trans*-polyacetylene, LC-DFTB clearly predicts the emergence of localized charge defects (bipolarons), while DFTB and local DFT is not able to account for this important effect. The study of the static longitudinal polarizabilities from the LC-DFTB method and the corresponding charge distribution reveal the presence of the field-counteracting term in the LC-DFTB. This finding is in particular important for the applications in the molecular electronics. We expect that the LC-DFTB method will be able to deal with the overestimation of conductance, which is usually observed in the transport calculations with local DFT. Finally, we showed that the LC-DFTB method can be successfully applied to large biological systems, such as proteins in zwitterionic conformation in gas-phase, which are in principle problematic for the standard DFTB method.





## CONCLUSIONS AND OPEN QUESTIONS

The main result of this work is the successful implementation of the DFTB method, based on the hybrid DFT-HF functionals. We provided the necessary parametrization tools and included the required algorithms in the development version of the DFTB+ code. Furthermore, we performed a parametrization of the method for the particular long-range corrected functional (part I of the thesis).

We applied the new method to a series of cases, where the delocalization problem of the local DFT and DFTB plays an important role (part II of the thesis). We found, that the new method, referred to as LC-DFTB, significantly improves the quality of the results as compared to the traditional DFTB and DFT with local xc-functionals. This improvement usually parallels that of the long-range corrected DFT with respect to the local DFT. Specifically, the new method outperforms the traditional DFTB in the description of orbital energies, HOMO-LUMO gaps, the response to electric fields, bond length alternation and description of polaronic states in conjugated polymers. Remarkable is the fact, that the LC-DFTB method is directly derived from the corresponding LC-DFT method by applying simple DFTB approximations. No empirical fits have been performed. The reduction of the self-interaction error is achieved due to the added non-local range-dependent term only.

The inclusion of this necessary term to the new Hamiltonian results in the increased computational requirements of LC-DFTB as compared to the standard DFTB. Although the present implementation already shows the expected quadratic scaling for the construction of the DFTB Hamiltonian with the non-local term, the employed algorithms can be further optimized. Especially the extension of the implementation to the parallel version is important to compete with first principles implementations, which usually exist as parallel versions. Of top priority is the combination of the neighbor list-based algorithm for the Hamiltonian construction with the

thresholding algorithm, which usually gives the speed-up factor of 2-3. The neighbor list is the basic concept of the DFTB+. Therefore, the implementation of the LC-DFTB method in this code should be based on the neighbor lists.

Important feature of the DFTB method, and in particular of its implementation in the DFTB+ code, is the ability to deal with both periodic and finite systems. The inclusion of non-local range-dependent HF exchange term in the Hamiltonian for the case of periodic systems would require only minor changes in the DFTB+ code, provided the neighbor list-based algorithm is used. This would allow the inclusion of the hybrid range-separated functional HSE06 [67–69], which has been successfully applied in the calculations of the lattice parameters, bulk moduli, adsorption energies, atomization energies and band gaps of metals, insulators and semiconductors [36, 78, 125, 142, 145, 157]. Usually the HSE06 functional outperforms the local DFT for periodic systems in the same way as the long-range corrected functionals do for the finite systems.

The LC-DFTB method in the present work has been formulated for the closed-shell case. This restricts the application area of the method. The extension to the spin-unrestricted case can in principle be carried out as in the standard DFTB [95, 96]. This would especially allow to use the procedure of non-empirical tuning of the exchange-correlation functional, where the condition  $|\epsilon_{\text{HOMO}}| = |E(N) - E(N - 1)| = \text{IP}$  is enforced and the value of the range-separation parameter for which this condition holds is used for the calculation. This approach further improves the performance of the long-range corrected DFT as compared to the exploitation of the universal value of the range-separation parameter [9, 93, 103, 161, 188].

Other branch of the further developments deals with parametrization of the method. At the present time the electronic Hamiltonian is available for the elements C,H,O,N and S. The repulsive potentials are available for the elements C and H. The results of this thesis show, that for the electronic Hamiltonian the parametrization can be done with the same basis set parameters (confinement radii) as in the standard DFTB. Nevertheless, the systematic optimization of the parameters would contribute to the understanding of the limits of the LC-DFTB method. On contrast, the repulsive potential should be definitely reoptimized for all elements. A careful in-depth investigations should find an optimal reference theory and figure out the optimal fit parameters.

## **Part III**

# **Appendix**





## $\gamma$ -INTEGRAL OVER YUKAWA INTERACTION

### A.1 REDUCTION TO ONE-DIMENSIONAL INTEGRAL

In this section we show how to reduce the  $\gamma$ -integral

$$\gamma_{AB}^{Y,\omega} = \frac{\tau_A^3 \tau_B^3}{(8\pi)^2} \int e^{-\tau_A |\mathbf{r}-\mathbf{R}_A|} e^{-\tau_B |\mathbf{r}'-\mathbf{R}_B|} \frac{\exp(-\omega |\mathbf{r}-\mathbf{r}'|)}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r} d\mathbf{r}' \quad (\text{A.1})$$

to an one-dimensional quadrature. Note, that a slightly different way has been shown in appendix of ref. [137]. We start with a general integral

$$\int f(\mathbf{r})g(\mathbf{r}')h(\mathbf{r}-\mathbf{r}') d\mathbf{r}d\mathbf{r}' = \int f(\mathbf{r})(g * h)(\mathbf{r})d\mathbf{r}, \quad (\text{A.2})$$

where  $(g * h)(\mathbf{r}) = \int g(\mathbf{r}')h(\mathbf{r}-\mathbf{r}')d\mathbf{r}'$  is the convolution of the functions  $g$  and  $h$ . Next we express the real space functions by the Fourier transformation integral, where  $\mathcal{F}[f](\mathbf{q})$  denotes the Fourier transformed function  $f(\mathbf{r})$

$$\int f(\mathbf{r})(g * h)(\mathbf{r})d\mathbf{r} = \int \left( \int \mathcal{F}[f](\mathbf{q})e^{-i\mathbf{q}\mathbf{r}} \frac{d\mathbf{q}}{(2\pi)^3} \right) \left( \mathcal{F}[(g * h)](\mathbf{k})e^{-i\mathbf{k}\mathbf{r}} \frac{d\mathbf{k}}{(2\pi)^3} \right) d\mathbf{r} \quad (\text{A.3})$$

$$= \int \left( \int \mathcal{F}[f](\mathbf{q})e^{-i\mathbf{q}\mathbf{r}} \frac{d\mathbf{q}}{(2\pi)^3} \right) \left( \mathcal{F}[g](\mathbf{k})\mathcal{F}[h](\mathbf{k})e^{-i\mathbf{k}\mathbf{r}} \frac{d\mathbf{k}}{(2\pi)^3} \right) d\mathbf{r} \quad (\text{A.4})$$

$$= \int \mathcal{F}[f](\mathbf{q})\mathcal{F}[g](\mathbf{k})\mathcal{F}[h](\mathbf{k}) \left( \int e^{-i(\mathbf{q}+\mathbf{k})\mathbf{r}} d\mathbf{r} \right) \frac{d\mathbf{k}d\mathbf{q}}{(2\pi)^6} \quad (\text{A.5})$$

$$= \int \mathcal{F}[f](\mathbf{q})\mathcal{F}[g](\mathbf{k})\mathcal{F}[h](\mathbf{k})(2\pi)^3 \delta(\mathbf{k}+\mathbf{q}) \frac{d\mathbf{k}d\mathbf{q}}{(2\pi)^6} \quad (\text{A.6})$$

$$= \frac{1}{(2\pi)^3} \int \mathcal{F}[f](\mathbf{k})\mathcal{F}[g](\mathbf{k})\mathcal{F}[h](\mathbf{k}) d\mathbf{k}, \quad (\text{A.7})$$

where in Eq. A.4 the convolution theorem was used. Thus the double integral in real space Eq. A.2 is now reduced to the a single integral in the Fourier space over a product of the Fourier

transforms of the functions  $f, g$  and  $h$ . Now we use the known Fourier transformed forms of the charge distribution  $F_A(\mathbf{r}) = \frac{\tau_A^3}{\pi} e^{-\tau_A |\mathbf{r} - \mathbf{R}_A|}$  and Yukawa interaction  $Y(r) = \frac{\exp(-\omega|r|)}{|r|}$

$$\mathcal{F}[F_A(\mathbf{r})] = \frac{\tau_A^4}{(k^2 + \tau_A^2)^2} e^{i\mathbf{k}\mathbf{R}_A} \quad (\text{A.8})$$

$$\mathcal{F}[Y(\mathbf{r})] = \frac{4\pi}{\omega^2 + k^2}. \quad (\text{A.9})$$

Inserting these expressions into the Eq. A.7 and integrating out the angular coordinate we obtain the expression for the one-dimensional quadrature

$$\gamma_{AB}^{Y,\omega} = \int F_A(\mathbf{r}) F_B(\mathbf{r}') Y(\mathbf{r} - \mathbf{r}') d\mathbf{r} d\mathbf{r}' = \frac{\tau_A^4 \tau_B^4 4\pi}{(2\pi)^3} \int \frac{e^{i\mathbf{k}\cdot(\mathbf{R}_A - \mathbf{R}_B)}}{(k^2 + \tau_A^2)^2 (k^2 + \tau_B^2)^2 (k^2 + \omega^2)} d\mathbf{k} \quad (\text{A.10})$$

$$= \frac{\tau_A^4 \tau_B^4 8\pi^2}{(2\pi)^3} \int_0^\infty \int_{-1}^1 \frac{k^2 e^{ikR_{AB} \cos\theta}}{(k^2 + \tau_A^2)^2 (k^2 + \tau_B^2)^2 (k^2 + \omega^2)} dk d\cos\theta \quad (\text{A.11})$$

$$= \frac{2\tau_A^4 \tau_B^4}{\pi R_{AB}} \int_0^\infty \frac{k \sin(kR_{AB})}{(k^2 + \tau_A^2)^2 (k^2 + \tau_B^2)^2 (k^2 + \omega^2)} dk. \quad (\text{A.12})$$

Here, we denote the interatomic distance as  $R_{AB} = |\mathbf{R}_A - \mathbf{R}_B|$ .

## A.2 ANALYTICAL EVALUATION OF THE OFF-SITE $\gamma$ -INTEGRAL

The one-dimensional integral Eq. A.12 can be rewritten

$$\gamma_{AB}^{Y,\omega} = \frac{2\tau_A^4 \tau_B^4}{\pi R_{AB}} \int_0^\infty \frac{q \sin(qR_{AB})}{(q^2 + \tau_A^2)^2 (q^2 + \tau_B^2)^2 (q^2 + \omega^2)} dq \quad (\text{A.13})$$

$$= \frac{2\tau_A^4 \tau_B^4}{2\pi i R_{AB}} \int_{-\infty}^\infty \frac{q e^{iqR_{AB}}}{(q^2 + \tau_A^2)^2 (q^2 + \tau_B^2)^2 (q^2 + \omega^2)} dq. \quad (\text{A.14})$$

The integrand has three poles at  $\pm i\tau_A, \pm i\tau_B, \pm i\omega$ . We close the contour in the upper half-plane and evaluate the residuals

$$\begin{aligned} \text{Res}(i\tau_A) &= \frac{d}{dq} \left[ \frac{q e^{iqR_{AB}} (q - i\tau_A)^2}{(q^2 + \tau_A^2)^2 (q^2 + \tau_B^2)^2 (q^2 + \omega^2)} \right] \Big|_{q=i\tau_A} \\ &= -\frac{e^{-\tau_A R_{AB}}}{2} \left[ \frac{\tau_B^2 - 3\tau_A^2 + 2\omega^2}{(\tau_B^2 - \tau_A^2)^3 (\omega^2 - \tau_A^2)^2} \right] + e^{-\tau_A R_{AB}} R_{AB} \left[ \frac{1}{4\tau_A (\tau_B^2 - \tau_A^2)^2 (\omega^2 - \tau_A^2)} \right] \end{aligned} \quad (\text{A.15})$$

$$\text{Res}(i\omega) = \frac{e^{-\omega R_{AB}}}{2(\tau_A^2 - \omega^2)^2 (\tau_B^2 - \omega^2)^2}. \quad (\text{A.16})$$

Note that the residual at the pole  $i\tau_B$  is obtained from the one at the pole  $i\tau_A$  by interchanging  $\tau_A \leftrightarrow \tau_B$ . Thus the integral reads

$$\begin{aligned}
 \gamma_{AB}^{Y,\omega} &= \frac{2\tau_A^4 \tau_B^4}{2\pi i R_{AB}} \int_{-\infty}^{\infty} \frac{q e^{iqR_{AB}}}{(q^2 + \tau_A^2)^2 (q^2 + \tau_B^2)^2 (q^2 + \omega^2)} dq \\
 &= \frac{2\tau_A^4 \tau_B^4}{2\pi i R_{AB}} 2\pi i [\text{Res}(i\tau_A) + \text{Res}(i\tau_B) + \text{Res}(i\omega)] \\
 &= \frac{\tau_A^4 \tau_B^4}{(\tau_A^2 - \omega^2)^2 (\tau_B^2 - \omega^2)^2} \frac{e^{-\omega R_{AB}}}{R_{AB}} \\
 &\quad - \left[ e^{-\tau_A R_{AB}} \left( \frac{\tau_A^2}{\tau_A^2 - \omega^2} \frac{\tau_A \tau_B^4}{2(\tau_B^2 - \tau_A^2)^2} - \frac{\tau_A^4}{(\omega^2 - \tau_A^2)^2} \frac{(\tau_B^6 - 3\tau_A^2 \tau_B^4 + 2\omega^2 \tau_B^4)}{(\tau_A^2 - \tau_B^2)^3 R_{AB}} \right) \right. \\
 &\quad \left. + e^{-\tau_B R_{AB}} \left( \frac{\tau_B^2}{\tau_B^2 - \omega^2} \frac{\tau_B \tau_A^4}{2(\tau_A^2 - \tau_B^2)^2} - \frac{\tau_B^4}{(\omega^2 - \tau_B^2)^2} \frac{(\tau_A^6 - 3\tau_B^2 \tau_A^4 + 2\omega^2 \tau_A^4)}{(\tau_B^2 - \tau_A^2)^3 R_{AB}} \right) \right]. \tag{A.17}
 \end{aligned}$$

In the limit  $\omega \rightarrow 0$  this formula is equivalent to the formula in ref. [42]. The long-range  $\gamma$ -integral can be calculated as  $\gamma_{AB}^{\text{lr},\omega}(R_{AB}) = \gamma_{AB}^{Y,0}(R_{AB}) - \gamma_{AB}^{Y,\omega}(R_{AB})$ .

### A.3 ANALYTICAL EVALUATION OF THE ON-SITE $\gamma$ -INTEGRAL

We show how to evaluate the on-site  $\gamma$ -integral. As has been shown the  $\gamma$ -integral over the Yukawa interaction can be reduced to the form

$$\gamma_{AB}^{Y,\omega}(R_{AB}) = \frac{2\tau_A^4 \tau_B^4}{\pi R_{AB}} \int_0^{\infty} \frac{q \sin(qR_{AB})}{(q^2 + \tau_A^2)^2 (q^2 + \tau_B^2)^2 (q^2 + \omega^2)} dq. \tag{A.18}$$

The on-site value is obtained from  $\gamma_{AA}^{Y,\omega} = \lim_{\substack{R_{AB} \rightarrow 0 \\ \tau_A = \tau_B}} \gamma_{AB}^{Y,\omega}(R_{AB})$ . We first perform the limit

$$\gamma_{AA}^{Y,\omega} = \lim_{\substack{R_{AB} \rightarrow 0 \\ \tau_A = \tau_B}} \gamma_{AB}^{Y,\omega}(R_{AB}) = \lim_{\substack{R_{AB} \rightarrow 0 \\ \tau_A = \tau_B}} \frac{2\tau_A^4 \tau_B^4}{\pi R_{AB}} \int_0^{\infty} \frac{q \sin(qR_{AB})}{(q^2 + \tau_A^2)^2 (q^2 + \tau_B^2)^2 (q^2 + \omega^2)} dq \tag{A.19}$$

$$= \lim_{\substack{R_{AB} \rightarrow 0 \\ \tau_A = \tau_B}} \frac{2\tau_A^4 \tau_B^4}{\pi R_{AB}} \int_0^{\infty} \frac{q}{(q^2 + \tau_A^2)^2 (q^2 + \tau_B^2)^2 (q^2 + \omega^2)} \left[ \sum_{k=0}^{\infty} (-1)^k \frac{(qR_{AB})^{2k+1}}{(2k+1)!} \right] dq \tag{A.20}$$

$$= \frac{2\tau_A^8}{\pi} \int_0^{\infty} \frac{q}{(q^2 + \tau_A^2)^4 (q^2 + \omega^2)} \lim_{R_{AB} \rightarrow 0} \frac{1}{R_{AB}} \left[ \sum_{k=0}^{\infty} (-1)^k \frac{(qR_{AB})^{2k+1}}{(2k+1)!} \right] dq \tag{A.21}$$

$$= \frac{2\tau_A^8}{\pi} \int_0^{\infty} \frac{q^2}{(q^2 + \tau_A^2)^4 (q^2 + \omega^2)} dq = \frac{\tau_A^8}{\pi} \int_{-\infty}^{\infty} \frac{q^2}{(q^2 + \tau_A^2)^4 (q^2 + \omega^2)} dq. \tag{A.22}$$

This integral can be evaluated, using the residue theorem. The integral has poles of the 4. degree at  $q = \pm i\tau_A$ , and simple poles at  $q = \pm i\omega$ . We choose the integration path to go from  $-\infty$  to  $\infty$

along the real axis and close the contour in the upper half-plane. The residues at both poles are

$$2\pi i \operatorname{Res}(i\tau_A) = \frac{(5\tau_A^6 + 15\tau_A^4\omega^2 - 5\tau_A^2\omega^4 + \omega^6)\pi}{16\tau_A^5(\tau_A^2 - \omega^2)^4} \quad (\text{A.23})$$

$$2\pi i \operatorname{Res}(i\omega) = -\frac{\omega\pi}{(\tau_A^2 - \omega^2)^4}. \quad (\text{A.24})$$

From this we obtain the value of the on-site integral

$$\gamma_{AA}^{Y,\omega} = \frac{\tau_A^8}{\pi} \int_{-\infty}^{\infty} \frac{q^2}{(q^2 + \tau_A^2)^4(q^2 + \omega^2)} dq = \frac{\tau_A^8}{\pi} \sum_{z \in \{i\tau_A, i\omega\}} 2\pi i \operatorname{Res}(z) \quad (\text{A.25})$$

$$= \frac{\tau_A^8}{(\tau_A^2 - \omega^2)^4} \left[ \frac{5\tau_A^6 + 15\tau_A^4\omega^2 - 5\tau_A^2\omega^4 + \omega^6}{16\tau_A^5} - \omega \right]. \quad (\text{A.26})$$

The limit  $\omega \rightarrow 0$ , which corresponds to the Coulomb interaction gives already known result, which is used in the standard DFTB

$$\lim_{\omega \rightarrow 0} \gamma_{AA}^{Y,\omega} = \frac{5}{16} \tau_A \quad (\text{A.27})$$

and the opposite limit  $\omega \rightarrow \infty$  vanishes as expected.



## AVERAGE POTENTIAL

The practical Kohn-Sham calculation results in a set of converged single particle orbitals  $\{\psi_i(\mathbf{r})\}$ ,  $\langle \psi_i | \psi_j \rangle = \delta_{ij}$ , which constitute the ground state density  $\rho(\mathbf{r}) = \sum_{i=1}^N |\psi_i(\mathbf{r})|^2$ . This density corresponds to the local Kohn-Sham potential  $v_{\text{KS}}(\mathbf{r})$ . If we know the exact potential, we can immediately obtain the orbitals. Now let us consider the reverse problem. Assume we have a set of converged orbitals, how to determine the local potential to which these orbitals correspond? In particular this can be interesting, if the orbitals have been optimized in a non-local potential. Baer et al. [9] suggested to define the deviance vector  $|D_j\rangle$

$$|D_j\rangle = \left( \epsilon_j - \left[ -\frac{1}{2}\nabla^2 + v_{\text{avg}}(\mathbf{r}) \right] \right) |\psi_j\rangle \quad (\text{B.1})$$

with local average potential  $v_{\text{avg}}(\mathbf{r})$  and effective orbital energies  $\epsilon_i$ . Minimizing the functional

$$\begin{aligned} L[v_{\text{avg}}, \{\epsilon_j\}] &= \sum_{j=1}^N \langle D_j | D_j \rangle \\ &= \sum_{j=1}^N \int \psi_j(\mathbf{r}') \left( \epsilon_j - \left[ -\frac{1}{2}\nabla^2 + v_{\text{avg}}(\mathbf{r}') \right] \right) \left( \epsilon_j - \left[ -\frac{1}{2}\nabla^2 + v_{\text{avg}}(\mathbf{r}') \right] \right) \psi_j(\mathbf{r}') d\mathbf{r}' \\ &= \sum_{j=1}^N \int \psi_j(\mathbf{r}') \left( \epsilon_j^2 - 2\epsilon_j \left[ -\frac{1}{2}\nabla^2 + v_{\text{avg}}(\mathbf{r}') \right] + \left[ -\frac{1}{2}\nabla^2 + v_{\text{avg}}(\mathbf{r}') \right]^2 \right) \psi_j(\mathbf{r}') d\mathbf{r}' \end{aligned} \quad (\text{B.2})$$

with respect to the average potential and the orbital energies we obtain the equations

$$0 \stackrel{!}{=} \frac{\partial L}{\partial \epsilon_k} = \sum_{j=1}^N \int \psi_j(\mathbf{r}') \left( 2\epsilon_j \delta_{jk} - 2\delta_{jk} \left[ -\frac{1}{2}\nabla^2 + v_{\text{avg}}(\mathbf{r}') \right] \right) \psi_j(\mathbf{r}') d\mathbf{r}' \quad (\text{B.3})$$

$$\Rightarrow \epsilon_j = \int \psi_j(\mathbf{r}') \left[ -\frac{1}{2}\nabla^2 + v_{\text{avg}}(\mathbf{r}') \right] \psi_j(\mathbf{r}') d\mathbf{r}' \quad (\text{B.4})$$

$$\begin{aligned}
 0 &\stackrel{!}{=} \frac{\delta L}{\delta v_{\text{avg}}(\mathbf{r})} = \sum_{j=1}^N \int \psi_j(\mathbf{r}') \left( -2\epsilon_j \delta(\mathbf{r} - \mathbf{r}') + 2 \left[ -\frac{1}{2} \nabla^2 + v_{\text{avg}}(\mathbf{r}') \right] \delta(\mathbf{r} - \mathbf{r}') \right) \psi_j(\mathbf{r}') d\mathbf{r}' \\
 &= 2 \sum_{j=1}^N \psi_j(\mathbf{r}) \left[ -\epsilon_j + \left[ -\frac{1}{2} \nabla^2 + v_{\text{avg}}(\mathbf{r}) \right] \right] \psi_j(\mathbf{r}) \tag{B.5}
 \end{aligned}$$

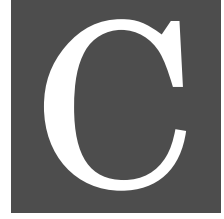
$$\Rightarrow \sum_{i=1}^N |\psi_i(\mathbf{r})|^2 v_{\text{avg}}(\mathbf{r}) = \sum_{j=1}^N \psi_j(\mathbf{r}) \left( \epsilon_j + \frac{1}{2} \nabla^2 \right) \psi_j(\mathbf{r}). \tag{B.6}$$

Simultaneous solution of these equations provides the average potential

$$v_{\text{avg}}(\mathbf{r}) = \frac{1}{\rho(\mathbf{r})} \sum_{j=1}^N \psi_j(\mathbf{r}) \left( \epsilon_j + \frac{1}{2} \nabla^2 \right) \psi_j(\mathbf{r}), \tag{B.7}$$

$$\epsilon_j = \langle \psi_j | v_{\text{avg}} - \frac{1}{2} \nabla^2 | \psi_j \rangle. \tag{B.8}$$

The potential is determined up to a constant. To eliminate this ambiguity we require  $\epsilon_N = \epsilon_{\text{HOMO}}$ , where  $\epsilon_{\text{HOMO}}$  is the HOMO eigenvalue of the molecular calculation, from which we obtained the orbitals.



## ALGORITHM FOR THE CORRECTED DECAY CONSTANTS

The equation 3.52 can be solved numerically in the following way. First we rescale the decay constant  $\tau$  by the range-separation parameter and introduce a new variable  $x = \tau/\omega$ . We rewrite then the Eq. 3.52 in terms of these quantities

$$xg_l(1 - P(x)) = x - \tilde{u}, \quad (\text{C.1})$$

where  $\tilde{u} = \frac{16U}{5\omega}$  and  $g_l = \frac{1}{2(2l+1)}$  and we define the function

$$P(\tau, \omega) = \frac{\tau^8 + 3\tau^6\omega^2 - \tau^4\omega^4 + 0.2\omega^6\tau^2 - 3.2\tau^7\omega}{(\tau^2 - \omega^2)^4} \quad (\text{C.2})$$

and further note that  $P(\tau, \omega) = P(\tau/\omega, 1) = P(x)$ . Simplifying this function, we obtain

$$P(x) = \frac{x^8 + 3x^6 - x^4 + 0.2x^2 - 3.2x^7}{(x^2 - 1)^4} = \frac{x^2(x^2 + 0.8x + 0.2)(x - 1)^4}{(x - 1)^4(x + 1)^4} = \frac{x^2(x^2 + 0.8x + 0.2)}{(x + 1)^4}. \quad (\text{C.3})$$

The variable  $x$  is by definition non-negative and can in principle take values in the interval  $x \in [0, \infty)$ . We expand the denominator  $(1+x)^4 = x^4 + 4x^3 + 6x^2 + 4x + 1 > x^4 + 0.8x^3 + 0.2x^2$ ,  $\forall x \in [0, \infty)$  thus  $P(x) < 1$ ,  $\forall x \in [0, \infty)$ . We investigate the limits of the expression

$$f(x) = x \left( 1 - \frac{x^4 + 0.8x^3 + 0.2x^2}{(1+x)^4} \right). \quad (\text{C.4})$$

The limit  $\lim_{x \rightarrow 0} f(x) \rightarrow 0$  obviously. To obtain the limit for  $x \rightarrow \infty$  we proceed as follows

$$\lim_{x \rightarrow \infty} f(x) = \lim_{x \rightarrow \infty} x \left( 1 - \frac{x^4 + 0.8x^3 + 0.2x^2}{(1+x)^4} \right) = \lim_{x \rightarrow \infty} x \left( 1 - \frac{1 + 0.8\frac{1}{x} + 0.2\frac{1}{x^2}}{(1 + \frac{1}{x})^4} \right) \quad (\text{C.5})$$

$$\approx \lim_{x \rightarrow \infty} x \left( 1 - \left( 1 + 0.8\frac{1}{x} + 0.2\frac{1}{x^2} \right) \left( 1 - 4\frac{1}{x} + 10\frac{1}{x^2} + \mathcal{O}\left(\frac{1}{x^3}\right) \right) \right) \quad (\text{C.6})$$

$$= \lim_{x \rightarrow \infty} x \left( 1 - \left( 1 + 0.8\frac{1}{x} - 4\frac{1}{x} + \mathcal{O}\left(\frac{1}{x^2}\right) \right) \right) = \lim_{x \rightarrow \infty} \left( 3.2 + \mathcal{O}\left(\frac{1}{x}\right) \right) = 3.2 \quad (\text{C.7})$$

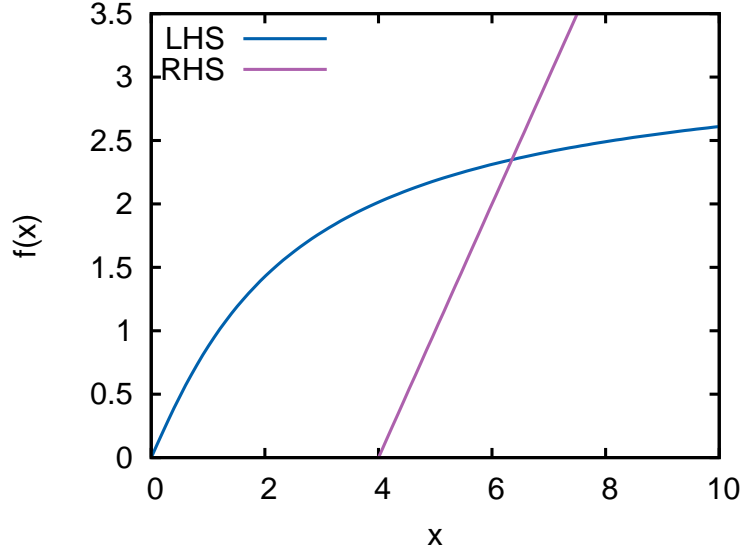


Figure C.1: The left-hand side (LHS) and right-hand side (RHS) of the equation C.1 for  $\tilde{u} = 4$  and  $g_l = 1$ .

Next we perform the derivative


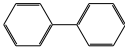
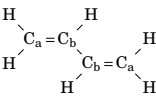
$$\frac{df}{dx} = 1 - \frac{x^5 + 1.8x^4 + x^3 + 0.2x^2}{(1+x)^5} \quad (\text{C.8})$$

Using the same analysis as in the case of the function  $P(x)$ , expanding the denominator  $(1+x)^5 = x^5 + 5x^4 + 10x^3 + 10x^2 + 5x + 1 > x^5 + 1.8x^4 + x^3 + 0.2x^2$  we conclude that the second term in Eq. C.8 takes values in the interval  $[0, 1)$ , thus  $\lim_{x \rightarrow \infty} \frac{df}{dx} = 0$ ,  $\lim_{x \rightarrow 0} \frac{df}{dx} = 1$  and it follows, that  $f(x)$  is monotonous on the positive real numbers. Thus the Eq. C.1 has at most one solution on the positive real numbers, which is the intersection point of the straight line (right-hand side) and the function  $f(x)g_l$  (left-hand side). As an example, we plot the function  $f(x)$  together with the right hand side of the Eq. C.1  $h(x) = x - \tilde{u}$  in Fig. C.1 for the values  $\tilde{u} = 4$  and  $g_l = 1$ . The function  $h(x)$  intersects the  $x$ -axis in the point  $\tilde{u}$  and the limit  $3.2g_l$  in the point  $3.2g_l + \tilde{u}$ . From the analysis above we conclude that the solution, the intersection point of functions  $g_l f(x)$  and  $h(x)$ , is in the interval  $x \in [\tilde{u}, \tilde{u} + 3.2g_l]$ .

To formulate a solution algorithm we use the nested interval principle. Given the values  $\omega, U, l$  the solution can be found in the interval  $x \in [\tilde{u}, \tilde{u} + 3.2g_l]$ . We evaluate thus the function  $g_l f(x)$  at the limits of this interval and obtain the new limits for the linear function. The solution interval in general reads  $I_n = [x_{\min}^n, x_{\max}^n] = [f(x_{\min}^{n-1}), f(x_{\max}^{n-1})]$ , where  $I_0 = [x_{\min}^0, x_{\max}^0] = [\tilde{u}, \tilde{u} + 3.2g_l]$ . The properties of the functions (monotonous, continuous, limited) lead to a converged solution (interval width reduces to a given numerical precision).

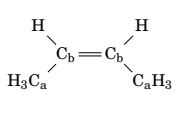
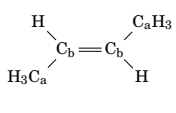
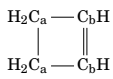
## GEOMETRY DATA FOR HYDROCARBONS

 Table D.1: Bond lengths in [ $\text{\AA}$ ] and angles in [deg] for selected hydrocarbons from LC-DFTB, standard DFTB, B3LYP/6-311G\*, BNL/6-311G\* and experiment.

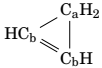
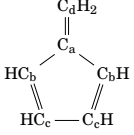
Molecule	Quantity	Exp. [115] ([65])	LC-DFTB	DFTB	B3LYP	BNL
<i>Allene</i>						
<chem>H2C=C=CH2</chem>	C-C	1.308	1.315	1.313	1.303	1.323
	C-H	1.087	1.106	1.096	1.087	1.122
	$\angle\text{HCH}$	118.2	116.1	117.5	116.9	117.7
	$\angle\text{HCC}$	120.9	122.0	121.2	121.5	121.2
<i>Benzene</i>						
	C-C	1.397 (1.399)	1.411	1.397	1.394	1.413
	C-H	1.084 (1.101)	1.111	1.099	1.086	1.122
<i>Biphenyl</i>						
	C-C (intra ring)	(1.396)	1.411	1.397	1.393	1.412
	C-C (inter ring)	1.489	1.526	1.476	1.485	1.510
	dihedral (rings)	41.6	30.3	29.8	40.9	43.0
<i>1,3-Butadiene</i>						
	C <sub>b</sub> -C <sub>b</sub>	1.476 (1.467)	1.495	1.455	1.456	1.487
	C <sub>a</sub> -C <sub>b</sub>	1.337 (1.349)	1.342	1.342	1.337	1.352
	C-H (average)	(1.108)	1.106	1.096	1.087	1.124
	$\angle\text{CCC}$	122.9 (124.4)	123.9	122.9	124.4	123.9

Continued on next page

## APPENDIX D. GEOMETRY DATA FOR HYDROCARBONS

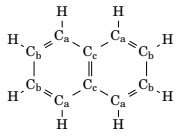
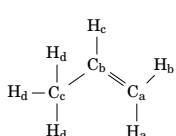
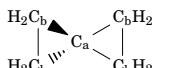
Molecule	Quantity	Exp. [115] ([65])	LC-DFTB	DFTB	B3LYP	BNL
	$\angle C_b C_a H$	(120.9)	121.4	121.4	121.8	121.9
<i>1,3-Butadiyne</i>						
	$C_a-C_b$	1.205 (1.218)	1.207	1.214	1.206	1.222
$HC_a \equiv C_b - C_b \equiv C_a H$	$C_b-C_b$	1.378 (1.384)	1.424	1.393	1.364	1.403
	C-H	1.058 (1.09)	1.087	1.075	1.064	1.102
<i>Butatriene</i>						
	$C_a-C_b$	1.318	1.328	1.327	1.314	1.330
	$C_b-C_b$	1.283	1.289	1.281	1.265	1.292
$H_2 C_a = C_b = C_b = C_a H_2$	C-H	1.083	1.108	1.098	1.087	1.122
	$\angle HCH$	117.0	115.7	117.3	116.7	117.4
	$\angle CCH$	121.5	122.1	121.3	121.6	121.3
<i>2-Butene (cis)</i>						
	$C_a-C_b$	1.506	1.521	1.486	1.502	1.523
	$C_b-C_b$	1.346	1.344	1.341	1.334	1.352
	$\angle C_a C_b C_b$	125.4	127.4	126.6	128.0	127.6
<i>2-Butene (trans)</i>						
	$C_a-C_b$	1.508	1.523	1.486	1.501	1.522
	$C_b-C_b$	1.347	1.345	1.340	1.331	1.349
	$\angle C_a C_b C_b$	123.8	124.4	123.5	125.4	125.1
<i>2-Butyne</i>						
	$C_b-C_b$	1.214	1.204	1.209	1.203	1.221
$H_3 C_a - C_b \equiv C_b - C_a H_3$	$C_a-C_b$	1.468	1.484	1.455	1.459	1.484
	C-H	1.116	1.105	1.100	1.094	1.129
	$\angle C_b C_a H$	110.7	111.3	110.8	111.4	111.3
<i>Cyclobutane</i>						
$(CH_2)_4$	C-H	1.092 (1.113)	1.110	1.103	1.092	1.128
	C-C	1.555	1.568	1.540	1.557	1.575
<i>Cyclobutene</i>						
	$C_b-C_b$	1.342	1.351	1.358	1.337	1.356
	$C_a-C_a$	1.566	1.588	1.569	1.572	1.588
	$C_a-C_b$	1.517	1.539	1.525	1.519	1.539
	$C_a-H$	1.094	1.112	1.104	1.095	1.130
	$C_b-H$	1.083	1.109	1.097	1.086	1.122
	$\angle C_a C_b C_b$	94.2	94.4	94.0	94.4	94.3
	$\angle C_b C_b H$	133.5	133.4	133.4	133.5	133.4
	$\angle C_a C_a H$	114.5	114.8	114.8	114.8	115.0

Continued on next page

Molecule	Quantity	Exp. [115] ([65])	LC-DFTB	DFTB	B3LYP	BNL
	$\angle C_a C_a C_b$	85.8	85.6	86.0	85.6	85.7
	$\angle HC_a H$	109.2	108.8	109.2	108.4	108.6
<i>Cyclopentane</i>						
	C-H	1.114	1.112	1.105	1.094	1.130
(CH <sub>2</sub> ) <sub>5</sub>	C-C	1.546	1.559	1.518	1.537	1.553
	$\angle CCH$	111.7	110.2	110.1	110.0	109.9
<i>Cyclopropane</i>						
	C-C	1.501 (1.512)	1.519	1.490	1.508	1.526
	C-H	1.083	1.102	1.096	1.084	1.120
(CH <sub>2</sub> ) <sub>3</sub>	$\angle HCH$	114.5 (114.0)	114.3	114.2	114.0	114.0
	$\angle HCC$	117.9	118.0	118.1	118.2	118.3
<i>Cyclopropene</i>						
	C <sub>a</sub> -H	1.088 (1.112)	1.111	1.107	1.092	1.128
	C <sub>b</sub> -C <sub>b</sub>	1.296 (1.304)	1.320	1.319	1.290	1.310
	C <sub>a</sub> -C <sub>b</sub>	1.509 (1.519)	1.531	1.495	1.509	1.529
	C <sub>b</sub> -H	1.072 (1.077)	1.100	1.090	1.078	1.114
	$\angle HC_a H$	114.57	113.4	113.2	113.3	113.5
	$\angle C_b C_b H$	149.85	148.8	148.3	149.7	149.8
<i>Fulvene</i>						
	C <sub>a</sub> -C <sub>d</sub>	1.349	1.352	1.350	1.341	1.354
	C <sub>a</sub> -C <sub>b</sub>	1.470	1.513	1.469	1.474	1.499
	C <sub>b</sub> -C <sub>c</sub>	1.355	1.362	1.360	1.351	1.365
	C <sub>c</sub> -C <sub>c</sub>	1.476	1.500	1.463	1.475	1.500
	C <sub>b</sub> -H	1.078	1.107	1.095	1.082	1.119
	C <sub>c</sub> -H	1.080	1.108	1.096	1.083	1.120
	C <sub>d</sub> -H	1.130	1.103	1.094	1.085	1.122
	$\angle C_a C_b C_c$	107.7	108.4	107.9	107.9	108.0
	$\angle C_a C_b H$	124.7	124.7	124.9	124.5	124.4
	$\angle HC_d H$	117.0	116.1	117.4	116.7	117.1
	$\angle C_b C_a C_b$	106.6	104.7	106.1	106.1	105.9
	$\angle C_b C_c C_c$	109.0	109.2	109.0	109.0	109.1
	$\angle C_b C_c H$	126.4	126.4	126.4	126.5	126.4
<i>Isobutane</i>						
	C <sub>a</sub> -C <sub>b</sub>	1.525 (1.535)	1.559	1.518	1.534	1.549
	C <sub>a</sub> -H	1.108 (1.122)	1.126	1.116	1.099	1.134
(C <sub>b</sub> H <sub>3</sub> ) <sub>2</sub> C <sub>a</sub> H	C <sub>b</sub> -H	1.100 (1.113)	1.102	1.098	1.094	1.130


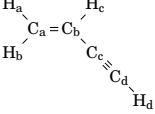

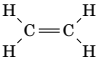

Continued on next page

## APPENDIX D. GEOMETRY DATA FOR HYDROCARBONS

Molecule	Quantity	Exp. [115] ([65])	LC-DFTB	DFTB	B3LYP	BNL
	$\angle C_a C_b H$	n/a (111.4)	110.9	110.9	111.5	111.7
	$\angle C_b C_a C_b$	111.2 (110.8)	111.2	110.9	111.0	110.7
	$\angle C_b C_a H_a$	109.4	107.7	108.1	107.8	108.1
	$\angle HC_b H$ (average)	108.2	107.7	107.9	107.6	107.6
<i>Methylene</i>						
CH <sub>2</sub>	C-H	1.085 (1.078)	1.120	1.115	1.115	1.154
	$\angle HCH$	135.5 (130)	103.8	100.1	101.0	100.7
<i>Naphtalene</i>						
	C <sub>a</sub> -C <sub>b</sub>	1.370	1.388	1.380	1.374	1.389
	C <sub>b</sub> -C <sub>b</sub>	1.410	1.437	1.415	1.415	1.441
	C <sub>a</sub> -C <sub>c</sub>	1.420	1.450	1.422	1.420	1.442
	C <sub>c</sub> -C <sub>c</sub>	1.420	1.449	1.428	1.432	1.438
	$\angle C_a C_c C_c$	119.4	118.8	119.1	118.8	119.1
<i>Neopentane</i>						
C(CH <sub>3</sub> ) <sub>4</sub>	C-C	1.537	1.575	1.526	1.539	1.553
	C-H	1.114	1.102	1.098	1.095	1.131
	$\angle CCH$	112.2	111.1	111.0	111.3	111.2
	$\angle HCH$	106.6	107.8	108.0	107.6	107.7
<i>Propene</i>						
	C <sub>a</sub> -H <sub>a</sub>	n/a (1.104)	1.103	1.094	1.087	1.124
	C <sub>a</sub> -C <sub>b</sub>	1.353 (1.341)	1.336	1.334	1.329	1.347
	C <sub>c</sub> -H <sub>d</sub>	n/a (1.117)	1.104	1.100	1.096	1.130
	C <sub>b</sub> -C <sub>c</sub>	1.488 (1.506)	1.520	1.486	1.501	1.521
	$\angle C_b C_c H_d$	n/a (110.7)	110.7	110.7	111.2	111.2
	$\angle C_b C_a H_a$	n/a (121.3)	121.3	121.3	121.8	121.8
	$\angle C_a C_b C_c$	124.8 (124.3)	125.2	123.9	125.4	124.9
<i>Propyne</i>						
H <sub>3</sub> C-C≡C-H	C <sub>c</sub> -C <sub>b</sub>	1.460 (1.459)	1.480	1.453	1.457	1.481
	C <sub>a</sub> -H	1.060 (1.056)	1.086	1.074	1.064	1.102
	C <sub>c</sub> -H	1.096 (1.105)	1.104	1.100	1.094	1.128
	C <sub>b</sub> -C <sub>a</sub>	1.207 (1.206)	1.201	1.206	1.201	1.220
	$\angle HC_c C_b$	110.6 (110.2)	111.1	110.6	111.2	111.2
	$\angle HC_c H$	108.3	107.8	108.3	107.7	107.8
	<i>Spiropentane</i>					
	C <sub>b</sub> -C <sub>b</sub>	1.520	1.543	1.508	1.530	1.547
	C <sub>a</sub> -C <sub>b</sub>	1.470	1.512	1.479	1.484	1.502

Continued on next page



Molecule	Quantity	Exp. [115] ([65])	LC-DFTB	DFTB	B3LYP	BNL
	C-H	1.090	1.104	1.097	1.086	1.121
	$\angle$ HCH	118	114.2	114.2	114.3	114.6
	$\angle$ C <sub>b</sub> C <sub>a</sub> C <sub>b</sub> (intra ring)	62	61.3	61.3	62.1	62.0
	$\angle$ C <sub>b</sub> C <sub>a</sub> C <sub>b</sub> (inter ring)	137.3	137.7	137.8	137.2	137.3
<i>Toluene</i>						
	C-C (ring)	1.395 (1.399)	1.410	1.396	1.392	1.411
	C-CH <sub>3</sub>	1.513 (1.524)	1.535	1.494	1.510	1.530
	C-H (average)	1.082 (1.11)	1.108	1.098	1.088	1.124
<i>Vinylacetylene</i>						
	C <sub>b</sub> -C <sub>c</sub>	1.434	1.465	1.433	1.423	1.455
	C <sub>a</sub> -C <sub>b</sub>	1.344	1.342	1.340	1.337	1.352
	C <sub>c</sub> -C <sub>d</sub>	1.215	1.205	1.211	1.204	1.221
	C <sub>a</sub> -H <sub>a</sub>	1.106	1.102	1.093	1.084	1.121
	C <sub>d</sub> -H <sub>d</sub>	1.090	1.086	1.075	1.064	1.102
	$\angle$ C <sub>b</sub> C <sub>c</sub> C <sub>d</sub>	177.9	178.8	177.8	178.1	179.1
	$\angle$ H <sub>b</sub> C <sub>a</sub> C <sub>b</sub>	121.6	122.7	121.7	121.8	121.3
	$\angle$ C <sub>c</sub> C <sub>d</sub> H <sub>d</sub>	177.7	179.6	178.5	179.2	179.5
	$\angle$ H <sub>a</sub> C <sub>a</sub> C <sub>b</sub>	118.7	121.0	120.9	121.0	121.1
	$\angle$ H <sub>a</sub> C <sub>b</sub> C <sub>a</sub>	121.7	119.8	120.8	119.5	120.1
	$\angle$ C <sub>a</sub> C <sub>b</sub> C <sub>c</sub>	123.1	124.4	123.1	124.6	123.6
<i>Methane</i>						
	C-H	1.087	1.091	1.089	1.091	1.126
	$\angle$ HCH	109.5	109.5	109.5	109.5	109.5
<i>Acetylene</i>						
HC≡CH	C-C	1.203	1.198	1.203	1.198	1.217
	C-H	1.063	1.087	1.075	1.065	1.103
<i>Ethylene</i>						
	C-C	1.339	1.327	1.327	1.327	1.346
	C-H	1.086	1.103	1.094	1.086	1.123
	$\angle$ HCC	121.2	122.1	121.6	121.9	121.7
<i>Ethane</i>						
	C-C	1.536	1.531	1.501	1.530	1.549
	C-H	1.091	1.102	1.098	1.094	1.129
	$\angle$ HCH	108.0	107.7	107.8	107.4	107.2
	$\angle$ HCC	110.9	111.2	111.1	111.4	111.6

## APPENDIX D. GEOMETRY DATA FOR HYDROCARBONS

Mode	LC-DFTB	DFTB	B3LYP	BNL	Exp.[115]	$\Delta$	$\Delta_{\text{rel}}[\%]$
<i>Methane</i>							
$\nu_1$ , $A_1$ , sym stretch	3040.0	2953.3	3032.2	2853.2	2917.0	123.0	4.2
$\nu_2$ , $E$ , deformation (2-fold)	1590.7	1514.5	1581.1	1468.5	1534.0	56.7	3.7
$\nu_3$ , $T_2$ , stretch (3-fold)	3329.0	3155.1	3136.8	2978.0	3019.0	310.0	10.3
$\nu_4$ , $T_2$ , deformation (3-fold)	1388.4	1336.7	1355.4	1246.9	1306.0	82.4	6.3
<i>Acetylene</i>							
$\nu_1$ , $\Sigma_g^+$ , CH stretch	3524.3	3427.7	3514.8	3297.4	3374.0	150.3	4.5
$\nu_2$ , $\Sigma_g^+$ , CC stretch	2190.0	2114.2	2070.3	2013.4	1974.0	216.0	10.9
$\nu_3$ , $\Sigma_u^+$ , CH stretch (2-fold)	3384.2	3309.0	3412.7	3187.8	3289.0	95.2	2.9
$\nu_4$ , $\pi_g$ , CH bend	708.9	496.7	565.6	620.3	612.0	96.9	15.8
$\nu_5$ , $\pi_u$ , CH bend	783.8	708.0	765.7	730.3	730.0	53.8	7.4
<i>Ethylene</i>							
$\nu_1$ , $A_g$ , $\text{CH}_2$ stretch	3119.2	3025.1	3138.1	2955.5	3026.0	93.2	3.1
$\nu_2$ , $A_g$ , CC stretch	1890.3	1823.8	1697.2	1648.5	1623.0	267.3	16.5
$\nu_3$ , $A_g$ , $\text{CH}_2$ scissor	1358.6	1316.7	1383.7	1303.6	1342.0	16.6	1.2
$\nu_4$ , $A_u$ , $\text{CH}_2$ twist	1138.1	1043.9	1066.9	1014.5	1023.0	115.1	11.3
$\nu_5$ , $B_{1g}$ , $\text{CH}_2$ A-stretch	3249.1	3126.4	3192.7	3010.6	3103.0	146.1	4.7
$\nu_6$ , $B_{1g}$ , $\text{CH}_2$ rocking	1325.1	1235.3	1247.1	1171.5	1236.0	89.1	7.2
$\nu_7$ , $B_{1u}$ , $\text{CH}_2$ wagging	971.5	844.3	967.5	937.9	949.0	22.5	2.4
$\nu_8$ , $B_{2g}$ , $\text{CH}_2$ wagging	984.4	908.9	956.2	923.9	943.0	41.4	4.4
$\nu_9$ , $B_{2u}$ , $\text{CH}_2$ A-stretch	3270.1	3148.7	3221.1	3032.2	3106.0	164.1	5.3
$\nu_{10}$ , $B_{2u}$ , $\text{CH}_2$ rocking	887.0	850.1	839.3	811.6	826.0	61.0	7.4
$\nu_{11}$ , $B_{3u}$ , $\text{CH}_2$ s-stretch	3114.2	3015.7	3123.1	2937.4	2989.0	125.2	4.2
$\nu_{12}$ , $B_{2u}$ , $\text{CH}_2$ scissor	1453.9	1386.3	1478.4	1371.6	1444.0	9.9	0.7
<i>Ethane</i>							
$\nu_1$ , $A_{1g}$ , $\text{CH}_3$ s-stretch	3032.3	2930.7	3031.6	2861.3	2954.0	78.3	2.7
$\nu_2$ , $A_{1g}$ , $\text{CH}_3$ s-deform	1475.3	1470.0	1437.5	1351.8	1388.0	87.3	6.3
$\nu_3$ , $A_{1g}$ , CC stretch	1083.3	1130.3	1001.8	988.7	995.0	88.3	8.9
$\nu_4$ , $A_{1u}$ , torsion	262.3	277.5	307.3	337.2	289.0	-26.7	-9.2
$\nu_5$ , $A_{2u}$ , $\text{CH}_3$ s-stretch	3012.3	2917.2	3031.6	2861.3	2896.0	116.3	4.0
$\nu_6$ , $A_{2u}$ , $\text{CH}_3$ s-deform	1453.3	1402.7	1420.3	1324.6	1379.0	74.3	5.4
$\nu_7$ , $E_g$ , $\text{CH}_3$ d-stretch	3216.2	3050.2	3075.6	2921.0	2969.0	247.2	8.3
$\nu_8$ , $E_g$ , $\text{CH}_3$ d-deform	1531.2	1461.4	1521.4	1413.5	1468.0	63.2	4.3
$\nu_9$ , $E_g$ , $\text{CH}_3$ rocking	1328.8	1253.3	1226.8	1148.9	1190.0	138.8	11.7
$\nu_{10}$ , $E_u$ , $\text{CH}_3$ d-stretch	3232.6	3068.1	3101.3	2938.4	2985.0	247.6	8.3
$\nu_{11}$ , $E_u$ , $\text{CH}_3$ d-deform	1548.1	1478.0	1524.1	1416.0	1469.0	79.1	5.4
$\nu_{12}$ , $E_u$ , $\text{CH}_3$ rocking	900.8	870.4	830.7	799.8	822.0	78.8	9.6
Mean signed error [ $\text{cm}^{-1}$ ]	106.1	25.9	61.4	-38.0			
Mean absolute error [ $\text{cm}^{-1}$ ]	107.7	50.7	64.1	45.2			

Table D.2: Vibrational frequencies in [ $\text{cm}^{-1}$ ] of reference molecules for LC-DFTB in present parametrization, standard DFTB and first principles B3LYP/6-311G\* and BNL/6-311G\*.  $\Delta$  and  $\Delta_{\text{rel}}$  are absolute and relative errors of LC-DFTB compared to experiment.



## NEIGHBOR LIST-BASED EVALUATION OF FORCES

As in the case of Hamiltonian evaluation (section 3.9.2) we write out the sum and collect the summands to the respective force contributions. We rewrite the sum over all force derivatives Eq. 3.75 in the following way (where for brevity  $\sum_{\alpha} = \sum_{\alpha \in A}$  and  $\mu \in C, \kappa \in K, \beta \in B, \alpha \in A$ )

$$-4 \sum_K F_K = \tag{E.1}$$

$$= \sum_K \sum_{C>K} \sum_B \sum_{A>B} \sum_{\mu\kappa\alpha\beta} S_{\alpha\beta} (\Delta P_{\kappa\beta} \Delta P_{\mu\alpha} + \Delta P_{\kappa\alpha} \Delta P_{\mu\beta}) (D_K S_{\kappa\mu}) (\gamma_{KB} + \gamma_{CB} + \gamma_{KA} + \gamma_{CA}) \tag{E.2}$$

$$+ \sum_K \sum_{C>K} \sum_B \sum_{A=B} \sum_{\mu\kappa\alpha\beta} S_{\alpha\beta} (\Delta P_{\kappa\beta} \Delta P_{\mu\alpha} + \Delta P_{\kappa\alpha} \Delta P_{\mu\beta}) (D_K S_{\kappa\mu}) (\gamma_{KB} + \gamma_{CB}) \tag{E.3}$$

$$+ \sum_K \sum_{C>K} \sum_B \sum_{A>B} \sum_{\mu\kappa\alpha\beta} S_{\alpha\beta} (\Delta P_{\mu\alpha} \Delta P_{\kappa\beta} + \Delta P_{\mu\beta} \Delta P_{\kappa\alpha}) (\gamma_{CB} + \gamma_{KB} + \gamma_{CA} + \gamma_{KA}) (D_C S_{\mu\kappa}) \tag{E.4}$$

$$+ \sum_K \sum_{C>K} \sum_B \sum_{A=B} \sum_{\mu\kappa\alpha\beta} S_{\alpha\beta} (\Delta P_{\mu\beta} \Delta P_{\kappa\alpha} + \Delta P_{\mu\alpha} \Delta P_{\kappa\beta}) (D_C S_{\mu\kappa}) (\gamma_{CB} + \gamma_{KB}) \tag{E.5}$$

$$+ \sum_K \sum_{C \geq K} \sum_B \sum_{\substack{A \geq B \\ A \neq K}} \sum_{\mu\kappa\alpha\beta} S_{\alpha\beta} S_{\kappa\mu} (\Delta P_{\kappa\beta} \Delta P_{\mu\alpha} + \Delta P_{\kappa\alpha} \Delta P_{\mu\beta}) (D_K \gamma_{KA}) \tag{E.6}$$

$$+ \sum_K \sum_{C \geq K} \sum_B \sum_{\substack{A > B \\ B \neq K}} \sum_{\mu\kappa\alpha\beta} S_{\beta\alpha} S_{\kappa\mu} (\Delta P_{\kappa\alpha} \Delta P_{\mu\beta} + \Delta P_{\kappa\beta} \Delta P_{\mu\alpha}) (D_K \gamma_{KB}) \tag{E.7}$$

$$+ \sum_K \sum_{C>K} \sum_B \sum_{\substack{A \geq B \\ A \neq C}} \sum_{\mu\kappa\alpha\beta} S_{\alpha\beta} S_{\kappa\mu} (\Delta P_{\kappa\beta} \Delta P_{\mu\alpha} + \Delta P_{\kappa\alpha} \Delta P_{\mu\beta}) (D_C \gamma_{CA}) \tag{E.8}$$

$$+ \sum_K \sum_{C>K} \sum_B \sum_{\substack{A > B \\ B \neq C}} \sum_{\mu\kappa\alpha\beta} S_{\beta\alpha} S_{\kappa\mu} (\Delta P_{\kappa\alpha} \Delta P_{\mu\beta} + \Delta P_{\kappa\beta} \Delta P_{\mu\alpha}) (D_C \gamma_{CB}). \tag{E.9}$$

Note that sums like  $\sum_{A=B}^M = \sum_{A=1}^M \delta_{AB}$  contain only one summand for  $A = B$ . In following we present the pseudo code of the routine as it is implemented in the DFTB+ code.

```

do K=1, M
  do C ∈ N(K)
    do B=1, M
      do A ∈ N(B)
         $\gamma_1 = \gamma_{KB} + \gamma_{CB}$ 
         $\gamma_2 = \gamma_{KB} + \gamma_{CB} + \gamma_{KA} + \gamma_{CA}$ 
         $F = \sum_{\mu \in C} \sum_{\kappa \in K} \sum_{\alpha \in A} \sum_{\beta \in B} S_{\beta\alpha} (\Delta P_{\beta\kappa} \Delta P_{\alpha\mu} + \Delta P_{\alpha\kappa} \Delta P_{\beta\mu}) (D_K S_{\kappa\mu})$ 
         $F_r = \sum_{\mu \in C} \sum_{\kappa \in K} \sum_{\alpha \in A} \sum_{\beta \in B} S_{\beta\alpha} (\Delta P_{\beta\kappa} \Delta P_{\alpha\mu} + \Delta P_{\alpha\kappa} \Delta P_{\beta\mu}) (D_C S_{\mu\kappa})$ 
         $F_2 = \sum_{\mu \in C} \sum_{\kappa \in K} \sum_{\alpha \in A} \sum_{\beta \in B} S_{\kappa\mu} S_{\beta\alpha} (\Delta P_{\beta\kappa} \Delta P_{\alpha\mu} + \Delta P_{\alpha\kappa} \Delta P_{\beta\mu})$ 
        if K ≠ C
          if B ≠ A
             $F = F \cdot \gamma_2 + F_2 \cdot (D_K \gamma_{KA} + D_K \gamma_{KB})$ 
             $F_r = F_r \cdot \gamma_2 + F_2 \cdot (D_C \gamma_{CA} + D_C \gamma_{CB})$ 
          else
             $F = F \cdot \gamma_1 + F_2 \cdot D_K \gamma_{KA}$ 
             $F_r = F_r \cdot \gamma_1 + F_2 \cdot D_C \gamma_{CA}$ 
          end if
        else
          if A ≠ B
             $F = F + F_2 \cdot (D_K \gamma_{KA} + D_K \gamma_{KB})$ 
          else
             $F = F + F_2 \cdot D_K \gamma_{KA}$ 
          end if
        end if
         $F_K = F_K - 0.25F$ 
         $F_C = F_C - 0.25F_r$ 
      end do
    end do
  end do
end do
end do
end do
    
```

## BIBLIOGRAPHY

- [1] C. ADAMO, M. ERNZERHOF, AND G. E. SCUSERIA, *The meta-GGA functional: Thermodynamics with a kinetic energy density dependent exchange-correlation functional*, *The Journal of Chemical Physics*, 112 (2000), pp. 2643–2649.
- [2] Y. AKINAGA AND S. TEN-NO, *Range-separation by the Yukawa potential in long-range corrected density functional theory with Gaussian-type basis functions*, *Chemical Physics Letters*, 462 (2008), pp. 348–351.
- [3] C.-O. ALMBLADH AND U. VON BARTH, *Exact results for the charge and spin densities, exchange-correlation potentials, and density-functional eigenvalues*, *Phys. Rev. B*, 31 (1985), pp. 3231–3244.
- [4] J. ALMLÖF, K. FAEGRI, AND K. KORSELL, *Principles for a direct SCF approach to LICAO-MO ab-initio calculations*, *Journal of Computational Chemistry*, 3 (1982), pp. 385–399.
- [5] E. ANDERSON, Z. BAI, C. BISCHOF, S. BLACKFORD, J. DEMMEL, J. DONGARRA, J. DU CROZ, A. GREENBAUM, S. HAMMARLING, A. MCKENNEY, AND D. SORENSEN, *LAPACK Users' Guide*, Society for Industrial and Applied Mathematics, Philadelphia, PA, third ed., 1999.
- [6] B. ARADI, B. HOURAHINE, AND T. FRAUENHEIM, *DFTB+, a sparse matrix-based implementation of the DFTB method*, *J. Phys. Chem. A*, 111 (2007), pp. 5678–5684.
- [7] F. ARYASETIAWAN AND O. GUNNARSSON, *The GW method*, *Reports on Progress in Physics*, 61 (1998), p. 237.
- [8] O. AYED, E. BERNARD, AND B. SILVI, *On the Mulliken approximation of multicentre integrals*, *Journal of Molecular Structure: THEOCHEM*, 135 (1986), pp. 159–168.
- [9] R. BAER, E. LIVSHITS, AND U. SALZNER, *Tuned range-separated hybrids in density functional theory*, *Annu. Rev. Phys. Chem.*, 61 (2010), pp. 85–109.
- [10] R. BAER AND D. NEUHAUSER, *Density functional theory with correct long-range asymptotic behavior*, *Phys. Rev. Lett.*, 94 (2005), p. 043002.

## BIBLIOGRAPHY

---

- [11] T. BALLY AND G. N. SASTRY, *Incorrect dissociation behavior of radical ions in density functional calculations*, *The Journal of Physical Chemistry A*, 101 (1997), pp. 7923–7925.
- [12] G. BARBARELLA, L. FAVARETTO, G. SOTGIU, M. ZAMBIANCHI, A. BONGINI, C. ARBIZZANI, M. MASTRAGOSTINO, M. ANNI, G. GIGLI, AND R. CINGOLANI, *Tuning solid-state photoluminescence frequencies and efficiencies of oligomers containing one central thiophene-S,S-dioxide unit*, *Journal of the American Chemical Society*, 122 (2000), pp. 11971–11978.
- [13] A. D. BECKE, *A multicenter numerical integration scheme for polyatomic molecules*, *The Journal of Chemical Physics*, 88 (1988), pp. 2547–2553.
- [14] ———, *Density-functional exchange-energy approximation with correct asymptotic behavior*, *Phys. Rev. A*, 38 (1988), pp. 3098–3100.
- [15] ———, *Density-functional thermochemistry. III. The role of exact exchange*, *The Journal of Chemical Physics*, 98 (1993), pp. 5648–5652.
- [16] ———, *Density-functional thermochemistry. IV. A new dynamical correlation functional and implications for exact-exchange mixing*, *The Journal of Chemical Physics*, 104 (1996), pp. 1040–1046.
- [17] ———, *Real-space post-Hartree-Fock correlation models*, *The Journal of Chemical Physics*, 122 (2005), p. 064101.
- [18] A. D. BECKE AND R. M. DICKSON, *Numerical solution of Poisson’s equation in polyatomic molecules*, *The Journal of Chemical Physics*, 89 (1988), pp. 2993–2997.
- [19] W. G. BICKLEY, *Math. Gaz.*, 25 (1941), p. 19.
- [20] X. BLASE, C. ATTACCALITE, AND V. OLEVANO, *First-principles GW calculations for fullerenes, porphyrins, phtalocyanine, and other molecules of interest for organic photovoltaic applications*, *Phys. Rev. B*, 83 (2011), p. 115103.
- [21] Z. BODROG, B. ARADI, AND T. FRAUENHEIM, *Automated repulsive parametrization for the DFTB method*, *Journal of Chemical Theory and Computation*, 7 (2011), pp. 2654–2664.
- [22] G. BOURHILL, J.-L. BREDAS, L.-T. CHENG, S. R. MARDER, F. MEYERS, J. W. PERRY, AND B. G. TIEMANN, *Experimental demonstration of the dependence of the first hyperpolarizability of donor-acceptor-substituted polyenes on the ground-state polarization and bond length alternation*, *Journal of the American Chemical Society*, 116 (1994), pp. 2619–2620.

- [23] C. J. BRABEC, N. S. SARICIFTCI, AND J. C. HUMMELEN, *Plastic solar cells*, *Advanced Functional Materials*, 11 (2001), pp. 15–26.
- [24] J. L. BREDAS, *Relationship between band gap and bond length alternation in organic conjugated polymers*, *The Journal of Chemical Physics*, 82 (1985), pp. 3808–3811.
- [25] J. L. BREDAS, R. R. CHANCE, AND R. SILBEY, *Comparative theoretical study of the doping of conjugated polymers: Polarons in polyacetylene and polyparaphenylene*, *Phys. Rev. B*, 26 (1982), pp. 5843–5854.
- [26] J. L. BREDAS AND G. B. STREET, *Polarons, bipolarons, and solitons in conducting polymers*, *Accounts of Chemical Research*, 18 (1985), pp. 309–315.
- [27] J. L. BREDAS, B. THEMANS, J. G. FRIPIAT, J. M. ANDRE, AND R. R. CHANCE, *Highly conducting polyparaphenylene, polypyrrole, and polythiophene chains: An ab initio study of the geometry and electronic-structure modifications upon doping*, *Phys. Rev. B*, 29 (1984), pp. 6761–6773.
- [28] S. CAPPONI, N. GUIHERY, J.-P. MALRIEU, B. MIGUEL, AND D. POILBLANC, *Bond alternation of polyacetylene as a spin-Peierls distortion*, *Chemical Physics Letters*, 255 (1996), pp. 238–243.
- [29] M. E. CASIDA AND D. R. SALAHUB, *Asymptotic correction approach to improving approximate exchange-correlation potentials: Time-dependent density-functional theory calculations of molecular excitation spectra*, *The Journal of Chemical Physics*, 113 (2000), pp. 8918–8935.
- [30] D. M. CEPERLEY AND B. J. ALDER, *Ground state of the electron gas by a stochastic method*, *Phys. Rev. Lett.*, 45 (1980), pp. 566–569.
- [31] B. CHAMPAGNE, E. A. PERPETE, S. J. A. VAN GISBERGEN, E.-J. BAERENDS, J. G. SNIJDERS, C. SOUBRA-GHAOUI, K. A. ROBINS, AND B. KIRTMAN, *Assessment of conventional density functional schemes for computing the polarizabilities and hyperpolarizabilities of conjugated oligomers: An ab initio investigation of polyacetylene chains*, *The Journal of Chemical Physics*, 109 (1998), pp. 10489–10498.
- [32] H. N. CHAPMAN, A. BARTY, M. J. BOGAN, S. BOUTET, M. FRANK, S. P. HAU-RIEGE, S. MARCHESINI, B. W. WOODS, S. BAJT, AND W. H. BENNER, *Femtosecond diffractive imaging with a soft-x-ray free-electron laser*, *Nature Physics*, 2 (2006), pp. 839–843.
- [33] D. P. CHONG, O. V. GRITSENKO, AND E. J. BAERENDS, *Interpretation of the Kohn-Sham orbital energies as approximate vertical ionization potentials*, *The Journal of Chemical Physics*, 116 (2002), pp. 1760–1772.

- [34] A. J. COHEN, P. MORI-SANCHEZ, AND W. YANG, *Challenges for density functional theory*, Chemical Reviews, 112 (2012), pp. 289–320.
- [35] Q. CUI AND M. ELSTNER, *Density functional tight binding: values of semi-empirical methods in an ab initio era*, Phys. Chem. Chem. Phys., 16 (2014), pp. 14368–14377.
- [36] P. DEÁK, B. ARADI, T. FRAUENHEIM, E. JANZÉN, AND A. GALI, *Accurate defect levels obtained from the HSE06 range-separated hybrid functional*, Phys. Rev. B, 81 (2010), p. 153203.
- [37] G. D. DICKENSON, M. L. NIU, E. J. SALUMBIDES, J. KOMASA, K. S. E. EIKEMA, K. PACHUCKI, AND W. UBACHS, *Fundamental vibration of molecular hydrogen*, Phys. Rev. Lett., 110 (2013), p. 193601.
- [38] N. DORI, M. MENON, L. KILIAN, M. SOKOŁOWSKI, L. KRONIK, AND E. UMBACH, *Valence electronic structure of gas-phase 3,4,9,10-perylene tetracarboxylic acid dianhydride: Experiment and theory*, Phys. Rev. B, 73 (2006), p. 195208.
- [39] T. H. DUNNING, *Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen*, The Journal of Chemical Physics, 90 (1989), pp. 1007–1023.
- [40] H. E. KATZ, *Organic molecular solids as thin film transistor semiconductors*, J. Mater. Chem., 7 (1997), pp. 369–376.
- [41] M. ELSTNER, *Weiterentwicklung quantenmechanischer Rechenverfahren für organische Moleküle und Polymere.*, PhD thesis, Universität Paderborn, 1998.
- [42] M. ELSTNER, D. POREZAG, G. JUNGNICHEL, J. ELSNER, M. HAUGK, T. FRAUENHEIM, S. SUHAI, AND G. SEIFERT, *Self-consistent-charge density-functional tight-binding method for simulations of complex materials properties*, Phys. Rev. B, 58 (1998), pp. 7260–7268.
- [43] M. ELSTNER AND G. SEIFERT, *Density functional tight binding*, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 372 (2014), p. 20120483.
- [44] ELSTNER M., *The SCC-DFTB method and its application to biological systems*, Theor. Chem. Acc., 116 (2006), pp. 316–325.
- [45] E. ENGEL AND R. N. SCHMID, *Insulating ground states of transition-metal monoxides from exact exchange*, Phys. Rev. Lett., 103 (2009), p. 036404.
- [46] H. ESCHRIG AND I. BERGERT, *An optimized LCAO version for band structure calculations application to copper*, physica status solidi (b), 90 (1978), pp. 621–628.



- [47] D. G. FEDOROV, T. NAGATA, AND K. KITaura, *Exploring chemistry with the fragment molecular orbital method*, Phys. Chem. Chem. Phys., 14 (2012), pp. 7562–7577.
- [48] J. FENN, M. MANN, C. MENG, S. WONG, AND C. WHITEHOUSE, *Electrospray ionization for mass spectrometry of large biomolecules*, Science, 246 (1989), pp. 64–71.
- [49] C. R. FINCHER, C. E. CHEN, A. J. HEEGER, A. G. MACDIARMID, AND J. B. HASTINGS, *Structural determination of the symmetry-breaking parameter in trans  $-(\text{CH})_x$* , Phys. Rev. Lett., 48 (1982), pp. 100–104.
- [50] T. FRAUENHEIM, G. SEIFERT, M. ELSTERNER, Z. HAJNAL, G. JUNGnickEL, D. POREZAG, S. SUHAI, AND R. SCHOLZ, *A self-consistent charge density-functional based tight-binding method for predictive materials simulations in physics, chemistry and biology*, physica status solidi (b), 217 (2000), pp. 41–62.
- [51] F. GARNIER, R. HAJLAOUI, A. YASSAR, AND P. SRIVASTAVA, *All-polymer field-effect transistor realized by printing techniques*, Science, 265 (1994), pp. 1684–1686.
- [52] M. GAUS, Q. CUI, AND M. ELSTNER, *DFTB3: Extension of the self-consistent-charge density-functional tight-binding method (SCC-DFTB)*, Journal of Chemical Theory and Computation, 7 (2011), pp. 931–948.
- [53] I. C. GERBER AND J. G. ANGYAN, *Hybrid functional with separated range*, Chemical Physics Letters, 415 (2005), pp. 100–105.
- [54] P. M. W. GILL, R. D. ADAMASON, AND J. A. POPLE, *Coulomb-attenuated exchange energy density functionals*, Molecular Physics, 88 (1996), pp. 1005–1009.
- [55] X. GONZE, P. GHOSEZ, AND R. W. GODBY, *Density-polarization functional theory of the response of a periodic insulating solid to an electric field*, Phys. Rev. Lett., 74 (1995), pp. 4035–4038.
- [56] C. M. GORINGE, D. R. BOWLER, AND E. HERNANDEZ, *Tight-binding modelling of materials*, Reports on Progress in Physics, 60 (1997), p. 1447.
- [57] A. GÖRLING, *Density-functional theory for excited states*, Phys. Rev. A, 54 (1996), pp. 3912–3915.
- [58] J. GRÄFENSTEIN, E. KRAKA, AND D. CREMER, *The impact of the self-interaction error on the density functional theory description of dissociating radical cations: Ionic and covalent dissociation limits*, The Journal of Chemical Physics, 120 (2004), pp. 524–539.
- [59] B. GRUNDKÖTTER-STOCK, V. BEZUGLY, J. KUNSTMANN, G. CUNIBERTI, T. FRAUENHEIM, AND T. A. NIEHAUS, *SCC-DFTB parametrization for boron and boranes*, Journal of Chemical Theory and Computation, 8 (2012), pp. 1153–1163.

## BIBLIOGRAPHY

---

- [60] O. GUNNARSSON AND B. I. LUNDQVIST, *Exchange and correlation in atoms, molecules, and solids by the spin-density-functional formalism*, Phys. Rev. B, 13 (1976), pp. 4274–4298.
- [61] M. HALIK, H. KLAUK, U. ZSCHIESCHANG, G. SCHMID, S. PONOMARENKO, S. KIRCHMEYER, AND W. WEBER, *Relationship between molecular structure and electrical performance of oligothiophene organic thin film transistors*, Advanced Materials, 15 (2003), pp. 917–922.
- [62] G. G. HALL, *The molecular orbital theory of chemical valency. VIII. a method of calculating ionization potentials*, Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 205 (1951), pp. 541–552.
- [63] F. E. HARRIS, *Analytic evaluation of two-center STO electron repulsion integrals via ellipsoidal expansion*, International Journal of Quantum Chemistry, 88 (2002), pp. 701–734.
- [64] M. HÄSER AND R. AHLRICHS, *Improvements on the direct SCF method*, Journal of Computational Chemistry, 10 (1989), pp. 104–111.
- [65] W. HAYNES, *CRC Handbook of Chemistry and Physics, 93rd Edition*, CRC Handbook of Chemistry and Physics, Taylor & Francis, 2012.
- [66] L. HEDIN, *New method for calculating the one-particle Green's function with application to the electron-gas problem*, Phys. Rev., 139 (1965), pp. A796–A823.
- [67] J. HEYD, G. SCUSERIA, AND M. ERNZERHOF, *Hybrid functionals based on a screened Coulomb potential*, J. Chem. Phys., 118 (2003), p. 8207.
- [68] J. HEYD AND G. E. SCUSERIA, *Efficient hybrid density functional calculations in solids: Assessment of the Heyd-Scuseria-Ernzerhof screened Coulomb hybrid functional*, The Journal of Chemical Physics, 121 (2004), pp. 1187–1192.
- [69] J. HEYD, G. E. SCUSERIA, AND M. ERNZERHOF, *Erratum: "Hybrid functionals based on a screened Coulomb potential" [J. Chem. Phys. 118, 8207 (2003)]*, The Journal of Chemical Physics, 124 (2006), p. 219906.
- [70] S. HIDEO, Y. MAEDA, AND M. KAMIYA, *Influence of the long-range exchange effect on dynamic polarizability*, Molecular Physics, 103 (2005), pp. 2183–2189.
- [71] C. HO CHOI, M. KERTESZ, AND A. KARPFFEN, *The effects of electron correlation on the degree of bond alternation and electronic structure of oligomers of polyacetylene*, The Journal of Chemical Physics, 107 (1997), pp. 6712–6721.
- [72] P. HOHENBERG AND W. KOHN, *Inhomogeneous electron gas*, Phys. Rev., 136 (1964), pp. B864–B871.

- [73] H. H. HOMEIER AND E. STEINBORN, *Some properties of the coupling coefficients of real spherical harmonics and their relation to Gaunt coefficients*, Journal of Molecular Structure: THEOCHEM, 368 (1996), pp. 31–37.
- [74] S. HONDA, K. YAMASAKI, Y. SAWADA, AND H. MORII, *10 residue folded peptide designed by segment statistics*, Structure, 12 (2004), pp. 1507–1518.
- [75] H. HOPPE AND N. S. SARICIFTCI, *Organic solar cells: An overview*, Journal of Materials Research, 19 (2004), pp. 1924–1945.
- [76] A. P. HORSFIELD AND A. M. BRATKOVSKY, *Ab initio tight binding*, Journal of Physics: Condensed Matter, 12 (2000), p. R1.
- [77] B. HOURAHINE, S. SANNA, B. ARADI, C. KÖHLER, T. NIEHAUS, AND T. FRAUENHEIM, *Self-interaction and strong correlation in DFTB+*, The Journal of Physical Chemistry A, 111 (2007), pp. 5671–5677.
- [78] K. HUMMER, A. GRÜNEIS, AND G. KRESSE, *Structural and electronic properties of lead chalcogenides from first principles*, Phys. Rev. B, 75 (2007), p. 195211.
- [79] W. HUMPHREY, A. DALKE, AND K. SCHULTEN, *VMD – Visual Molecular Dynamics*, Journal of Molecular Graphics, 14 (1996), pp. 33–38.
- [80] G. R. HUTCHISON, Y.-J. ZHAO, B. DELLEY, A. J. FREEMAN, M. A. RATNER, AND T. J. MARKS, *Electronic structure of conducting polymers: Limitations of oligomer extrapolation approximations and effects of heteroatoms*, Phys. Rev. B, 68 (2003), p. 035204.
- [81] H. IIKURA, T. TSUNEDA, T. YANAI, AND K. HIRAO, *A long-range correction scheme for generalized-gradient-approximation exchange functionals*, J. Chem. Phys., 115 (2001), p. 3540.
- [82] D. JACQUEMIN AND C. ADAMO, *Bond length alternation of conjugated oligomers: Wave function and DFT benchmarks*, Journal of Chemical Theory and Computation, 7 (2011), pp. 369–376.
- [83] D. JACQUEMIN, A. FEMENIAS, H. CHERMETTE, I. CIOFINI, C. ADAMO, J.-M. ANDRE, AND E. A. PERPETE, *Assessment of several hybrid DFT functionals for the evaluation of bond length alternation of increasingly long oligomers*, The Journal of Physical Chemistry A, 110 (2006), pp. 5952–5959.
- [84] D. JACQUEMIN, E. A. PERPETE, I. CIOFINI, AND C. ADAMO, *Assessment of recently developed density functional approaches for the evaluation of the bond length alternation in polyacetylene*, Chemical Physics Letters, 405 (2005), pp. 376–381.

- [85] D. JACQUEMIN, E. A. PERPETE, G. SCALMANI, M. J. FRISCH, R. KOBAYASHI, AND C. ADAMO, *Assessment of the efficiency of long-range corrected functionals for some properties of large compounds*, The Journal of Chemical Physics, 126 (2007), p. 144105.
- [86] M. JAIN, J. R. CHELIKOWSKY, AND S. G. LOUIE, *Reliability of hybrid functionals in predicting band gaps*, Phys. Rev. Lett., 107 (2011), p. 216806.
- [87] J. F. JANAK, *Proof that  $\frac{\partial e}{\partial n_i} = \epsilon$  in density-functional theory*, Phys. Rev. B, 18 (1978), pp. 7165–7168.
- [88] S. KAMINSKI, M. GAUS, AND M. ELSTNER, *Improved electronic properties from third-order SCC-DFTB with cost efficient post-SCF extensions*, The Journal of Physical Chemistry A, 116 (2012), pp. 11927–11937.
- [89] S. KAMINSKI, T. J. GIESE, M. GAUS, D. M. YORK, AND M. ELSTNER, *Extended polarization in third-order SCC-DFTB from chemical-potential equalization*, The Journal of Physical Chemistry A, 116 (2012), pp. 9131–9141.
- [90] M. KAMIYA, H. SEKINO, T. TSUNEDA, AND K. HIRAO, *Nonlinear optical property calculations by the long-range-corrected coupled-perturbed Kohn-Sham method*, The Journal of Chemical Physics, 122 (2005), p. 234111.
- [91] R. KAR, J.-W. SONG, AND K. HIRAO, *Long-range corrected functionals satisfy Koopmans’ theorem: Calculation of correlation and relaxation energies*, Journal of Computational Chemistry, 34 (2013), pp. 958–964.
- [92] A. KAROLEWSKI, R. ARMIENTO, AND S. KÜMMEL, *Polarizabilities of polyacetylene from a field-counteracting semilocal functional*, Journal of Chemical Theory and Computation, 5 (2009), pp. 712–718.
- [93] A. KAROLEWSKI, L. KRONIK, AND S. KÜMMEL, *Using optimally tuned range separated hybrid functionals in ground-state calculations: Consequences and caveats*, The Journal of Chemical Physics, 138 (2013), p. 204115.
- [94] F. KJELDSSEN, O. A. SILIVRA, AND R. A. ZUBAREV, *Zwitterionic states in gas-phase polypeptide ions revealed by 157-nm ultra-violet photodissociation*, Chemistry - A European Journal, 12 (2006), pp. 7920–7928.
- [95] C. KÖHLER, T. FRAUENHEIM, B. HOURAHINE, G. SEIFERT, AND M. STERNBERG, *Treatment of collinear and noncollinear electron spin within an approximate density functional based method*, The Journal of Physical Chemistry A, 111 (2007), pp. 5622–5629.
- [96] C. KOHLER, G. SEIFERT, U. GERSTMANN, M. ELSTNER, H. OVERHOF, AND T. FRAUENHEIM, *Approximate density-functional calculations of spin densities in large molecular systems and complex solids*, Phys. Chem. Chem. Phys., 3 (2001), pp. 5109–5114.

- [97] W. KOHN, *Nobel lecture: Electronic structure of matter - wave functions and density functionals*, Rev. Mod. Phys., 71 (1999), pp. 1253–1266.
- [98] W. KOHN AND L. J. SHAM, *Self-consistent equations including exchange and correlation effects*, Phys. Rev., 140 (1965), pp. A1133–A1138.
- [99] T. KÖRZDÖRFER, R. M. PARRISH, N. MAROM, J. S. SEARS, C. D. SHERRILL, AND J.-L. BREDAS, *Assessment of the performance of tuned range-separated hybrid density functionals in predicting accurate quasiparticle spectra*, Phys. Rev. B, 86 (2012), p. 205110.
- [100] T. KÖRZDÖRFER, R. M. PARRISH, J. S. SEARS, C. D. SHERRILL, AND J.-L. BREDAS, *On the relationship between bond-length alternation and many-electron self-interaction error*, The Journal of Chemical Physics, 137 (2012), p. 124305.
- [101] T. KÖRZDÖRFER, S. TRETIAK, AND S. KÜMMEL, *Fluorescence quenching in an organic donor-acceptor dyad: A first principles study*, The Journal of Chemical Physics, 131 (2009), p. 034310.
- [102] L. KRONIK, A. MAKMAL, M. L. TIAGO, M. M. G. ALEMANY, M. JAIN, X. HUANG, Y. SAAD, AND J. R. CHELIKOWSKY, *PARSEC - the pseudopotential algorithm for real-space electronic structure calculations: Recent advances and novel applications to nanostructures*, physica status solidi (b), 243 (2006), pp. 1063–1079.
- [103] L. KRONIK, T. STEIN, S. REFAELY-ABRAMSON, AND R. BAER, *Excitation gaps of finite-sized systems from optimally tuned range-separated hybrid functionals*, Journal of Chemical Theory and Computation, 8 (2012), pp. 1515–1531.
- [104] T. KUBAR, Z. BODROG, M. GAUS, C. KÖHLER, B. ARADI, T. FRAUENHEIM, AND M. ELSTNER, *Parametrization of the SCC-DFTB method for halogens*, Journal of Chemical Theory and Computation, 9 (2013), pp. 2939–2949.
- [105] S. KÜMMEL AND L. KRONIK, *Orbital-dependent density functionals: Theory and applications*, Rev. Mod. Phys., 80 (2008), pp. 3–60.
- [106] S. KURTH, J. P. PERDEW, AND P. BLAHA, *Molecular and solid-state tests of density functional approximations: LSD, GGAs, and meta-GGAs*, International Journal of Quantum Chemistry, 75 (1999), pp. 889–909.
- [107] D. LANGRETH AND J. PERDEW, *The exchange-correlation energy of a metallic surface*, Solid State Communications, 17 (1975), pp. 1425–1429.
- [108] V. LEBEDEV, Zh. Vychisl. Mat. Mat. Fiz., 15 (1975), p. 48.
- [109] C. LEE, W. YANG, AND R. G. PARR, *Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density*, Phys. Rev. B, 37 (1988), pp. 785–789.

## BIBLIOGRAPHY

---

- [110] T. LEININGER, H. STOLL, H.-J. WERNER, AND A. SAVIN, *Combining long-range configuration interaction with short-range density functionals*, Chem. Phys. Lett., 275 (1997), p. 151.
- [111] M. LEVY, *Universal variational functionals of electron densities, first-order density matrices, and natural spin-orbitals and solution of the  $v$ -representability problem*, Proceedings of the National Academy of Sciences, 76 (1979), pp. 6062–6065.
- [112] ———, *Electron densities in search of hamiltonians*, Phys. Rev. A, 26 (1982), pp. 1200–1208.
- [113] ———, *Excitation energies from density-functional orbital energies*, Phys. Rev. A, 52 (1995), pp. R4313–R4315.
- [114] S. M. LINDSAY AND M. A. RATNER, *Molecular transport junctions: Clearing mists*, Advanced Materials, 19 (2007), pp. 23–31.
- [115] P. LINSTROM AND W. MALLARD, *NIST chemistry webbook, NIST standard reference database number 69*, National Institute of Standards and Technology.
- [116] E. LIVSHITS AND R. BAER, *A well-tempered density functional theory of electrons in molecules*, Phys. Chem. Chem. Phys., 9 (2007), pp. 2932–2941.
- [117] LONGUET-HIGGINS H. C. AND SALEM L., *The alternation of bond lengths in long conjugated chain molecules*, Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 251 (1959), pp. 172–185.
- [118] M. LUNDBERG, Y. NISHIMOTO, AND S. IRLE, *Delocalization errors in a Hubbard-like model: Consequences for density-functional tight-binding calculations of molecular systems*, International Journal of Quantum Chemistry, 112 (2012), pp. 1701–1711.
- [119] V. LUTSKER, B. ARADI, AND T. A. NIEHAUS, *Implementation and benchmark of a long-range corrected functional in the density functional based tight-binding method*, ArXiv e-prints, (2015).
- [120] R. MARCHESE, R. GRANDORI, P. CARLONI, AND S. RAUGEI, *On the zwitterionic nature of gas-phase peptides and protein ions*, PLoS Comput. Biol., 6 (2010), p. e1000775.
- [121] S. R. MARDER, C. B. GORMAN, F. MEYERS, J. W. PERRY, G. BOURHILL, J.-L. BREDAS, AND B. M. PIERCE, *A unified description of linear and nonlinear polarization in organic polymethine dyes*, Science, 265 (1994), pp. 632–635.
- [122] S. R. MARDER, J. W. PERRY, B. G. TIEMANN, C. B. GORMAN, S. GILMOUR, S. L. BIDDLE, AND G. BOURHILL, *Direct observation of reduced bond-length alternation in donor/acceptor polyenes*, Journal of the American Chemical Society, 115 (1993), pp. 2524–2526.

- [123] N. MAROM, F. CARUSO, X. REN, O. T. HOFMANN, T. KÖRZDÖRFER, J. R. CHELIKOWSKY, A. RUBIO, M. SCHEFFLER, AND P. RINKE, *Benchmark of GW methods for azabenzenes*, Phys. Rev. B, 86 (2012), p. 245127.
- [124] M. A. MARQUES, M. J. OLIVEIRA, AND T. BURNUS, *Libxc: A library of exchange and correlation functionals for density functional theory*, Computer Physics Communications, 183 (2012), pp. 2272–2281.
- [125] M. MARSMAN, J. PAIER, A. STROPPA, AND G. KRESSE, *Hybrid functionals applied to extended systems*, Journal of Physics: Condensed Matter, 20 (2008), p. 064201.
- [126] T. MEYER, V. GABELICA, H. GRUBMÜLLER, AND M. OROZCO, *Proteins in the gas phase*, Wiley Interdisciplinary Reviews: Computational Molecular Science, 3 (2013), pp. 408–425.
- [127] U. MITSCHKE AND P. BAUERLE, *The electroluminescence of organic materials*, J. Mater. Chem., 10 (2000), pp. 1471–1507.
- [128] C. MØLLER AND M. S. PLESSET, *Note on an approximation treatment for many-electron systems*, Phys. Rev., 46 (1934), pp. 618–622.
- [129] M. B. MONAGAN, K. O. GEDDES, K. M. HEAL, G. LABAHN, S. M. VORKOETTER, J. MCCARRON, AND P. DEMARCO, *Maple 10 Programming Guide*, Maplesoft, Waterloo ON, Canada, 2005.
- [130] H. J. MONKHORST, *Calculation of properties with the coupled-cluster method*, International Journal of Quantum Chemistry, 12 (1977), pp. 421–432.
- [131] P. MORI-SANCHEZ, A. J. COHEN, AND W. YANG, *Self-interaction-free exchange-correlation functional for thermochemistry and kinetics*, The Journal of Chemical Physics, 124 (2006), p. 091102.
- [132] ———, *Localization and delocalization errors in density functional theory and implications for band-gap prediction*, Phys. Rev. Lett., 100 (2008), p. 146401.
- [133] G. MORO, G. SCALMANI, U. COSENTINO, AND D. PITEA, *On the structure of bipolaronic defects in thiophene oligomers: A density functional study*, Synthetic Metals, 92 (1998), pp. 69–73.
- [134] ———, *On the structure of polaronic defects in thiophene oligomers: A combined Hartree-Fock and density functional theory study*, Synthetic Metals, 108 (2000), pp. 165–172.
- [135] T. NAGATA, K. BRORSEN, D. G. FEDOROV, K. KITaura, AND M. S. GORDON, *Fully analytic energy gradient in the fragment molecular orbital method*, The Journal of Chemical Physics, 134 (2011), p. 124115.

## BIBLIOGRAPHY

---

- [136] J. W. NEIDIGH, R. M. FESINMEYER, AND N. H. ANDERSEN, *Designing a 20-residue protein*, Nat Struct Mol Biol, 9 (2002), pp. 425–430.
- [137] T. NIEHAUS AND F. DELLA SALA, *Range-separated functionals in the density-functional based tight-binding method: Formalism*, physica status solidi (b), 249 (2012), p. 237.
- [138] T. A. NIEHAUS, A. DI CARLO, AND T. FRAUENHEIM, *Effect of self-consistency and electron correlation on the spatial extension of bipolaronic defects*, Org. Elec., 5 (2004), pp. 167–174.
- [139] T. A. NIEHAUS, M. ELSTNER, T. FRAUENHEIM, AND S. SUHAI, *Application of an approximate density-functional method to sulfur containing compounds*, J. Mol. Struct. - Theochem, 541 (2001), pp. 185–194.
- [140] Y. NISHIMOTO, D. G. FEDOROV, AND S. IRLE, *Density-functional tight-binding combined with the fragment molecular orbital method*, Journal of Chemical Theory and Computation, 10 (2014), pp. 4801–4812.
- [141] Y. NISHIMOTO AND S. IRLE.  
private communication.
- [142] J. PAIER, M. MARSMAN, K. HUMMER, G. KRESSE, I. C. GERBER, AND J. G. ANGYAN, *Screened hybrid density functionals applied to solids*, The Journal of Chemical Physics, 124 (2006), p. 154709.
- [143] R. PARR AND W. YANG, *Density-functional theory of atoms and molecules*, International Series of Monographs on Chemistry, Oxford University Press, USA, 1989.
- [144] A. PATRIKSSON, C. M. ADAMS, F. KJELDSSEN, R. A. ZUBAREV, AND D. VAN DER SPOEL, *A direct comparison of protein structure in the gas and solution phase: The Trp-cage*, The Journal of Physical Chemistry B, 111 (2007), pp. 13147–13150.
- [145] J. E. PERALTA, J. HEYD, G. E. SCUSERIA, AND R. L. MARTIN, *Spin-orbit splittings and energy band gaps calculated with the Heyd-Scuseria-Ernzerhof screened hybrid functional*, Phys. Rev. B, 74 (2006), p. 073101.
- [146] J. P. PERDEW, K. BURKE, AND M. ERNZERHOF, *Generalized gradient approximation made simple*, Phys. Rev. Lett., 77 (1996), pp. 3865–3868.
- [147] J. P. PERDEW, M. ERNZERHOF, AND K. BURKE, *Rationale for mixing exact exchange with density functional approximations*, The Journal of Chemical Physics, 105 (1996), pp. 9982–9985.



- [148] J. P. PERDEW, R. G. PARR, M. LEVY, AND J. L. BALDUZ, *Density-functional theory for fractional particle number: Derivative discontinuities of the energy*, Phys. Rev. Lett., 49 (1982), pp. 1691–1694.
- [149] J. P. PERDEW, A. RUZSINSZKY, G. I. CSONKA, O. A. VYDROV, G. E. SCUSERIA, V. N. STAROVEROV, AND J. TAO, *Exchange and correlation in open systems of fluctuating electron number*, Phys. Rev. A, 76 (2007), p. 040501.
- [150] J. P. PERDEW AND A. ZUNGER, *Self-interaction correction to density-functional approximations for many-electron systems*, Phys. Rev. B, 23 (1981), pp. 5048–5079.
- [151] I. PEREPICHKA, D. PEREPICHKA, H. MENG, AND F. WUDL, *Light-emitting polythiophenes*, Advanced Materials, 17 (2005), pp. 2281–2305.
- [152] D. PINCHON AND P. E. HOGGAN, *New index functions for storing Gaunt coefficients*, International Journal of Quantum Chemistry, 107 (2007), pp. 2186–2196.
- [153] J. A. POPLÉ, D. L. BEVERIDGE, AND P. A. DOBOSH, *Approximate self-consistent molecular-orbital theory. V. intermediate neglect of differential overlap*, The Journal of Chemical Physics, 47 (1967), pp. 2026–2033.
- [154] J. A. POPLÉ AND G. A. SEGAL, *Approximate self-consistent molecular orbital theory. II. calculations with complete neglect of differential overlap*, The Journal of Chemical Physics, 43 (1965), pp. S136–S151.
- [155] D. POREZAG, T. FRAUENHEIM, T. KÖHLER, G. SEIFERT, AND R. KASCHNER, *Construction of tight-binding-like potentials on the basis of density-functional theory: Application to carbon*, Phys. Rev. B, 51 (1995), pp. 12947–12957.
- [156] W. D. PRICE, R. A. JOCKUSCH, AND E. R. WILLIAMS, *Is arginine a zwitterion in the gas phase?*, J. Am. Chem. Soc., 119 (1997), pp. 11988–11989.
- [157] A. RAMASUBRAMANIAM, D. NAVEH, AND E. TOWE, *Tunable band gaps in bilayer transition-metal dichalcogenides*, Phys. Rev. B, 84 (2011), p. 205325.
- [158] M. RAPACIOLI, F. SPIEGELMAN, A. SCEMAMA, AND A. MIRTSCHINK, *Modeling charge resonance in cationic molecular clusters: Combining DFT-tight binding with configuration interaction*, Journal of Chemical Theory and Computation, 7 (2011), pp. 44–55.
- [159] J. RASCH AND A. C. H. YU, *Efficient storage scheme for precalculated Wigner 3j, 6j and Gaunt coefficients*, SIAM Journal on Scientific Computing, 25 (2004), pp. 1416–1428.
- [160] S. REFAELY-ABRAMSON, R. BAER, AND L. KRONIK, *Fundamental and excitation gaps in molecules of relevance for organic photovoltaics from an optimally tuned range-separated hybrid functional*, Phys. Rev. B, 84 (2011), p. 075144.

- [161] S. REFAELY-ABRAMSON, S. SHARIFZADEH, N. GOVIND, J. AUTSCHBACH, J. B. NEATON, R. BAER, AND L. KRONIK, *Quasiparticle spectra from a nonempirical optimally tuned range-separated hybrid density functional*, Phys. Rev. Lett., 109 (2012), p. 226405.
- [162] J. E. ROBINSON, F. BASSANI, R. S. KNOX, AND J. R. SCHRIEFFER, *Screening correction to the Slater exchange potential*, Phys. Rev. Lett., 9 (1962), pp. 215–217.
- [163] M. A. ROHRDANZ, K. M. MARTINS, AND J. M. HERBERT, *A long-range-corrected density functional that performs well for both ground-state properties and time-dependent density functional theory excitation energies, including charge-transfer excited states*, The Journal of Chemical Physics, 130 (2009), p. 054112.
- [164] C. C. J. ROOTHAAN, *New developments in molecular orbital theory*, Rev. Mod. Phys., 23 (1951), pp. 69–89.
- [165] ———, *Self-consistent field theory for open shells of electronic systems*, Rev. Mod. Phys., 32 (1960), pp. 179–185.
- [166] C. C. J. ROOTHAAN, L. M. SACHS, AND A. W. WEISS, *Analytical self-consistent field functions for the atomic configurations  $1s^2$ ,  $1s^22s$ , and  $1s^22s^2$* , Rev. Mod. Phys., 32 (1960), pp. 186–194.
- [167] E. RUNGE AND E. K. U. GROSS, *Density-functional theory for time-dependent systems*, Phys. Rev. Lett., 52 (1984), pp. 997–1000.
- [168] U. SALZNER, J. B. LAGOWSKI, P. G. PICKUP, AND R. A. POIRIER, *Design of low band gap polymers employing density functional theory-hybrid functionals ameliorate band gap problem*, Journal of Computational Chemistry, 18 (1997), pp. 1943–1953.
- [169] U. SALZNER, P. G. PICKUP, R. A. POIRIER, AND J. B. LAGOWSKI, *Accurate method for obtaining band gaps in conducting polymers using a DFT/hybrid approach*, The Journal of Physical Chemistry A, 102 (1998), pp. 2572–2578.
- [170] J. SAUTHER, J. WÜSTEN, S. LACH, AND C. ZIEGLER, *Gas phase and bulk ultraviolet photoemission spectroscopy of 3,4,9,10-perylene-tetracarboxylic dianhydride, 1,4,5,8-naphthalene-tetracarboxylic dianhydride, and 1,8-naphthalene-dicarboxylic anhydride*, The Journal of Chemical Physics, 131 (2009), p. 034711.
- [171] A. SAVIN AND H.-J. FLAD, *Density functionals for the Yukawa electron-electron interaction*, International Journal of Quantum Chemistry, 56 (1995), pp. 327–332.
- [172] B. M. SAVOIE, N. E. JACKSON, T. J. MARKS, AND M. A. RATNER, *Reassessing the use of one-electron energetics in the design and characterization of organic photovoltaics*, Phys. Chem. Chem. Phys., 15 (2013), pp. 4538–4547.

- [173] K. SCHULTEN AND R. GORDON, *Recursive evaluation of 3j and 6j coefficients*, Computer Physics Communications, 11 (1976), pp. 269–278.
- [174] A. SEIDL, A. GÖRLING, P. VOGL, J. A. MAJEWSKI, AND M. LEVY, *Generalized Kohn-Sham schemes and the band-gap problem*, Phys. Rev. B, 53 (1996), pp. 3764–3774.
- [175] G. SEIFERT, H. ESCHRIG, AND W. BIEGER, *Z. Phys. Chem. (Leipzig)*, 267 (1986), p. 529.
- [176] H. SEKINO, Y. MAEDA, M. KAMIYA, AND K. HIRAO, *Polarizability and second hyperpolarizability evaluation of long molecules by the density functional theory with long-range correction*, The Journal of Chemical Physics, 126 (2007), p. 014107.
- [177] M. SETH AND T. ZIEGLER, *Range-separated exchange functionals with Slater-type functions*, Journal of Chemical Theory and Computation, 8 (2012), pp. 901–907.
- [178] S. SHARIFZADEH, A. BILLER, L. KRONIK, AND J. B. NEATON, *Quasiparticle and optical spectroscopy of the organic semiconductors pentacene and ptcda from first principles*, Phys. Rev. B, 85 (2012), p. 125307.
- [179] R. T. SHARP AND G. K. HORTON, *A variational approach to the unipotential many-electron problem*, Phys. Rev., 90 (1953), pp. 317–317.
- [180] K. B. SHELIMOV, D. E. CLEMMER, R. R. HUDGINS, AND M. F. JARROLD, *Protein structure in vacuo: Gas-phase conformations of BPTI and cytochrome c*, Journal of the American Chemical Society, 119 (1997), pp. 2240–2248.
- [181] C. D. SHERRILL AND H. SCHAEFER, *The configuration interaction method: Advances in highly correlated approaches*, Advances in Quantum Chemistry, 34 (1999), pp. 143–269.
- [182] S. I. SIMDYANKIN, T. A. NIEHAUS, G. NATARAJAN, T. FRAUENHEIM, AND S. R. ELLIOTT, *New type of charged defect in amorphous chalcogenides*, Phys. Rev. Lett., 94 (2005), p. 086401.
- [183] J. C. SLATER AND G. F. KOSTER, *Simplified LCAO method for the periodic potential problem*, Phys. Rev., 94 (1954), pp. 1498–1524.
- [184] J. M. SMITH, Y. JAMI ALAHMADI, AND C. N. ROWLEY, *Range-separated DFT functionals are necessary to model Thio-Michael additions*, Journal of Chemical Theory and Computation, 9 (2013), pp. 4860–4865.
- [185] P. J. STEPHENS, F. J. DEVLIN, C. F. CHABALOWSKI, AND M. J. FRISCH, *Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields*, The Journal of Physical Chemistry, 98 (1994), pp. 11623–11627.

## BIBLIOGRAPHY

---

- [186] W. P. SU, J. R. SCHRIEFFER, AND A. J. HEEGER, *Solitons in polyacetylene*, Phys. Rev. Lett., 42 (1979), pp. 1698–1701.
- [187] R. SUENRAM AND F. LOVAS, *Millimeter wave spectrum of glycine. A new conformer*, J. Am. Chem. Soc., 102 (1980), pp. 7180–7184.
- [188] H. SUN AND J. AUTSCHBACH, *Influence of the delocalization error and applicability of optimal functional tuning in density functional calculations of nonlinear optical properties of organic donor-acceptor chromophores*, ChemPhysChem, 14 (2013), pp. 2450–2461.
- [189] ———, *Electronic energy gaps for  $\pi$ -conjugated oligomers and polymers calculated with density functional theory*, Journal of Chemical Theory and Computation, 10 (2014), pp. 1035–1047.
- [190] M. TAKAHASHI AND J. PALDUS, *Bond-length alternation and vibrational spectra of polyacetylene*, Canadian Journal of Physics, 62 (1984), pp. 1226–1231.
- [191] J. D. TALMAN AND W. F. SHADWICK, *Optimized effective atomic central potential*, Phys. Rev. A, 14 (1976), pp. 36–40.
- [192] W. M. TEMMERMAN, Z. SZOTEK, AND H. WINTER, *Self-interaction-corrected electronic structure of  $\text{La}_2\text{CuO}_4$* , Phys. Rev. B, 47 (1993), pp. 11533–11536.
- [193] D. J. TOZER AND N. C. HANDY, *Improving virtual Kohn-Sham orbitals and eigenvalues: Application to excitation energies and static polarizabilities*, The Journal of Chemical Physics, 109 (1998), pp. 10180–10189.
- [194] G. TU, V. CARRAVETTA, O. VAHTRAS, AND H. AGREN, *Core ionization potentials from self-interaction corrected Kohn-Sham orbital energies*, The Journal of Chemical Physics, 127 (2007), p. 174110.
- [195] M. VALIEV, E. J. BYLASKA, N. GOVIND, K. KOWALSKI, T. P. STRAATSMA, H. J. VAN DAM, D. WANG, J. NIEPLOCHA, E. APRA, T. L. WINDUS, AND W. A. DE JONG, *NWChem: A comprehensive and scalable open-source solution for large scale molecular simulations*, Comput. Phys. Commun., 181 (2010), pp. 1477–1489.
- [196] S. J. A. VAN GISBERGEN, P. R. T. SCHIPPER, O. V. GRITSENKO, E. J. BAERENDS, J. G. SNIJDERS, B. CHAMPAGNE, AND B. KIRTMAN, *Electric field dependence of the exchange-correlation potential in molecular chains*, Phys. Rev. Lett., 83 (1999), pp. 694–697.
- [197] S. H. VOSKO, L. WILK, AND M. NUSAIR, *Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis*, Canadian Journal of Physics, 58 (1980), pp. 1200–1211.

- [198] O. A. VYDROV AND G. E. SCUSERIA, *Assessment of a long-range corrected hybrid functional*, The Journal of Chemical Physics, 125 (2006), p. 234109.
- [199] M. WEIMER, F. DELLA SALA, AND A. GÖRLING, *The Kohn-Sham treatment of anions via the localized Hartree-Fock method*, Chemical Physics Letters, 372 (2003), pp. 538–547.
- [200] T. YANAI, D. P. TEW, AND N. C. HANDY, *A new hybrid exchange-correlation functional using the Coulomb-attenuating method (CAM-B3LYP)*, Chemical Physics Letters, 393 (2004), pp. 51–57.
- [201] W. YANG, *Generalized adiabatic connection in density functional theory*, The Journal of Chemical Physics, 109 (1998), pp. 10107–10110.
- [202] C. S. YANNONI AND T. C. CLARKE, *Molecular geometry of cis- and trans-polyacetylene by nutation NMR spectroscopy*, Phys. Rev. Lett., 51 (1983), pp. 1191–1193.
- [203] S. S. ZADE AND M. BENDIKOV, *From oligomers to polymer: Convergence in the HOMO-LUMO gaps of conjugated oligomers*, Organic Letters, 8 (2006), pp. 5243–5246.
- [204] ———, *Theoretical study of long oligothiophene dications: Bipolaron vs polaron pair vs triplet state*, The Journal of Physical Chemistry B, 110 (2006), pp. 15839–15846.
- [205] Y. ZHANG AND W. YANG, *A challenge for density functionals: Self-interaction error increases for systems with a noninteger number of electrons*, The Journal of Chemical Physics, 109 (1998), pp. 2604–2608.



## ACKNOWLEDGEMENTS

In following I would like to thank

- *Prof. Dr. Thomas Niehaus*  
for providing me an opportunity to work on this exciting topic. His excellent supervision was essential for the success of the project.
- *Dr. Balint Aradi*  
for his support and useful discussions regarding the implementation of the Range-Separated Hybrids in the DFTB+ code.
- *Dr. Manohar Awasthi*  
for technical support and sage advise.
- *Prof. Dr. Klaus Richter*  
for refereeing this thesis.
- *Robert Hrdina, Angela Reisser and Toni Bienert*  
for support in administrative issues.
- *Deutsche Forschungsgemeinschaft (RTG 1570)*  
*and President of Universität Regensburg (Prof. Dr. Udo Hebel)*  
for financial support.

Finally I thank my parents and my brother for their patience and help.





## ERRATUM

Changes with respect to the refereed version

(Änderungen gegenüber der Abgabe der Dissertation):

- Added Acknowledgements
- p. 108, Eq. A.12: added missed factor 2
- p. 80, Fig. 7.2: label y-axis  $|\epsilon_H| - |IP|$  instead of  $|IP| - |\epsilon_H|$
- p. 80, bottom:  $\Delta = |\epsilon_{\text{HOMO}}| - |IP_{\text{exp}}|$  instead of  $\Delta = |IP_{\text{exp}}| - |\epsilon_{\text{HOMO}}|$

