N° d'ordre : 41026

THÉSE

Développement de modèle de potentiels effectifs d'interactions interatomiques pour la modélisation d'oxydes

Development of effective interatomic potentials for computer simulation of oxides

présentée à

l'Université Lille 1 - Sciences et Technologies

École doctorale : Science de la Matière, du Rayonnement et de l'Environnement

en vue de l'obtention du titre de

DOCTEUR

Mention : Optique et Lasers – Physico-Chimie – Atmosphère

par

Sergey SUKHOMLINOV

soutenue le 18 décembre 2012, devant la comission d'examen :

Président :	Prof. Sylvain CRISTOL, Professeur à l'Université Lille 1
Rapporteurs :	Dr. Andrey KALINICHEV, Directeur de Recherche, Ecole des Mines de Nantes Prof. Dr. Martin H. MÜSER, Professeur à l'Université des Saarlandes, Saarbrücken, Allemagne
Examinateurs :	Dr. Simona ISPAS, Maître de Conférences, HDR, Université Montpellier 2 Dr. Mathieu SALANNE, Maître de Conférences, Université Pierre et Marie Curie – Paris VI
Directeur de thèse :	Dr. Konstantin S. SMIRNOV, Chargé de Recherche au CNRS, HDR, Université Lille 1

Contents

1 Introduction. 5

2 Models and Methods. 9

2.1 Quantum-chemical models. 9

2.1.1 Schrödinger equation. 10

2.1.2 Adiabatic approximation. 11

2.1.3 Solving the Schrödinger equation. 12

2.1.4 Dealing with periodic systems. 21

2.1.5 Basis sets. 24

2.1.6 Pseudopotentials. 26

2.2 Computational methods. 27

2.2.1 Methods of energy minimization. 28

2.2.2 Molecular dynamics. 32

2.3 Numerical methods of parameters optimization. 38

2.3.1 Cost function. 38

2.3.2 Minimization methods without gradients. 38

2.3.3 Minimization methods using gradients. 40

3 Models of effective potentials in molecular modeling. 45

3.1 Types of effective potential model. 45

3.1.1 Ionic force fields. 45

3.1.2 Molecular mechanics force fields. 47

3.2 Introduction of polarization effect in force field. 49

3.2.1 Induced point dipoles model. 49

3.2.2 Shell model. 52

3.2.3 Deformable and polarizable ions model. 54

3.2.4 Chemical potential equalization model. 57

3.3 Polarizable force field: model and approaches. 61

4 Modelling long-range interactions in silicates 63

4.1 Introduction 63

4.2 Theoretical models and computational details . . . 65
4.2.1 EEM and SQE models . . . 65
4.2.2 Computational schemes . . . 68
4.2.3 Training and validation sets . . . 68
4.2.4 Computational details . . . 69
4.2.5 Parametrization strategy . . . 70
4.3 Results and discussion . . . 72
4.3.1 Choice of the best set of parameters . . . 72
4.3.2 Charge calculation with EEM and SQE models . . . 74
4.3.3 Electrostatic potential . . . 81
4.4 Conclusions . . . 87

5 Computation of dispersion energy in oxides 89
5.1 Introduction. . . . 89
5.2 Theory. . . . 91
5.2.1 Maximally localized Wannier functions. . . . 91
5.2.2 Dispersion coefficient and MLWFs. . . . 92
5.2.3 Localization of MLWFs and character of chemical bond in system. . . . 93
5.2.4 Effective atomic orbitals. . . . 94
5.3 Calculations. . . . 95
5.4 Results and discussion. . . . 96
5.4.1 Effective atomic orbitals method and combination rules. . . . 96
5.4.2 Atomic dispersion coefficients. . . . 98
5.4.3 Implication for energy and forces. . . . 106
5.5 Conclusions. . . . 107

6 Short-range potentials by force-matching procedure. 109
6.1 Introduction. . . . 109
6.2 Minimization criteria. . . . 111
6.3 Creating the database of reference values. . . . 112
6.4 Effective interatomic potentials by cubic splines. . . . 114
6.5 Force-matching approach. . . . 120
6.5.1 Ansatz for short-range interaction potential. . . . 121
6.5.2 Basic expressions for derivatives. . . . 124
6.5.3 Parameters of the short-range potential. . . . 126

7 Polarizable force field models for silica. 129
7.1 Introduction. . . . 129
7.2 Computations. . . . 131
7.3 Results and Discussion. . . . 133
7.3.1 *exp-4* potential. . . . 133

7.3.2 *Morse-SR* potential. 135
7.4 Conclusions. 141

8 Conclusions and Perspectives. 143

Appendices. 147

A Electrostatic interactions for distributed charges and dipoles. 149

B Hardness matrix elements of the CPE and SQE models for periodic systems. 151

C Static and response cost functions. 153

D Details on structural models. 157

E Forces and Hessian matrix in CPE models. 159

F Short-range forces and virial for the environment-dependent ansatz. 161

Chapter 1

Introduction.

Nowadays molecular modeling has become an indispensable tool in such fields as physical chemistry, solid state physics, materials science, geochemistry, molecular biology, pharmaceutical research, *etc.* In the broadest sense, molecular modeling can be defined as the use of computers and numerical algorithms to obtain the information about properties and processes in the matter at the atomic level. With the development of computer hardware, computational algorithms, and of theoretical models, this research activity has rapidly evolved over the past 50 years into a domain that successfully complements the traditional areas of scientific studies, such as experiment and theory. The major advantage of molecular modeling is that it is capable of providing exact results for complex models of physico-chemical phenomena and thus, to be of great help not only for the interpretation of experimental data, but also for validation of theoretical models.

A key quantity in molecular modeling is the energy that is intimately related to the geometrical arrangement of atoms constituting the system in the space, *i.e.* the structure. The finding of the relationship between these two characteristics, as well as between the response of the structure and the energy to a perturbation, provides access to the understanding of the behaviour of the systems at the macroscopic level.

Models underlying molecular modeling and related methods can be classified according to the approach they use to compute the energy and related characteristics for a given structure. The most sophisticated and reliable approaches use the quantum-mechanical ansatz, *i.e.* these methods attempt to solve, possibly without any simplifications, the Schrödinger equation for the nuclei and electrons constituting the system under study. Of course, for the systems of practical interest the task is insurmountable and some approximations are inevitably to be done, *e.g.* Born-Oppenheimer (adiabatic) approximation separating the degrees of freedom of nuclei from those of electrons. Methods based on the quantum mechanics of many-electron systems, which form the family of quantum chemistry methods, yield a wealth of information about the electronic structure, from which many of other important characteristics can be derived. An important advantage of quantum-chemical calculations is that the only information, which needs to be given as input, are the positions of atoms, their nuclear charge, and the charge state of the system. The

price to pay for the quality of obtained results is a large computational expenses related to the calculations. Consequently, the models and methods of quantum chemistry, especially at a high level, commonly deal with systems whose size does not exceed tens to hundred of atoms.

The second way of computing the energy-structure relationship is the use of a set of effective potential functions for describing the interatomic interactions. Such an approach greatly reduces the computational burden, thus allowing systems of thousands (hundreds of thousands) of atoms to be studied. The set of potential energy functions and their parameters is often called force field, this term is widely employed in the literature on molecular modeling. In contrast to the quantum-chemical models, the models of effective potentials rely on parameters that need to be determined and the transferability of the parameters between different systems has to be carefully tested. The force field approach introduces significant simplifications, as compared to the quantum-chemical models. Disadvantages of force fields lie in their parametric nature and the absence of a universal functional form of interatomic potentials. In fact, the analytical form of effective potential functions depends on both the nature of chemical bonding in the system and the properties of interest in a given study. Thus, the vibrational dynamics can well be described with a simple harmonic force field, while such a model fails to simulate phase transitions upon a variation of temperature and/or pressure.

In spite of spectacular progress of the computer hardware and computational algorithms in the last two decades*, one can put the question: *"do we still need force fields?"*. The answer to this question *"yes, we do"* has many facets. Besides the possibility of studying large systems at a modest computation cost, force fields have several attractive features. Thus, a properly constructed force field provides a physically meaningful model of interatomic interactions allowing to condense the intricate many-electron wavefunction and complex many-center integrals of quantum-mechanical model into a (relatively) small number of parameters. Secondly, as any parametric model, force fields can be used for "playing", *i.e.* to answer the question *"what if?"* and are, therefore, capable of studying hypothetical structures by varying geometrical and/or force field parameters.

As it was mentioned above, force fields use parametrized functions to represent interatomic interactions in a system. The calibration of potential functions remains a challenging task, especially if one demands the parameters to be transferable in a wide range of systems with different structures and to have a physical meaning. Generally, once being fitted the parameters remain fixed from system to system, although the environment of atoms in the systems may be different. One of the promising routes of improving the transferability of force fields and of increasing their predictive power is the development of so-called polarizable force fields. The term "polarizable" signifies that parameters of potential functions depend on the geometry, chemical environment of atoms, and/or on an external perturbation applied to the system. In a strict sense, only pa-

*At the end 1980s, a geometry optimization of a system of ten heavy atoms at HF/6-31G level necessitated the use of mainframe computers, while nowadays this task is easily done with a modest laptop (why not with a smartphone).

rameters of potentials describing electrostatic interactions are changeable, whereas in the largest sense, parameters related to other types of interactions can vary as well in the response to the environmental changes. Of course, the parameter variations cannot be random and should rely on a model based on physical ground. Establishing the relationships between different parameters can therefore shed a new light on the linkages existing between different types interatomic interactions, that could previously be hidden in the average values of parameters of nonpolarizable force fields.

The subject of the work presented in the thesis is the development of a polarizable force field model in a broad sense of this term. Since the quality of force field crucially depends on the parametrization and as the quantity of experimental data is often limited, the work uses purely computational way of the force field development. That is high-quality quantum-chemical calculations are employed to generate a database of "observables" and these observables are then used in the design of effective potential functions and in the calibration of force field parameters. It is clear that such an approach contain errors intrinsic to the quantum-chemical methods used in the database generation. The advantage is, however, that being tested for the general validity of underlying interaction model, the force field can easily be improved with the use of results of quantum-chemical calculations of higher level.

The systems of interest in the present work are semi-ionic solids, more specifically oxide materials. These systems are of significant interest from both the fundamental viewpoint and their industrial applications. Thus, silica and silicates are the most abundant compounds in the Earth's mantle. Complex oxides, *e.g.* ceramic materials, find numerous applications in such areas as catalysis, electronic devices, new functional materials, *etc.* A model system in the work is silicon dioxide, SiO_2. The choice of silica as the test system is due to different reasons. First, silica has a number of crystalline polymorphs and also exists in amorphous phase. This feature simplifies the test of transferability of developed force field, which ideally should be able to describe the complete phase diagram. Second, a wealth of experimental data is available for silica structures that allows validation of the model. Third, silica is, perhaps, one of the best studied solids from the modelling viewpoint. There exist numerous potential models for silica that allows an easy comparison of computational results with the data of other modelling studies, thus revealing all *pros* and *cons* of the model.

The manuscript is structured as follows. The next chapter (Chapter 2) presents models and computational methods used in the work. The chapter discusses not only the models and methods of molecular modeling, but also some of the methods used in the data fitting. Chapter 3 provides an overview of force field models and then discusses the ways of how the polarization effects can be incorporated in the force fields, such polarizable point dipoles model, shell model, *etc.* The final part of the chapter states the problems the thesis aimed to answer at. It also presents the main lines of the developed force field. The following chapter (Chapter 4) describes the model used to

deal with the long-range electrostatic interactions. The model is based on the chemical potential equalization approach with both atomic and bond resolution. Factors influencing the transferability of the model parameters are discussed and the best scheme is chosen for the subsequent use in the complete force field model. Chapter 5 deals with the calculation of dispersion interactions and makes use of one of the approaches that have recently emerged from the density functional theory. A simplified model that can be incorporated to a force field is presented. In the following chapter (Chapter 6) the choice of short-range interaction potentials and their parametrization is discussed. Results of the developed force field to the model systems are reported in Chapter 7. The last chapter of the thesis summarizes results of the work and envisages directions of future studies. Lengthy derivation of equations and auxiliary formulas are given in Appendix. The bibliographic references to the chapters are collected after the Appendices. The list of figures and the list of tables are presented at the end of the manuscript.

Chapter 2

Models and Methods.

The quality of a force field depends not only on the underlying model of interatomic interactions, but also on the parameters of the potential functions. These parameters are generally obtained in a fitting procedure in which some characteristics of the system are reproduced with the chosen set of effective potentials as closely as possible. The construction of a force field therefore needs a database of characteristics that will be used in the fitting procedure. Characteristics commonly used for this purpose are the structural parameters, vibrational spectrum, elastic constants, and bulk modulus for solids, energy differences between structural conformations, etc. However, the quantity of such experimental information is often limited that causes the fitting procedure to be an ill-conditioned as the number of parameters to be found exceeds the number of observables. Furthermore, the reference data are subjected to some (often unknown or uncontrollable) errors.

The development of computer hardware assisted by new numerical algorithms allows nowadays the construction of databases of reference quantities using purely *in silico* approach. This way has the advantage that both the quality and quantity of the reference data are by no means limited by the experimental setup and can systematically be improved by using more sophisticated model in the calculations. This is just the way which was used in the present work for the development of a polarizable force field model. The subsequent sections of this chapter briefly discuss the models and computational methods that were employed to obtain the databases of reference characteristics for the parameterization and validation of the force field. The last part of the chapter is a succinct description of the numerical methods used for the optimization of parameters of the effective potential functions. Some additional information on details of the calculations can be found in the chapters dealing with the results of the work (Chapter 4 to Chapter 7).

2.1 Quantum-chemical models.

Models employing the quantum-mechanical description of interatomic interactions can be classified into two main groups. The division is based on the way the models use to solve the Schrödinger equation for a many-electron system. Those models that employ a many-electron wavefunction as a basic quantity form the family of wavefunction models. The most known member of this family

is the Hartree-Fock model. The second group of models is based on the density functional theory (DFT) and, as it follows from the name, these models use the electron density as a basic variable.

Further classification of the QC models (and corresponding methods) is related to the simplifications introduced by the models to the description of interactions in the system. Certain models neglect some less important interactions and employ (semi-) empirical parameters to describe others. These models, e.g. tight-binding model, are referred to as semi-empirical. Other models either do not introduce any simplifications at all or use a minimal set of parameters. Such models and methods relying on first principles are often named *ab initio*. The results described in the thesis were obtained using methods employing *ab initio* models. The content of this section briefly presents the underlying theory. Atomic units are used throughout the presentation, unless otherwise is explicitly stated.

2.1.1 Schrödinger equation.

The evolution of a non-relativistic quantum-mechanical system consisting of nuclei and electrons is described by the time-dependent Schrödinger equation

$$i\frac{\partial}{\partial t}\Psi(\{\mathbf{r}_i\},\{\mathbf{R}_I\},t) = \hat{H}(t)\Psi(\{\mathbf{r}_i\},\{\mathbf{R}_I\},t), \tag{2.1.1}$$

where $\Psi(\{\mathbf{r}_i\},\{\mathbf{R}_I\},t)$ is the wave-function of the system, $\hat{H}(t)$ denotes the Hamiltonian operator, and $\{\mathbf{r}_i\}$ and $\{\mathbf{R}_I\}$ stand for the coordinates of electrons and nuclei, respectively. It is assumed that the electronic coordinates $\{\mathbf{r}_i\}$ include both the spatial and spin components.

If the Hamiltonian does not depend on time, the time-dependent equation (2.1.1) can be reduced to the time-independent Schrödinger equation, which describes the stationary states of the system

$$\hat{H}\Psi(\{\mathbf{r}_i\},\{\mathbf{R}_I\}) = E\Psi(\{\mathbf{r}_i\},\{\mathbf{R}_I\}). \tag{2.1.2}$$

The total Hamiltonian of the system (2.1.2) is given as the sum of kinetic and Coulombic contributions

$$\hat{H} = -\sum_{i=1}^{n}\frac{1}{2}\boldsymbol{\nabla}_i^2 - \sum_{I=1}^{N}\frac{1}{2M_I}\boldsymbol{\nabla}_I^2 + \sum_{i=1}^{n-1}\sum_{j>i}^{n}\frac{1}{|\mathbf{r}_i-\mathbf{r}_j|} - \sum_{i=1}^{n}\sum_{I=1}^{N}\frac{Z_I}{|\mathbf{r}_i-\mathbf{R}_I|} + \sum_{I=1}^{N-1}\sum_{J>I}^{N}\frac{Z_IZ_J}{|\mathbf{R}_I-\mathbf{R}_J|}, \tag{2.1.3}$$

where $\mathbf{r}_i$ is the coordinate of the electron i and M_I, Z_I, and $\mathbf{R}_I$ are the mass, charge, and the coordinate of the nucleus I, respectively. One sees that the Hamiltonian (2.1.3) depends on both the coordinates of nuclei and electrons that complicates the solving of equation (2.1.2). The simplification of the problem relies upon an approximation that separates electronic and nuclear degrees of freedom and is known as adiabatic or Born-Oppenheimer approximation.

2.1.2 Adiabatic approximation.

The electrostatic forces on both electrons and nuclei are of the same magnitude. If one takes into account the fact that the masses of nuclei $M_I \gg 1$, then the velocities of nuclei are much smaller than those of electrons, and thus, the kinetic energy of nuclei T_N can be viewed as a small perturbation to the electronic Hamiltonian, and the positions of nuclei $\mathbf{R}_I$ can be seen as slowly changing parameters. In other words, in the time-scale of electron dynamics the nuclei can be considered as fixed such that the electrons will rapidly relax to their instantaneous ground state configuration at a given configuration of nuclei. Such a separation of electronic and nuclear degrees of freedom, known as adiabatic approximation, has been originally proposed by M. Born and R. Oppenheimer [1]. It is also referred to as Born-Oppenheimer approximation.

According to the adiabatic approximation, we will state that the total wavefunction of the system can be represented as

$$\Psi(\{\mathbf{r}_i\},\{\mathbf{R}_I\}) = \psi(\{\mathbf{r}_i\}:\{\mathbf{R}_I\})\,\phi(\{\mathbf{R}_I\}), \tag{2.1.4}$$

where $\psi(\{\mathbf{r}_i\}:\{\mathbf{R}_I\})$ is the electronic wavefunction for a given configuration of nuclei $\{\mathbf{R}_I\}$ and $\phi(\{\mathbf{R}_I\})$ is the nuclear wavefunction. The electronic wavefunction is to be the solution of the following Schrödinger equation

$$\hat{H}_e\,\psi(\{\mathbf{r}_i\}:\{\mathbf{R}_I\}) = \varepsilon_e(\{\mathbf{R}_I\})\,\psi(\{\mathbf{r}_i\}:\{\mathbf{R}_I\}), \tag{2.1.5}$$

where $\hat{H}_e$ is the electronic Hamiltonian defined as a sum of the kinetic energy operator for electrons $\hat{T}_e$ and of the potential energy operators of electron-electron $\hat{U}_{ee}$ and electrons-nuclei $\hat{U}_{eN}$ interactions.

$$\hat{H}_e = \hat{T}_e + \hat{U}_{ee} + \hat{U}_{eN} = -\sum_{i=1}^{n}\frac{1}{2}\boldsymbol{\nabla}_i^2 + \sum_{i=1}^{n}\sum_{j>i}^{n}\frac{1}{|\mathbf{r}_i-\mathbf{r}_j|} - \sum_{i=1}^{n}\sum_{I=1}^{N}\frac{Z_I}{|\mathbf{r}_i-\mathbf{R}_I|}. \tag{2.1.6}$$

Energy $\varepsilon_e(\{\mathbf{R}_I\})$ is called adiabatic contribution of electrons to the energy of system and its dependence on the nuclear configuration can be obtained by solving eq. (2.1.5) for a set of configurations of nuclei.

Applying the full Hamiltonian (2.1.3) to the wavefunction (2.1.4) one obtains

$$\begin{aligned}\hat{H}\Psi(\{\mathbf{r}_i\},\{\mathbf{R}_I\}) &= \psi(\{\mathbf{r}_i\}:\{\mathbf{R}_I\})\Big[\hat{H}_N + \varepsilon_e(\{\mathbf{R}_I\})\Big]\,\phi(\{\mathbf{R}_I\}) \\ &- \sum_{J=1}^{N}\frac{1}{2M_J}\Big[2\boldsymbol{\nabla}_J\phi(\{\mathbf{R}_I\})\cdot\boldsymbol{\nabla}_J\psi(\{\mathbf{r}_i\}:\{\mathbf{R}_I\}) \\ &+ \phi(\{\mathbf{R}_I\})\boldsymbol{\nabla}_J^2\psi(\{\mathbf{r}_i\}:\{\mathbf{R}_I\})\Big],\end{aligned} \tag{2.1.7}$$

where $\hat{H}_N$ denotes the nuclear Hamiltonian

$$\hat{H}_N = \hat{T}_N + \hat{U}_{NN} = -\sum_{I=1}^{N} \frac{1}{2M_I} \boldsymbol{\nabla}_I^2 + \sum_{I=1}^{N} \sum_{J>I}^{N} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|}. \tag{2.1.8}$$

The second term in eq. (2.1.7) is a non-adiabatic term. Because $M_J \gg 1$ and $\boldsymbol{\nabla}_J \phi(\{\mathbf{R}_I\}) \gg \boldsymbol{\nabla}_J \psi(\{\mathbf{r}_i\} : \{\mathbf{R}_I\})$ (wavefunction of heavy nuclei rises much more steeply than that of light electrons), the non-adiabatic term can be neglected in most cases. Therefore, the nuclear wavefunction satisfies the following equation

$$\left[\hat{H}_N + \varepsilon_e(\{\mathbf{R}_I\})\right] \phi(\{\mathbf{R}_I\}) = \varepsilon_n \phi(\{\mathbf{R}_I\}), \tag{2.1.9}$$

that corresponds to the separation of electronic and nuclear degrees of freedom. From the viewpoint of solving the time-independent Schrödinger equation (2.1.2) this means that one can first solve the Schrödinger equation for electrons (2.1.5) for a set of "clamped" nuclei to obtain the potential energy surface $\varepsilon_e(\{\mathbf{R}_I\})$, and then solve (2.1.9) to find the wavefunction of the nuclei.

2.1.3 Solving the Schrödinger equation.

As it follows from the above presentation, the key-point of quantum-chemical description of interatomic interactions is the solving of the electronic Schrödinger equation (2.1.5). Two models have been developed for this purpose that are briefly presented below: the Hartree-Fock model and Density Functional Theory.

Hartree-Fock model.

The Hartree-Fock (HF) model (and method) is based on the idea of D. R. Hartree and has been later extended by V. Fock [2]. The HF model belongs to the family of wavefunction models and uses effective one-particle wavefunctions.

The Pauli principle demands that the wavefunction for a system of identical particles with half-integer spin is to be antisymmetric with respect to permutation of any two particles

$$\psi(\mathbf{r}_1, \ldots, \mathbf{r}_i, \ldots, \mathbf{r}_j, \ldots, \mathbf{r}_n) = -\psi(\mathbf{r}_1, \ldots, \mathbf{r}_j, \ldots, \mathbf{r}_i, \ldots, \mathbf{r}_n). \tag{2.1.10}$$

To satisfy the condition (2.1.10) the many-electron wavefunction is constructed in the form of determinant (Slater determinant):

$$\psi(\{\mathbf{r}_i\}) = \frac{1}{\sqrt{n!}} \begin{vmatrix} \chi_1(\mathbf{r}_1) & \chi_1(\mathbf{r}_2) & \ldots & \chi_1(\mathbf{r}_n) \\ \chi_2(\mathbf{r}_1) & \chi_2(\mathbf{r}_2) & \ldots & \chi_2(\mathbf{r}_n) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_n(\mathbf{r}_1) & \chi_n(\mathbf{r}_2) & \ldots & \chi_n(\mathbf{r}_n) \end{vmatrix}, \tag{2.1.11}$$

where $\chi_i(\mathbf{r})$ are one-electron orthonormal wavefunctions (spin orbitals); the wavefunctions $\chi_i(\mathbf{r})$

should minimize the electronic energy.

The energy of the system ε_e can be obtained as

$$\varepsilon_e = \langle\psi|\hat{H}_e|\psi\rangle, \tag{2.1.12}$$

where the electronic Hamitonian $\hat{H}_e$ is given by (2.1.6). Making use of the many-electron wavefunction in the form of Slater determinant (2.1.11), the energy ε_e can be recast as

$$\varepsilon_e = \sum_{i=1}^{n} H_{ii} + \frac{1}{2}\sum_{i,j=1}^{n}\Big(J_{ij} - K_{ij}\Big) \tag{2.1.13}$$

with the H_{ii}, J_{ij}, and K_{ij} contributions defined as

$$H_{ii} = \int d\mathbf{r}\,\chi_i^*(\mathbf{r})\Big(-\frac{1}{2}\boldsymbol{\nabla}^2 - \sum_{I=1}^{N}\frac{Z_I}{|\mathbf{r}-\mathbf{R}_I|}\Big)\chi_i^*(\mathbf{r}), \tag{2.1.14}$$

$$J_{ij} = \int d\mathbf{r}_1 d\mathbf{r}_2\,\chi_i^*(\mathbf{r}_1)\chi_i(\mathbf{r}_1)\frac{1}{|\mathbf{r}_1-\mathbf{r}_2|}\chi_j(\mathbf{r}_2)\chi_j^*(\mathbf{r}_2), \tag{2.1.15}$$

$$K_{ij} = \int d\mathbf{r}_1 d\mathbf{r}_2\,\chi_i^*(\mathbf{r}_1)\chi_j(\mathbf{r}_1)\frac{1}{|\mathbf{r}_1-\mathbf{r}_2|}\chi_j^*(\mathbf{r}_2)\chi_i(\mathbf{r}_2). \tag{2.1.16}$$

These integrals H_{ii}, J_{ij}, and K_{ij} are called core integral, Coulomb integral, and exchange integral, respectively. The latter does not have a classical counterpart and is a direct consequence of the Pauli principle (2.1.10).

The energy (2.1.12) must be minimized with respect to the one-electron wavefunctions $\chi_i(\mathbf{r})$ under the constraint of orthonormality $\int d\mathbf{r}\,\chi_i^*(\mathbf{r})\chi_j(\mathbf{r}) = \delta_{ij}$. The use of the variational principle results in a set of equations

$$\hat{f}_i\,\chi_i(\mathbf{r}) = \sum_{j=1}^{n}\lambda_{ij}\chi_j(\mathbf{r}),\ i = 1, 2, \ldots, n, \tag{2.1.17}$$

where the operator $\hat{f}_i$ is defined as

$$\hat{f}_i = \hat{h}_i + \sum_{j=1}^{n}\Big(\hat{j}_j - \hat{k}_j\Big) \tag{2.1.18}$$

with the Hamitonian core operator

$$\hat{h}_i = -\frac{1}{2}\boldsymbol{\nabla}_i^2 - \sum_{I=1}^{N}\frac{Z_I}{|\mathbf{r}_i-\mathbf{R}_I|}, \tag{2.1.19}$$

Coulomb operator $\hat{j}_j$

$$\hat{j}_j\chi_i(\mathbf{r}) = \left[\int \chi_j^*(\mathbf{r}')\frac{1}{|\mathbf{r}-\mathbf{r}'|}\chi_j(\mathbf{r}')d\mathbf{r}'\right]\chi_i(\mathbf{r}), \tag{2.1.20}$$

and exchange operator $\hat{k}_j$

$$\hat{k}_j \chi_i(\mathbf{r}) = \left[\int \chi_j^*(\mathbf{r}') \frac{1}{|\mathbf{r} - \mathbf{r}'|} \chi_i(\mathbf{r}') d\mathbf{r}' \right] \chi_j(\mathbf{r}). \tag{2.1.21}$$

One sees that the Fock operator (2.1.18) is an effective one-electron Hamiltonian operator of a many-electron system.

Eq. (2.1.17) can further be simplified by applying an unitary transformation to it in the space of occupied orbitals. Then, the Hartree-Fock equations (2.1.17) take their canonical form

$$\hat{f}\, \chi_i(\mathbf{r}) = \varepsilon_i \chi_i(\mathbf{r}),\ i = 1, 2, \ldots, n. \tag{2.1.22}$$

Note that eqs. (2.1.22) represent a standard eigenvalue problem and thus, all one-electron wavefunctions $\chi_i(\mathbf{r})$ are orthogonal to each other. The eigenvalues ε_i contain the kinetic energy of the i-th electron, its energy in an external field due to the nuclei, and the interaction energy with all other electrons.

The system of Hartree-Fock equations (2.1.22) is nonlinear with respect to the one-electron wavefunctions because the Fock operator itself (2.1.18) depends on the wavefunctions $\chi_i(\mathbf{r})$. Consequently, the solution of (2.1.22) should proceed via an iterative procedure, i.e. making use of a set of known one-electron wavefunctions $\chi_i^{(m)}(\mathbf{r}),\ i = 1, 2, \ldots, n$ at the iteration m, one constructs the Fock operator $\hat{f}^{(m)}$. The eigenfunctions $\chi_i^{(m+1)}(\mathbf{r}),\ i = 1, 2, \ldots, n$ of this operator obtained as the solution of equations

$$\hat{f}^{(m)}\, \chi_i^{(m+1)}(\mathbf{r}) = \varepsilon_i^{(m+1)} \chi_i^{(m+1)}(\mathbf{r}),\ i = 1, 2, \ldots, n, \tag{2.1.23}$$

are then used in the next iteration until a convergence criterion is satisfied. This iterative procedure is called self-consistent field (SCF) method.

Density functional theory.

The basic idea of the DFT is to replace the many-electron wavefunction with the electron density, that reduces the number of spatial coordinates from $3n$ to 3. The electron density can be directly constructed from the corresponding many-body wavefunction, if the latter is known, as

$$\rho(\mathbf{r}) = \int \ldots \int d\mathbf{r}_2 \ldots d\mathbf{r}_n \psi^*(\mathbf{r}, \mathbf{r}_2, \ldots, \mathbf{r}_n) \psi(\mathbf{r}, \mathbf{r}_2, \ldots, \mathbf{r}_n). \tag{2.1.24}$$

The modern formulation of DFT is due to P. Hohenberg and W. Kohn [3]. They have shown that for many-electron system in its ground state the electron density can be considered as a basic variable, which uniquely defines all characteristics of the system.

Hohenberg-Kohn theorems. Let us consider the total electronic Hamiltonian (2.1.6) that can be rewritten in the form

$$\hat{H}_e = \hat{F} + \hat{V}_{\text{ext}}, \tag{2.1.25}$$

where the operator $\hat{F}$ is defined as

$$\hat{F} = -\sum_{i=1}^{n} \frac{1}{2}\boldsymbol{\nabla}_i^2 + \sum_{i=1}^{n}\sum_{j>i}^{n} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}, \tag{2.1.26}$$

and $\hat{V}_{\text{ext}}$ stands for

$$\hat{V}_{\text{ext}} = \sum_{i=1}^{n} V_{\text{ext}}(\mathbf{r}_i) \tag{2.1.27}$$

with $V_{\text{ext}}(\mathbf{r})$ being the Coulomb potential due to the nuclei

$$V_{\text{ext}}(\mathbf{r}) = -\sum_{I=1}^{N} \frac{Z_I}{|\mathbf{r} - \mathbf{R}_I|}. \tag{2.1.28}$$

It is clear that the functional $\hat{F}$ in eq. (2.1.25) has the same form for any system of n electrons. Hence, the total electronic Hamiltonian and thus, the ground state wavefunction $|\psi_0\rangle$ are fully determined by n and $V_{\text{ext}}(\mathbf{r})$. Using eq. (2.1.24) the ground state $|\psi_0\rangle$ yields the ground state density $\rho_0(\mathbf{r})$. Consequently, the ground state $|\psi_0\rangle$ and density $\rho_0(\mathbf{r})$ are both the functionals of the number of electrons n and of the external potential $V_{\text{ext}}(\mathbf{r})$.

P. Hohenberg and W. Kohn [3] introduced two remarkable statements, which are also known as Hohenberg-Kohn theorems.

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

   The proof can be given by *reductio ad absurdum*. Assume that there exists another external potential $V'_{\text{ext}}(\mathbf{r})$ with ground state $|\psi'_0\rangle$, and it corresponds to the same density $\rho_0(\mathbf{r})$. The ground state energies are $E_0 = \langle\psi_0|\hat{H}|\psi_0\rangle$ and $E'_0 = \langle\psi'_0|\hat{H}'|\psi'_0\rangle$, where the Hamiltonians are $\hat{H} = \hat{F} + \hat{V}_{\text{ext}}$ and $\hat{H}' = \hat{F} + \hat{V}'_{\text{ext}}$.

   By taking $|\psi'_0\rangle$ as a trial wavefunction for $\hat{H}$ one obtains

$$\begin{aligned} E_0 < \langle\psi'_0|\hat{H}|\psi'_0\rangle &= \langle\psi'_0|\hat{H}'|\psi'_0\rangle + \langle\psi'_0|(\hat{H} - \hat{H}')|\psi'_0\rangle \\ &= E'_0 + \int d\mathbf{r}\rho_0(\mathbf{r})\{V_{\text{ext}}(\mathbf{r}) - V'_{\text{ext}}(\mathbf{r})\}, \end{aligned} \tag{2.1.29}$$

   while using $|\psi_0\rangle$ as a trial wavefunction for $\hat{H}'$ gives

$$\begin{aligned} E'_0 < \langle\psi_0|\hat{H}'|\psi_0\rangle &= \langle\psi_0|\hat{H}|\psi_0\rangle + \langle\psi_0|(\hat{H}' - \hat{H})|\psi_0\rangle \\ &= E_0 + \int d\mathbf{r}\rho_0(\mathbf{r})\{V'_{\text{ext}}(\mathbf{r}) - V_{\text{ext}}(\mathbf{r})\}. \end{aligned} \tag{2.1.30}$$

By taking the sum of the two above equations one obtains

$$E_0 + E_0' < E_0 + E_0' \tag{2.1.31}$$

that is obviously wrong and therefore, the $V_{\text{ext}}(\mathbf{r})$ for the given density is unique.

2. *A universal functional for energy $E_V[\rho]$ can be defined, that is valid for any external potential $V(\mathbf{r})$. For all V-representable densities $\rho(\mathbf{r})$ $E_V[\rho] \geq E_0$, where E_0 is the ground-state energy for n electrons in the external potential $V(\mathbf{r})$.*

   For any density which is the ground state for some external potential (V-representability) the functional $F[\rho] = \langle\psi_0|\hat{F}|\psi_0\rangle$ is unique. Therefore, a functional for an arbitrary external potential $V(\mathbf{r})$, determined by $\rho(\mathbf{r})$, can be defined as

$$E_V[\rho] = F[\rho] + \int d\mathbf{r} V(\mathbf{r})\rho(\mathbf{r}). \tag{2.1.32}$$

   By virtue of the first theorem, the given density $\rho(\mathbf{r})$ determines its external potential $V_{\text{ext}}(\mathbf{r})$ and the ground state $|\psi\rangle$. If this ground state is then used as a trial wavefunction for the Hamiltonian with external potential $V(\mathbf{r})$, one obtains

$$\langle\psi|\hat{H}|\psi\rangle = \langle\psi|\hat{F}|\psi\rangle + \langle\psi|\hat{V}|\psi\rangle = F[\rho] + \int d\mathbf{r} V(\mathbf{r})\rho(\mathbf{r}) = E_V[\rho] \geq E_0 \tag{2.1.33}$$

   by the variational principle.

Thus, the problem of solving the Schrödinger equation for a non-degenerate ground state can be revised as a minimization of the functional $E_V[\rho]$ with respect to all V-representable densities $\rho(\mathbf{r})$. It should be underlined that the functional determines only the ground state, and gives no guidance concerning excited states. It is also important to note that the functional $\hat{F}$ (2.1.26) is defined independently on the external potential and therefore, $\hat{F}$ is a universal functional of $\rho(\mathbf{r})$.

The Kohn-Sham equations. The Hohenberg-Kohn theorems are the existence theorems and they provide no functional forms or solutions. In order to make the problem tractable, W. Kohn and L. J. Sham [4] have proposed to replace the many-body interacting system with an auxiliary *noninteracting* system. Such a trick leads to the independent-particle equations, that can be solved exactly. The two key assumptions of the Kohn-Sham model are

- The exact ground state density $\rho_0(\mathbf{r})$ of a system can be represented by the ground state density of some auxiliary noninteracting system. This is known as "*noninteracting-V-representability*", i.e. $E_{\text{KS}}[\rho_0] = E[\rho_0]$, where E_{KS} is the energy of the auxiliary system.
- The auxiliary Hamiltonian operator is given as a sum of kinetic operator of noninteracting system and an auxiliary potential $v_{\text{eff}}(\mathbf{r})$ acting on an electron at position $\mathbf{r}$.

The one-electron orbitals of the system of noninteracting electrons (Kohn-Sham orbitals) are then found as a solution of one-electron Schrödinger equation

$$\big(-\frac{1}{2}\boldsymbol{\nabla}_i^2 + v_{\text{eff}}(\mathbf{r})\big)\chi_i(\mathbf{r}) = \varepsilon_i\chi_i(\mathbf{r}) \tag{2.1.34}$$

that yields the density

$$\rho(\mathbf{r}) = \sum_{i=1}^{n} |\chi_i(\mathbf{r})|^2. \tag{2.1.35}$$

The energy functional $E[\rho]$ (2.1.26) of the original interacting system is given by the sum (2.1.32)

$$E[\rho] = T[\rho] + V_{\text{ee}}[\rho] + V_{\text{Ne}}[\rho], \tag{2.1.36}$$

where the first two terms define the universal functional $F[\rho]$ (2.1.26). Kohn and Sham proposed to rewrite (2.1.36) in the form

$$E[\rho] = T_{\text{a}}[\rho] + J[\rho] + V_{\text{Ne}}[\rho] + \Big(T[\rho] - T_{\text{a}}[\rho] + V_{\text{ee}}^{\text{NC}}[\rho]\Big), \tag{2.1.37}$$

and defined an exchange-correlation functional $E_{\text{xc}}[\rho]$ as

$$E_{\text{xc}}[\rho] = T[\rho] - T_{\text{a}}[\rho] + V_{\text{ee}}^{\text{NC}}[\rho]. \tag{2.1.38}$$

In (2.1.37) and (2.1.38) $T_{\text{a}}[\rho]$ is the kinetic energy of the auxiliary system of noninteracting electrons and $V_{\text{ee}}[\rho] = J[\rho] + V_{\text{ee}}^{\text{NC}}[\rho]$ with $J[\rho]$ being the Coulomb electron-electron energy and $V_{\text{ee}}^{\text{NC}}[\rho]$ including all nonclassical contributions to the energy $V_{\text{ee}}[\rho]$ due to exchange and correlation interactions.

Let us now define $v_{\text{eff}}(\mathbf{r})$ in (2.1.34) as

$$v_{\text{eff}}(\mathbf{r}) = \frac{\delta J[\rho]}{\delta\rho(\mathbf{r})} + \frac{\delta V_{\text{Ne}}[\rho]}{\delta\rho(\mathbf{r})} + \frac{\delta E_{\text{xc}}[\rho]}{\delta\rho(\mathbf{r})} = \int d\mathbf{r}' \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + \sum_{I=1}^{N} \frac{Z_I}{|\mathbf{r}-\mathbf{R}_I|} + v_{\text{xc}}(\mathbf{r}). \tag{2.1.39}$$

Therefore, by solving n one-electron equations (2.1.34) with $v_{\text{eff}}(\mathbf{r})$ given by (2.1.39) one can find the electron density of the original system of interacting particles as (2.1.35). Note that v_{eff} itself depends on $\rho(\mathbf{r})$ and thus, the (2.1.34) and (2.1.39) should be solved iteratively. Starting with a guessed $\rho(\mathbf{r})$ construct $v_{\text{eff}}(\mathbf{r})$ from (2.1.39), then find new $\rho(\mathbf{r})$ from (2.1.34) and (2.1.35), etc. until the self-consistent solution is found.

As it follows from (2.1.38), exchange-correlation functional $E_{\text{xc}}[\rho]$ includes the difference between the kinetic energies of the interacting and noninteracting systems and the contributions from the exchange and correlation electron-electron energies. The main problem, however, is that the exact form of XC functional is unknown and therefore, approximations have to be employed to construct the functional for the use of the Kohn-Sham model in practical applications.

Local density approximation. The simplest approximation to the exact exchange-correlation functional is the local density approximation (LDA), which was proposed by Kohn and Sham [4]. In LDA the exchange-correlation energy $E_{\rm xc}^{LDA}$ is taken simply as an integral over all space with the exchange-correlation energy density $\varepsilon_{\rm xc}^{\rm hom}$ assumed to be that of the homogeneous electron gas with the same density

$$E_{\rm xc}^{LDA}[\rho] = \int d\mathbf{r}\varepsilon_{\rm xc}(\rho(\mathbf{r}))\rho(\mathbf{r}). \tag{2.1.40}$$

The exchange-correlation potential is then given by

$$\frac{\delta E_{\rm xc}^{LDA}[\rho]}{\delta\rho(\mathbf{r})} = \varepsilon_{\rm xc}(\rho(\mathbf{r})) + \rho(\mathbf{r})\frac{\partial\varepsilon_{\rm xc}(\rho(\mathbf{r}))}{\partial\rho(\mathbf{r})} \tag{2.1.41}$$

and it is common to represent the XC potential as the sum of exchange (x) and correlation (c) contributions

$$\varepsilon_{\rm xc} = \varepsilon_{\rm x} + \varepsilon_{\rm c}. \tag{2.1.42}$$

The exchange part of LDA is given by Dirac functional [5]

$$\varepsilon_{\rm x}(\rho(\mathbf{r})) = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{1/3}\rho(\mathbf{r})^{1/3} \tag{2.1.43}$$

so that the exchange energy functional has the following analytical form

$$E_{\rm x}^{LDA}[\rho] = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{\frac{1}{3}}\int d\mathbf{r}\rho(\mathbf{r})^{4/3}. \tag{2.1.44}$$

The analytical expression for the correlation energy density $\varepsilon_{\rm c}(\rho(\mathbf{r}))$ is unknown except for the high- and low-density limits. For an unpolarized electron gas in the high-density limit the result is [6, 7]

$$\varepsilon_{\rm c}(r_s) \to A\ln(r_s) + B + r_s(C\ln(r_s) + D) + \dots, \tag{2.1.45}$$

whereas in the low-density limit the expression is [8]

$$\varepsilon_{\rm c}(r_s) \to \frac{a_1}{r_s} + \frac{a_2}{r_s^{3/2}} + \frac{a_3}{r_s^2} + \dots. \tag{2.1.46}$$

In eqs. (2.1.45) and (2.1.46) r_s states for Wigner-Seitz radius and it is related to the density ρ as

$$\frac{4\pi}{3}r_s^3 = \rho^{-1}. \tag{2.1.47}$$

The accurate values for the energy density $\varepsilon_{\rm c}(\rho(\mathbf{r}))$ have been determined from quantum Monte Carlo calculations by Ceperley and Alder [9]. The results have then been interpolated by different analytic forms to provide the analytical expression for $\varepsilon_{\rm c}(\rho(\mathbf{r}))$. The best known parameterizations are those by Vosko-Wilk-Nusair(VWN) [10], Perdew-Zunger (PZ81) [11], and by Perdew-Wang (PW92) [12].

Generalized gradient approximation. The local density approximation is expected to perform well in situations, where the density varies slowly, while it causes large errors for systems with rapidly changing density. The next step in the development of an approximate XC functional is to include not only the density value at the given point, but also the density gradient at that point into consideration, i.e.

$$E_{\rm xc} = E_{\rm xc}[\rho(\mathbf{r}), \boldsymbol{\nabla}\rho(\mathbf{r}).] \tag{2.1.48}$$

This approximation is known as generalized gradient approximation (GGA). While there exists only one LDA XC functional, there are many different GGA flavours: semi-empirical and fully from first principles. The GGA XC functionals considered below belong to those based on first principles.

PBE and revPBE functionals. The exchange-correlation energy within GGA is given by

$$E_{\rm xc}^{\rm GGA}[\rho_\uparrow, \rho_\downarrow] = \int d\mathbf{r}\, f(\rho_\uparrow, \rho_\downarrow, \boldsymbol{\nabla}\rho_\uparrow, \boldsymbol{\nabla}\rho_\downarrow) = E_{\rm c}^{\rm GGA}[\rho_\uparrow, \rho_\downarrow] + E_{\rm x}^{\rm GGA}[\rho_\uparrow, \rho_\downarrow], \tag{2.1.49}$$

where $\rho_\uparrow$ and $\rho_\downarrow$ are the spin densities. PBE [13] and revPBE [14] XC functionals have been developed such that they satisfy some universal physical constraints and contain parameters based on fundamental constants.

The correlation part of the functionals is chosen in the form of uniform gas correlation energy plus an additive term

$$E_{\rm c}^{\text{GGA-PBE}}[\rho_\uparrow, \rho_\downarrow] = \int d\mathbf{r}\, \rho(\mathbf{r}) \left[\epsilon_{\rm c}^{\rm hom}(r_s, \zeta) + H(r_s, \zeta, t)\right], \tag{2.1.50}$$

where ζ is the relative spin polarization

$$\zeta = (\rho_\uparrow - \rho_\downarrow)/\rho \quad \text{and} \quad \rho = \rho_\uparrow + \rho_\downarrow \tag{2.1.51}$$

and t is the dimensionless density gradient

$$t = \frac{|\boldsymbol{\nabla}\rho|}{2\phi k_s \rho}. \tag{2.1.52}$$

In eq. (2.1.52) ϕ is a spin scaling factor

$$\phi = \left[(1+\zeta)^{2/3} + (1-\zeta)^{2/3}\right]/2 \tag{2.1.53}$$

and k_s is the Thomas-Fermi screening wavenumber.

The form for the function H is chosen to satisfy three conditions [13]: known functional forms of H at (i) slow ($t \to 0$) and (ii) rapidly ($t \to \infty$) varying limits, and (iii) behaviour of the correlation energy at uniform scaling to the high-density limit ($\rho(\mathbf{r}) \to \lambda^3\rho(\lambda\mathbf{r})$ and $\lambda \to \infty$).

Analytical form of H satisfying these conditions is given by

$$H = \gamma\,\phi^3\,\ln\left\{1+\frac{\beta}{\gamma}t^2\frac{1+At^2}{1+At^2+A^2t^4}\right\}, \tag{2.1.54}$$

where

$$A = \frac{\beta}{\gamma}\left[\exp\left(-\frac{\epsilon_{\rm c}^{\rm hom}}{\gamma\,\phi^3}\right)\right]^{-1}, \quad \gamma = \frac{1-\ln 2}{\pi^2} \quad \text{and} \quad \beta \simeq 0.066725. \tag{2.1.55}$$

The exchange part of PBE XC functional is written as

$$E_{\rm x}^{\text{GGA-PBE}}[\rho_\uparrow,\rho_\downarrow] = \int d\mathbf{r}\,\rho(\mathbf{r})\,\epsilon_{\rm x}^{\rm hom}(\rho)F_x(s), \tag{2.1.56}$$

where $F_{\rm x}$ is the enhancement factor that is chosen to satisfy four other conditions based on physical constraints. The factor has the following analytical form:

$$F_{\rm x}(s) = 1+\kappa-\frac{\kappa}{1+\mu s^2/\kappa}, \tag{2.1.57}$$

where $s = 0.235$ and $\kappa = 0.804$.

Y. Zhang and W. Yang noticed [14] that there remained a flexibility in the choice of the parameter κ value in the exchange part of PBE functional (2.1.57). In the original parameterization the value of $\kappa = 0.804$ was chosen such that the local Lieb-Oxford bound for the exchange potential was satisfied. The authors [14] proposed the value $\kappa = 1.245$ based on the fit of exchange-only atomic energies to exchange-only results from an optimized method. It appeared that this change resulted in a significant improvements over PBE, while calculating atomic total energies and molecule atomization energies. The corresponding XC functional is referred to as revPBE functional.

Hybrid functionals. The basic idea for introduction of hybrid XC functionals is to mix the exact exchange energies from Hartree-Fock method with those obtained from DFT in order to improve the performance. The concept of hybrid functionals was introduced by Becke [15].

Probably the most widely used hybrid functional is B3LYP (Becke, 3-parameter, Lee-Yang-Parr) exchange-correlation functional, which was proposed by Stephens *et al.* [16]. B3LYP is a generalization of the B3P86 functional by Becke [15]. The XC energy in B3LYP is written as

$$E_{\rm xc}^{\rm B3LYP} = E_{\rm xc}^{\rm LDA} + a_0\left(E_{\rm x}^{\rm HF}-E_{\rm x}^{\rm LDA}\right) + a_x\left(E_{\rm x}^{\rm B88}-E_{\rm x}^{\rm LDA}\right) + a_c\left(E_{\rm c}^{\rm LYP}-E_{\rm c}^{\rm VWN}\right), \tag{2.1.58}$$

where B88 and LYP stand for the Becke 88 exchange functional [17] and the correlation functional of Lee, Yand and Parr [18], respectively. The parameters $a_0 = 0.20$, $a_x = 0.72$ and $a_c = 0.81$ are obtained through fitting of thermochemical data in the G1 set [19] of molecules.

Hybrid exchange-correlation functionals give a significant improvement with respect to GGA ones for many molecular properties. However, they are not generally used in solid state calculations because of the difficulty in computing the exact exchange using plane waves. Hybrid functionals

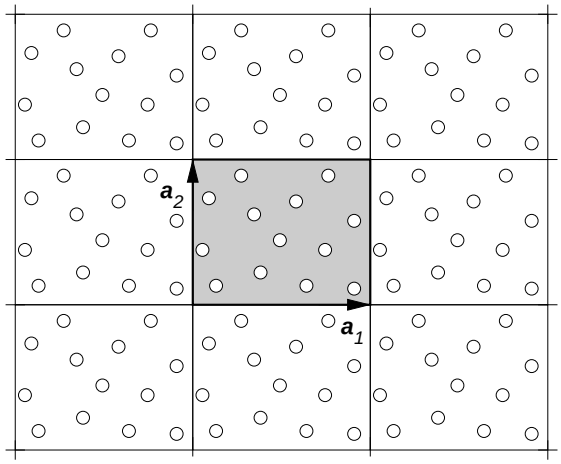

Figure 2.1: Periodic boundary conditions in two-dimensions. The system of particles is enclosed into a cell (shown as shaded rectangle) with the cell parameters $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$. The principal cell is then replicated in the space (plane of the figure) to form an infinite 2D system. The particles in the principal cell interact not only with each other, but also with particles in the images of the cell.

demonstrate the need of including fully non-local information in order to improve the accuracy of calculations.

2.1.4 Dealing with periodic systems.

The models discussed above are straightforward to apply to isolated structures. However, an additional consideration has to be taken into account, while dealing with condensed matter systems. In fact, the number of particles that can be treated in the calculations ranges from few hundreds (with quantum-chemical models) to few tens of thousand (effective potential models). This means that the size of a system is still small as compared to the real world. Therefore, to avoid any finite size effects on the properties of the system, periodic boundary conditions are applied so that the system becomes effectively infinite. The use of periodic boundary conditions is illustrated in Figure 2.1. The system of interest is placed in a cell (often called simulation box) and this cell is replicated in the space in three (or two for 2D systems) dimensions. Both the positions and dynamics of particles (atoms, molecules) in all cells are identical to those in the principal cell. Basically, this means that the system has become periodic with the spatial periodicity determined by the size of simulation box. Two cases should be distinguished. The first one is the case of crystalline systems that are already periodic in space. In this case, the simulation box should contain an integer number of the crystallographic unit cells along each direction. In the second case corresponding to disordered systems (amorphous solids, liquids), the size of cell should be taken large enough such that the artificial periodicity does not affect the characteristics of interest. Nevertheless, regardless the system, the existence of periodicity leads to specific properties of the one-electron wavefunction as discussed below.

Bloch theorem.

For the sake of simplicity let us consider the case of an ideal crystal with the nuclei fixed in a regular periodic array described by a set of Bravais lattice vectors $\{\mathbf{R}\}$. Such an infinite periodic system is invariant with respect to any translation along the lattice vectors and thus the potential $V(\mathbf{r})$ induced by the nuclei is also periodic

$$V(\mathbf{r}+\mathbf{R}) = V(\mathbf{r}). \tag{2.1.59}$$

The Schrödinger equation for a single particle in the potential $V(\mathbf{r})$ reads

$$\hat{H}|\chi\rangle = \Big\{-\frac{1}{2}\boldsymbol{\nabla}^2 + V(\mathbf{r})\Big\}|\chi\rangle = \varepsilon|\chi\rangle \tag{2.1.60}$$

and we are interested in the properties of the one-particle wavefunction $\chi(\mathbf{r})$.

For each vector $\mathbf{R}$

$$\mathbf{R} = n_1\boldsymbol{a}_1 + n_2\boldsymbol{a}_2 + n_3\boldsymbol{a}_3 \tag{2.1.61}$$

with $\boldsymbol{a}_i$ being the primitive lattice vectors and n_i the integers, we define a translation operator $\hat{T}_{\mathbf{R}}$, which acts on an arbitrary function $f(\mathbf{r})$ as follows

$$\hat{T}_{\mathbf{R}} f(\mathbf{r}) = f(\mathbf{r}+\mathbf{R}). \tag{2.1.62}$$

Since the potential $V(\mathbf{r})$ is periodic, the Hamiltonian $\hat{H}$ is periodic too, i.e. $\hat{H}(\mathbf{r}) = \hat{H}(\mathbf{r}+\mathbf{R})$. Therefore, the Hamiltonian operator commutes with the translation operator ($[\hat{H}, \hat{T}_{\mathbf{R}}] = 0$) and

$$\hat{T}_{\mathbf{R}}\hat{H}(\mathbf{r})\chi(\mathbf{r}) = \hat{H}(\mathbf{r}+\mathbf{R})\chi(\mathbf{r}+\mathbf{R}) = \hat{H}(\mathbf{r})\chi(\mathbf{r}+\mathbf{R}) = \hat{H}(\mathbf{r})\hat{T}_{\mathbf{R}}\chi(\mathbf{r}). \tag{2.1.63}$$

It is also clear that the translation operators commute with each other

$$\hat{T}_{\mathbf{R}}\hat{T}_{\mathbf{R}'} = \hat{T}_{\mathbf{R}'}\hat{T}_{\mathbf{R}} = \hat{T}_{\mathbf{R}+\mathbf{R}'}. \tag{2.1.64}$$

Consequently, it should be possible to choose the eigenstates of the Hamiltonian to be simultaneously the eigenstates of all translation operators corresponding to the $\mathbf{R}$ vectors:

$$\hat{H}|\chi\rangle = \varepsilon|\chi\rangle, \tag{2.1.65}$$

$$\hat{T}_{\mathbf{R}}|\chi\rangle = C(\mathbf{R})|\chi\rangle. \tag{2.1.66}$$

From eqs. (2.1.64) and (2.1.66) it follows that

$$C(\mathbf{R}+\mathbf{R}') = C(\mathbf{R})C(\mathbf{R}'), \tag{2.1.67}$$

$$|C(\mathbf{R})| = 1. \tag{2.1.68}$$

and therefore, $C(\mathbf{R})$ can be chosen in the following form

$$C(\mathbf{R}) = \exp(i\mathbf{k}\cdot\mathbf{R}) \tag{2.1.69}$$

that satisfies both the condition (2.1.68) and (2.1.68). The vector $\mathbf{k}$ in (2.1.69) is given by

$$\mathbf{k} = x_1\mathbf{b}_1 + x_2\mathbf{b}_2 + x_3\mathbf{b}_3, \tag{2.1.70}$$

where x_i are some complex numbers and $\boldsymbol{b}_i$ are the reciprocal lattice vectors defined as

$$\boldsymbol{a}_i \cdot \boldsymbol{b}_j = 2\pi\delta_{ij}. \tag{2.1.71}$$

Thus, we have obtained the following relation

$$\chi(\mathbf{r}+\mathbf{R}) = \exp(i\mathbf{k}\cdot\mathbf{R})\chi(\mathbf{r}). \tag{2.1.72}$$

Let us now consider a function $u_{\mathbf{k}}(\mathbf{r})$ which is defined as

$$u_{\mathbf{k}}(\mathbf{r}) = \exp(-i\mathbf{k}\cdot\mathbf{r})\chi(\mathbf{r}). \tag{2.1.73}$$

Making use of (2.1.72), it is straightforward to show that $u_{\mathbf{k}}(\mathbf{r})$ has the periodicity of lattice, i. e.

$$u_{\mathbf{k}}(\mathbf{r}+\mathbf{R}) = u_{\mathbf{k}}(\mathbf{r}). \tag{2.1.74}$$

Therefore, the wavefunction $\chi(\mathbf{r})$ can be represented with a lattice-periodic function $u_{\mathbf{k}}(\mathbf{r})$ multiplied by the factor $\exp(i\mathbf{k}\cdot\mathbf{r})$. This statement is known as Bloch's theorem.

Consequently, instead of solving the Schrödinger equation (2.1.60) for the wavefunction in an infinite space, the problem is recast to finding the wavefunction only within the cell defined by the vectors $\boldsymbol{a}_i$. Furthermore, the periodicity of the wavefunction (2.1.72) restricts the allowed values of the vectors $\mathbf{k}$. Let us consider the vector $\mathbf{k}' = \mathbf{k} + \boldsymbol{b}$, where $\boldsymbol{b} = m_1\boldsymbol{b}_1 + m_2\boldsymbol{b}_2 + m_3\boldsymbol{b}_3$ with $\{m_i\}$ being integers. We will label the wavefunction with its reciprocal vector for the convenience. Then

$$\chi_{\mathbf{k}}(\mathbf{r}+\mathbf{R}) = \exp(i\mathbf{k}\cdot\mathbf{R})\chi_{\mathbf{k}}(\mathbf{r}) = \exp(i(\mathbf{k}'-\boldsymbol{b})\cdot\mathbf{R})\chi_{\mathbf{k}}(\mathbf{r}) = \exp(i\mathbf{k}'\cdot\mathbf{R})\chi_{\mathbf{k}}(\mathbf{r}). \tag{2.1.75}$$

The state $\chi_{\mathbf{k}}$ satisfies the Bloch's theorem as if it had the wavevector $\mathbf{k}'$. Therefore, the original label $\mathbf{k}$ is not unique, and every state has a set of possible wavevectors that differ from each other by a vector of reciprocal lattice. This means that one may use only the wavevectors with the components defined in the range

$$-\frac{\pi}{a_i} < k_i \leq \frac{\pi}{a_i} \tag{2.1.76}$$

that defines the first Brillouin zone.

Now, to construct the electron density of periodic system one has to calculate the occupied

Hamiltonian eigenstates for each allowed **k**-vector within the first Brillouin zone. Hopefully, the wavefunctions and other properties, such as Hamiltonian eigenvalues, vary smoothly over the Brillouin zone [20] so that in practice, only a finite set of points needs to be chosen. The efficient choice of **k**-points in the reciprocal space is discussed in Refs. [21–26] in detail. Note that an increase of the cell size results in the contraction of Brillouin zone and therefore, the calculations of system with a large cell can be limited only to the center of Brillouin zone (Γ-point).

2.1.5 Basis sets.

In both the wavefunction and DFT-based methods one has to find a way of representing one-electron orbitals $\chi_i(\mathbf{r})$, whose analytical form is generally unknown. The natural solution of this problem is to expand the sought orbitals in a basis of known functions. Following this approach, the one-electron wavefunction $\chi_i(\mathbf{r})$ is written as

$$\chi_i(\mathbf{r}) = \sum_{j=1}^{\infty} c_j \phi_j(\mathbf{r}), \tag{2.1.77}$$

where $\{\phi_i(\mathbf{r})\}$ is a complete set of known basis functions and c_j are the expansion coefficients. It is, however, clear that for numerical applications one cannot use the infinite sum in (2.1.77) and therefore, the sum in (2.1.77) includes a limited number of terms.

In general, any complete set of function can be used as the basis functions $\phi_i(\mathbf{r})$. However, following the chemical sense and for the practical reasons two considerations can be applied while choosing the basis functions: (i) the basis functions should match the solution of one-electron Schrödinger equation as closely as possible (that would reduce the number of terms in (2.1.77)) and (ii) the basis functions should be computationally inexpensive (if by mischance the number of terms in (2.1.77) is still large).

Localized basis set.

The wavefunctions of an isolated atom exponentially decay at large distances and consequently, the function written as (2.1.77) is expected to mimic this exponential dependence. The most widely used basis sets that have such a behaviour are Slater-type orbitals (STOs) [27] and Gaussian-type orbitals (GTOs) [28]. These functions belong to the family of localized basis sets since they are written in a coordinate frame with the origin placed on the position of nucleus. The STOs and GTOs can be represented with the following general form

$$\phi_{nlm\zeta}(r, \theta, \varphi) = A\, R_{nl}(r) Y_{lm}(\theta, \varphi), \tag{2.1.78}$$

where A is a normalization factor, and $R_{nl}(r)$ and $Y_{lm}(\theta, \varphi)$ denote the radial and angular parts, respectively, with n, l, m being the quantum numbers. For both Slater and Gaussian basis functions the angular part is given by spherical harmonics and the difference between STOs and GTOs lies in the form of radial part.

Slater-type orbitals have the following analytical form for the radial part

$$R_{nl}(r) = r^{n-1}e^{-\zeta r}, \tag{2.1.79}$$

where parameter ζ determines the rate of exponential decay and the value of the parameter can be related to an effective charge of the nucleus.

The radial part of GTOs is written in Cartesian coordinates as

$$R_{nl}(r) = x^p y^q z^s \, e^{-\alpha r^2}, \tag{2.1.80}$$

with the exponents p, q, and s that sum to the value of the quantum number l ($p + q + s = l$) and the parameter α determines the spread of the Gaussian. Obviously, the function (2.1.80) does not follow the exponential law as STO (2.1.79) does. The reason for using the GTOs is purely practical. In fact, the GTOs are much more computationally efficient than the STOs and many quantities can be calculated analytically [29] with the Gaussian functions (2.1.80). Consequently, many of the modern quantum-chemical methods employ the Gaussian-type orbitals as the basis set. More exactly, linear combinations of such functions with different exponents and expansion coefficients are formed to approximate the one-electron orbitals $\chi_i(\mathbf{r})$.

Plane-wave basis set.

The localized basis sets in the form of STOs or GTOs is a natural choice in the calculations of isolated systems. However, in a periodic structure the one-electron wavefunction should satisfy the Bloch theorem and thus, is to be written in the form (see 2.1.4)

$$\chi(\mathbf{r}) = \exp(i\mathbf{k}\cdot\mathbf{r})\, u_{\mathbf{k}}(\mathbf{r}), \tag{2.1.81}$$

where the function $u_{\mathbf{k}}(\mathbf{r})$ has the periodicity of the system in the direct space. The function is constructed by expanding it in the basis of plane waves

$$u_{\mathbf{k}}(\mathbf{r}) = A \sum_{\boldsymbol{b}} c_{\boldsymbol{b}} \exp(i\boldsymbol{b}\cdot\mathbf{r}), \tag{2.1.82}$$

with A and $\boldsymbol{b}$ being the normalization constant and the reciprocal lattice vector, respectively. Note that plane waves form a complete and orthonormal basis and, in contrast to atom-centered STOs and GTOs, the plane waves are independent on atomic positions (delocalized). This feature has important consequences for applications. Combining eqs. (2.1.81) and (2.1.82), one obtains the following expression for the one-electron function χ

$$\chi(\mathbf{r}) = \sum_{\boldsymbol{b}} c_{\boldsymbol{b}} \exp(i(\mathbf{k}+\boldsymbol{b})\cdot\mathbf{r}). \tag{2.1.83}$$

As in the case of the localized basis sets, only a finite number of plane waves can be taken in the expansion (2.1.83). Consequently, for each **k**-point in the Brillouin zone, only those $\boldsymbol{b}$ vectors are considered that satisfy the cutoff criterion

$$\frac{\mathbf{k}^2 + \boldsymbol{b}^2}{2} \leq E_{\text{cut}}. \tag{2.1.84}$$

The plane waves basis set has several advantages over localized basis sets. These are

- the same basis set can be used for different atomic species;
- the convergence of results with respect to the size of basis set can easily be tested by increasing the E_{cut} criterion;
- the plane waves are independent on the positions of nuclei and therefore, an error in Hellmann-Feynman forces, which is called the Pulay force [30], is exactly zero.

The major disadvantage of plane wave basis set is that a very large number of plane waves is necessary to represent the atomic wavefunction and because of this reason the plane wave basis set is almost never used in all-electron calculations. The number of plane waves can, however, be drastically reduced by the use of pseudopotential approach. Consequently, the plane wave basis set combined with pseudopotentials has become a widely used computational scheme in the calculations of periodic systems.

In principle, the delocalized plane wave basis set can also be used for isolated systems, e.g. molecules. For this purpose, the system should be placed into a periodic cell large enough such that the interactions of the system with its images due to the use of periodic boundary conditions were negligible.

2.1.6 Pseudopotentials.

The atomic wavefunctions $\chi_i(\mathbf{r})$ are eigenfunctions of the atomic Hamiltonian and therefore, the functions must all be orthogonal to each other. Since the wavefunctions of low-lying states (core states) are localized in the vicinity of nucleus, the valence wavefunctions must oscillate rapidly in the region close to the nuclei in order to maintain the orthogonality with the core states. These rapid oscillations result in a large kinetic energy of valence electrons in the core region, but the energy is roughly canceled by potential energy due to strong Coulomb potential of nucleus. Consequently, the valence electrons are much more weakly bound than the core electrons.

The pseudopotential approximation is based on the observation that the core electrons remain relatively unaffected by the chemical environment of atom. The key idea of pseudopotential is to replace the interaction of valence states with the potential of nucleus and with chemically inert core states by a more weak effective potential (pseudopotential). Consequently, the rapidly oscillating exact wavefunction of the valence state in the core region is replaced by a smooth pseudo-wavefunction, whereas this effective wavefunction remains identical to the real wavefunction in the chemically relevant outer region.

The pseudopotential approach allows reducing the number of electrons that have to be explicitly treated in the calculations. Furthermore, when combined with the plane wave basis set, lower cutoffs E_{cut} can be used since there is no need to use a large number of plane waves in (2.1.82) to describe the rapid oscillations of real valence wavefunctions in the core region.

A "good" pseudopotential should satisfy a number of conditions first formulated by Hamann *et al.* [31]:

- real and pseudo valence eigenvalues agree for a chosen reference configuration;
- real and pseudo valence wavefunctions agree beyond the chosen core radius R_{c};
- the integrals over the core region of the real and valence charge densities agree for each valence state (norm conservation condition);
- the logarithmic derivatives of the real and pseudo wavefunction and their first energy derivatives agree in the region $r \geq R_{\text{c}}$.

The pseudopotentials that obey all these conditions are norm-conserving pseudopotentials. There exist several methods of generating the norm-conserving pseudopotentials such as Bachelet-Hamann-Schlüter [32], Troullier-Martins [33], or Rappé *et al.* schemes [34]. In a simplified form an algorithm for the generation of pseudopotential includes the following steps:

i) choose XC functional to work with;

ii) solve all-electron Schrödinger equation for atom with the chosen XC functional;

iii) choose which states will belong to the core and which to the valence shell;

iv) choose the reference state for the construction of pseudopotential;

v) choose the core radii for each angular momentum (and any other parameters);

vi) generate pseudoptential making use of one of the methods mentioned above;

vii) test pseudopotential for transferability to other states.

Other than norm-conserving pseudopotentials exist too. Thus, Vanderbilt [35] proposed to relax the norm-conservation condition that allows generation of pseudopotentials that behave more soft in the core region. These "ultrasoft" pseudopotentials permit to use very small energy cutoffs in the calculations with the drawback of necessity to perform some additional computations because of the absence of norm-conservation condition.

2.2 Computational methods.

The models described above are realized in methods that translate the model into a set of equations thus providing practical recipes of how one can obtain the desired results. These recipes are then coded in computer programs using numerical algorithms that solve the basic equations with computers. The methods used in the present work can generally be divided into two groups. The

first group includes the methods based on the energy minimization followed by the computation of system characteristics once the energy minimum (stable state of the system) was reached. The second group of methods is the molecular dynamics method that computes the trajectory of system in the phase space and then uses the ergodicity theorem to yield the mean system characteristics from their instantaneous values.

To be operational, the methods of both groups need to know the potential energy surface in the space of atomic coordinates. This potential energy surface, which generally has the dimension of $3N$ (N being the number of atoms), can be obtained with either a quantum-chemical model or an effective potential model. However, the equations that need to be solved are generally independent on the underlying model of interatomic interactions and the discussion below presents the basics of both group of methods, briefly touches some issues related to the use of specific model of interatomic interactions, and illustrates, by few examples, the computation of system characteristics.

2.2.1 Methods of energy minimization.

Regardless the computational method, the first information that needs to be known about the system is its structure, i.e. the arrangement of atoms in the space. This information is commonly taken from available experimental data, but this initial guess of system geometry is not necessarily the one, which corresponds to the energy minimum for a given model of interatomic interactions. As the minima on the potential energy surface correspond to the stable states of the system, the methods of energy minimization try to find these minima with the use of different computational algorithms. The energy minimization method combined with a model of effective potentials functions, known as molecular mechanics, has been widely used from the beginning of 1970s for studying the conformational space of organic molecules and later on, it has been increasingly applied to other systems, such as polymers, biomolecules, and heterogeneous systems. In the case of *static* calculations the energy minimization is the mandatory step prior to computing properties of the system.

The problem of energy minimization can be stated as

$$\forall\, \xi_i : \quad \frac{\partial E}{\partial \xi_i} = 0, \quad \frac{\partial^2 E}{\partial \xi_i^2} > 0, \tag{2.2.1}$$

where E is the energy, which is a function of coordinates ξ_i. The latter can be either $3N$ Cartesian coordinates of atoms or internal coordinates, such as bond length, valence angles, etc. The algorithms used to explore the potential energy surface $E(\xi_1, \xi_2, \ldots\, \xi_{3N})$ can be classified into two groups. Those which use the derivatives of energy with respect to coordinates and those which do not. The most modern methods employ one of the algorithms of the first group that also differ in the order of derivatives used. The first-order algorithm use the information only about the derivative of energy, whereas the second-order algorithms take both the first and second derivatives into account.

Newton-Raphson method.

The energy minimization algorithm is illustrated below with the simplest second-order algorithm, Newton-Raphson method. For this purpose, let us first consider an one-dimensional case and write the energy $E(\xi)$ as a Taylor series about the point ξ_k

$$E(\xi) = E(\xi_k) + (\xi - \xi_k)\left(\frac{\partial E}{\partial \xi}\right)_{\xi_k} + \frac{1}{2}(\xi - \xi_k)^2\left(\frac{\partial^2 E}{\partial \xi^2}\right)_{\xi_k} + \dots . \tag{2.2.2}$$

The first derivative of the energy is then given by

$$\frac{\partial E(\xi)}{\partial \xi} = \left(\frac{\partial E}{\partial \xi}\right)_{\xi_k} + (\xi - \xi_k)\left(\frac{\partial^2 E}{\partial \xi^2}\right)_{\xi_k} \tag{2.2.3}$$

and if the system has an energy minimum in the point ξ^* then $(\partial E/\partial \xi) = 0$ at $\xi = \xi^*$ and the point ξ^* can be found as

$$\xi^* = \xi_k - \left(\frac{\partial E}{\partial \xi}\right)_{\xi_k}\left[\left(\frac{\partial^2 E}{\partial \xi^2}\right)_{\xi_k}\right]^{-1} . \tag{2.2.4}$$

In the multi-dimensional case, eqn. (2.2.4) reads

$$\boldsymbol{\xi}^* = \boldsymbol{\xi}_k - \mathbf{g}_k\,\mathbf{H}_k^{-1}, \tag{2.2.5}$$

where $\mathbf{g}_k$ and $\mathbf{H}_k$ denote the vector of derivatives and the matrix of second derivatives of the energy (Hessian matrix), respectively, in the point $\boldsymbol{\xi}_k$.

The advantage of the second-order methods is that they use the information about not only the direction to move (vector $\mathbf{g}$), but also the slope of the potential energy surface (matrix $\mathbf{H}$). The disadvantage is that the matrix of second derivatives needs to be computed (and inverted) at each minimization step that can be rather expensive, especially in quantum-chemical calculations. Furthermore, the analytical second derivatives of the energy with respect to coordinates might not be known for some quantum-chemical models, while computation of the Hessian by the finite differences can be very time-consuming procedure. The problem is circumvented by the use of quasi-Newton methods that construct the Hessian matrix at a given minimization step using the Hessian matrix of previous step, and the vectors $\boldsymbol{\xi}$ and $\mathbf{g}$ of the current and previous steps. Examples of such methods, which are most widely used in the quantum-chemical calculations, are Davidson-Fletcher-Powell (DFP) or Broyden-Fletcher-Goldfarb-Shanno (BFGS) methods.

Computation of characteristics. Vibrational analysis.

Once the system has reached the energy minimum (local or global), its state corresponds to a stable state at $T = 0$ K, where no motion occurs. It should be noted that the energy of a system is non-zero even at the temperature $T = 0$ K. This energy E_{zp} is called the zero-point energy, whose value is computed from the vibrational analysis, and should always be taken into account

when comparing the relative energies of conformations.

$$E_{zp} = \sum_{k=1}^{3N-3} \frac{1}{2}\hbar\omega_k, \tag{2.2.6}$$

where the sum runs over all vibrations of the system.

Many characteristics can be computed for the system in the state and the following presentation deals with the analysis of the vibrational dynamics of a periodic system. The analysis relies on the fact that the vibrational dynamics of a system of heavy atoms ($M_I > M_H$) at relatively low temperature is well described by the harmonic approximation.

Let us consider a system of atoms replicated in the space along the three dimensions such that the system is surrounded by infinite number of its images (cf. Figure 2.1). The vector $\mathbf{u}_{s\boldsymbol{l}}$ is the displacement vector of the s-th atom in the $\boldsymbol{l}$-th cell from its equilibrium position (this implies that at the equilibrium all $\mathbf{u}_{s\boldsymbol{l}}$ vectors are zero). If one assumes that the displacements are small, the potential energy $V(\mathbf{u}_{s\boldsymbol{l}})$ can be expanded into a Taylor series with respect to the Cartesian components i ($i = x, y, z$) of the displacements $u^i_{s\boldsymbol{l}}$ up to the second order

$$V = V_0 + \sum_{s,\boldsymbol{l},i} u^i_{s\boldsymbol{l}} \left[\frac{\partial V}{\partial u^i_{s\boldsymbol{l}}}\right]_0 + \frac{1}{2} \sum_{ss',\boldsymbol{ll'},ii'} u^i_{s\boldsymbol{l}} u^{i'}_{s'\boldsymbol{l'}} \left[\frac{\partial^2 V}{\partial u^i_{s\boldsymbol{l}} \partial u^{i'}_{s'\boldsymbol{l'}}}\right]_0 + \dots . \tag{2.2.7}$$

The constant term in (2.2.7) is irrelevant for the dynamics and can be omitted, whereas the linear terms vanish because the system is in the energy minimum. Thus, only the last terms quadratic in the atomic displacements remain in the right-hand side. Kinetic energy of the system can be written as

$$E_{\text{kin}} = \frac{1}{2} \sum_{s\boldsymbol{l}} M_s |\dot{\mathbf{u}}_{s\boldsymbol{l}}|^2, \tag{2.2.8}$$

where M_s is the mass of atom s and the index s runs over all atoms in a unit cell, and $\boldsymbol{l}$ runs over all unit cells.

The equations of motion for a generalized coordinate q and the conjugated momentum $\dot{q}$ are obtained from the Lagrangian $\mathcal{L} = E_{\text{kin}} - V$ as

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}}\right) - \frac{\partial \mathcal{L}}{\partial q} = 0. \tag{2.2.9}$$

Making use of eqs. (2.2.7) and (2.2.8) one can obtain the equations of motion for atoms in the system in the form

$$M_s \ddot{u}^i_{sl} = - \sum_{s',l',i'} \left[\frac{\partial^2 V}{\partial u^i_{sl} \partial u^{i'}_{s'l'}}\right]_0 u^{i'}_{s'l'}. \tag{2.2.10}$$

Eq. (2.2.10) is valid for all atoms s in the cell, for all cells $\boldsymbol{l}$, and for each Cartesian component of the displacement vector. If one introduces the Cartesian tensor defined through its component

ii' as

$$G^{ii'}_{s\boldsymbol{l},s'\boldsymbol{l}'} = \left[\frac{\partial^2 V}{\partial u^i_{s\boldsymbol{l}} \partial u^{i'}_{s'\boldsymbol{l}'}}\right]_0, \tag{2.2.11}$$

then eq. (2.2.10) can be rewritten as

$$M_s \ddot{\mathbf{u}}_{s\boldsymbol{l}} = -\sum_{s',\boldsymbol{l}'} \mathbf{G}_{s\boldsymbol{l},s'\boldsymbol{l}'} \cdot \mathbf{u}_{s'\boldsymbol{l}'}. \tag{2.2.12}$$

Each term of the sum in the right-hand side of eq. (2.2.12) can be viewed as force acting on the atom s in the $\boldsymbol{l}$-th cell due to displacement $\mathbf{u}_{s'\boldsymbol{l}'}$ of atom s' in the cell $\boldsymbol{l}'$.

The elements of tensor $\mathbf{G}$ should depend not on absolute values of cell indexes $\boldsymbol{l}$ and $\boldsymbol{l}'$, but on their relative value $\Delta\boldsymbol{l} = \boldsymbol{l} - \boldsymbol{l}'$ such that

$$\mathbf{G}_{s\boldsymbol{l},s'\boldsymbol{l}'} = \mathbf{G}_{ss'}(\Delta\boldsymbol{l}). \tag{2.2.13}$$

Eq. (2.2.12) can be recast as

$$M_s \ddot{\mathbf{u}}_{s\boldsymbol{l}} = -\sum_{s',\Delta\boldsymbol{l}} \mathbf{G}_{ss'}(\Delta\boldsymbol{l}) \cdot \mathbf{u}_{s',\boldsymbol{l}+\Delta\boldsymbol{l}}. \tag{2.2.14}$$

The equation described by (2.2.14) is translationally invariant and its solution must satisfy Bloch's theorem:

$$\mathbf{u}_{s\boldsymbol{l}}(t) = \exp(i\boldsymbol{b}\cdot\boldsymbol{l})\mathbf{u}_{s\mathbf{0}}(t), \tag{2.2.15}$$

where t is time, $\boldsymbol{b}$ is the reciprocal lattice vector, and the subscript zero denotes the principal cell.

After substitution of eq. (2.2.15) into eq. (2.2.14) one obtains

$$M_s \ddot{\mathbf{u}}_s(t;\boldsymbol{b}) = -\sum_{s'} \mathbf{G}_{ss'}(\boldsymbol{b}) \cdot \mathbf{u}_{s'}(t;\boldsymbol{b}), \tag{2.2.16}$$

where $\mathbf{u}_s(t;\boldsymbol{b}) = \mathbf{u}_{s\mathbf{0}}$ and

$$\mathbf{G}_{ss'}(\boldsymbol{b}) = \sum_{\Delta\boldsymbol{l}} \mathbf{G}_{ss'}(\Delta\boldsymbol{l}) \exp(i\boldsymbol{b}\cdot\Delta\boldsymbol{l}) \tag{2.2.17}$$

is the Fourier transform of the force constant tensor $\mathbf{G}_{ss'}(\Delta\boldsymbol{l})$. It is noteworthy, that the starting set of equations (2.2.12) for a system containing N cells and n atoms per cell contains $3nN$ equations, whereas the final set has only $3n$ equations (2.2.16).

The solution of (2.2.16) is sought in the form

$$\mathbf{u}_s(t;\boldsymbol{b}) = \mathbf{u}_s(\boldsymbol{b})\, \exp(i\omega t), \tag{2.2.18}$$

that leads to the the following set of equations

$$\sum_{s'i'} \{\mathbf{G}^{ii'}_{ss'}(\boldsymbol{b}) - \omega^2 M_s \delta_{ss'}\delta_{ii'}\} u^{i'}_{s'}(\boldsymbol{b}) = 0. \tag{2.2.19}$$

The last $3n$ equations represent an eigenvalue problem with respect to ω^2 that can easily be solved to yield the vibrational frequencies of the system for the given reciprocal lattice vector $\boldsymbol{b}$.

2.2.2 Molecular dynamics.

The molecular dynamics (MD) method was first introduced by Alder and Wainwright [36,37] and applied to a hard spheres system. The next step in the development of the MD technique was done by A. Rahman who used a more realistic Lennard-Jones (12-6) potential to describe the interatomic interactions in liquid argon [38]. The basic idea of MD method is to use computer for numerical integration of the classical equations of motion of particles in a system with the known interaction potential. The procedure yields the trajectory of the system in the phase space and the desired properties of the system can be calculated by averaging their instantaneous values over sufficiently long trajectory (time). Such a procedure implies the validity of ergodic theorem.

The points visited by the system in the phase space are distributed according to a specific distribution function. The function is determined by a set of fixed macroscopic parameters, such as temperature, pressure, volume, number of particles, etc. These macroscopic parameters are said to define specific statistical ensemble. Thus, if one considers a system described by the Lagrangian $\mathcal{L} = T - V$, where T and V are the classical kinetic and potential energies, respectively, the time evolution of the system is described by the Newton's equation of motion

$$\ddot{\mathbf{R}}_i = -\frac{1}{M_i}\boldsymbol{\nabla}_i V(\{\mathbf{R}_j\}) \tag{2.2.20}$$

relating the acceleration of particle i with mass M_i to the force acting on the particle (the negative of the potential energy gradient). Since the Hamiltonian of the system $H = T + V$ does not explicitly depend on time, $\partial H/\partial t = 0$ and therefore, the total energy, which is identical to the Hamiltonian in this case, is a conserved quantity. The macroscopic characteristics fixed in such calculations are the number of particles N, the volume V, and the total energy E, i.e. the calculations generate trajectory corresponding to the microcanonical (NVE) statistical ensemble. The NVE ensemble is the natural statistical ensemble in MD calculations. Simulations in other statistical ensembles require a modification of the equations of motion to generate the corresponding trajectory of the system in the phase space.

To perform a molecular dynamics simulation, the following items should be specified

- The model of interatomic interactions. Two approaches are mostly used. In the first approach, called classical MD, the interactions are described by a set of effective potential functions. The second approach uses quantum-chemical methods to obtain forces acting on atoms. One refers to such MD calculations as *ab initio* molecular dynamics if they use *ab initio* quantum-chemical methods.
- A statistical ensemble has to be chosen to keep specified macroscopic parameters fixed. This choice assumes that the equations of motion corresponding to the chosen ensemble have to be derived from the expression for the Lagrangian of the system.

- Given the equations of motion for the specific statistical ensemble, an integration algorithm should be developed to propagate the system in the phase space from time t to $t + \Delta t$. Since, the numerical integration uses a finite difference scheme, the integration time-step Δt should guarantee the stability of the integration procedure.
- The initial configuration of the system should be given. Even if the thermal motion will later destroy the initial configuration, it is desirable that the configuration is close to a real structure. Otherwise, the equilibration of the system can take a long time; this point has a particular importance for *ab initio* simulations.
- Finally, a model relating the position of the system in the phase space, i.e. atomic coordinates and velocities, to the quantities of interest should be designed. The mean values of these quantities are then obtained by averaging their instantaneous values along the MD trajectory.

As it was mentioned above the natural statistical ensemble in MD simulations is the NVE ensemble. Other statistical ensembles widely used in the MD simulations are the canonical (isothermic) NVT ensemble and isothermic-isobaric NPT ensemble. Equations of motion generating the trajectory in the phase space corresponding to these ensembles are often obtained with the help of the extended Lagrangian approach. Thus, for modelling the system in the NVT ensemble the most popular scheme is the Nosé-Hoover thermostat [39–41].

Nosé-Hoover thermostat. The main idea of the method is to consider the heat bath as an integral part of the system by introduction of a variable s with a mass $Q > 0$ and velocity $\dot{s}$. The variable s can be viewed as a time-scaling parameter: the time in the extended system is stretched by the factor s

$$d\tilde{t} = s\, dt. \tag{2.2.21}$$

Since the atomic coordinates are identical in both systems, the scaled, $\tilde{\mathbf{r}}$ and $\tilde{s}$, and the original, $\mathbf{r}$ and s, variables are related to each other via

$$\tilde{\mathbf{r}} = \mathbf{r},\ \dot{\tilde{\mathbf{r}}} = s^{-1}\dot{\mathbf{r}},\ \tilde{s} = s,\ \text{and}\ \dot{\tilde{s}} = s^{-1}\dot{s}. \tag{2.2.22}$$

The Lagrangian of the extended system $\tilde{\mathcal{L}}$ is written in the scaled variables as

$$\tilde{\mathcal{L}} = \sum_{i=1}^{N} \frac{\tilde{\mathbf{p}}_i^2}{2M_i\tilde{s}^2} - U(\{\tilde{\mathbf{r}}\}) + \frac{\tilde{p}_s^2}{2Q} - g\, k_b T \ln \tilde{s}, \tag{2.2.23}$$

where $\tilde{\mathbf{p}}_i$ and $\tilde{p}_s$ are the scaled momenta of particle i and of the variable s, respectively, and g is the number of independent degrees of freedom of the system. The first two terms in (2.2.23) are the classical kinetic and potential energies, respectively, whereas the last two terms can be viewed as the kinetic and potential energies of the variable s.

The equations of motion that can be derived from (2.2.23) by using (2.2.9) read (in real

coordinates $\mathbf{r}$, $\mathbf{p}$, s, and p_s)

$$\begin{aligned}
\dot{\mathbf{r}_i} &= \frac{\mathbf{p}_i}{M_i}, && (2.2.24)\\
\dot{\mathbf{p}_i} &= -\frac{\partial U}{\partial \mathbf{r}} - \frac{s p_s \mathbf{p}_i}{Q}, && (2.2.25)\\
\dot{s} &= \frac{s^2 p_s}{Q}, && (2.2.26)\\
\dot{p_s} &= \frac{1}{s}\left(\sum_{i=1}^{N}\frac{\mathbf{p}_i^2}{M_i} - g k_b T\right) - \frac{s p_s^2}{Q}. && (2.2.27)
\end{aligned}$$

The equations (2.2.24) – (2.2.27) generate a trajectory corresponding to the canonical (isothermic) statistical ensemble in the phase space of particles. The quantity conserved in the MD simulations using the Nosé-Hoover thermostat is the *extended* Hamiltonian of the system

$$H_{\text{Nosé}} = \sum_{i=1}^{N}\frac{\mathbf{p}_i^2}{2M_i} + U(\{\mathbf{r}\}) + \frac{s^2 p_s^2}{2Q} + g k_b T \ln s. \qquad (2.2.28)$$

A special care must be taken in choosing the correct value for the fictitious mass variable Q. In the case of very large values of Q the temperature control is poor (in the limit of $Q \to \infty$ the method generates the microcanonical ensemble). On the other hand, too small Q may cause too large and rapid oscillations of the temperature.

To be solved with computers, the equations of motion (2.2.20) or (2.2.24) – (2.2.27) need to be given in a discrete form. The standard approach is to solve these equations using finite differences methods. The most widely used algorithm to integrate the Newton's equations of motion (2.2.20) is the velocity Verlet integrator

$$\begin{aligned}
\mathbf{r}_i(t+\Delta t) &= \mathbf{r}_i(t) + \mathbf{v}_i(t)\,\Delta t + \frac{1}{2}\,\mathbf{a}_i(t)\,\Delta t^2,\\
\mathbf{v}_i(t+\Delta t) &= \mathbf{v}_i(t) + \frac{1}{2}\left[\mathbf{a}_i(t) + \mathbf{a}_i(t+\Delta t)\right]\Delta t, && (2.2.29)
\end{aligned}$$

where $\mathbf{r}_i(t)$, $\mathbf{v}_i(t)$, $\mathbf{a}_i(t)$ are the coordinate, velocity, and the acceleration of particle i at time t, respectively. The velocity Verlet algorithm is suited for the equations of motion, where the force acting on particle is a function of coordinates. However, if the force also depends on the particle's velocity or explicitly coupled to other degrees of freedom, other integration schemes, such as predictor-corrector, have to be used. Size of the time-step Δt in (2.2.29) is a matter of trade between the accuracy of the integration and the length of MD trajectory to follow. The use of a small Δt results in a stable integration, but only a small part of the phase space can be sampled in reasonable computing time, that questions the validity of the ergodic hypothesis. A large time-step allows a better exploration of the phase space at the risk of system "exploding" due to unstable integration of the equations of motion. Typically, the time-step Δt has to be taken by an order of magnitude smaller than the smallest characteristic time in the system, e.g.

the shortest vibrational period.

Ab initio molecular dynamics.

The main idea of any *ab initio* MD method is to compute forces acting on the nuclei from electronic structure calculations performed "on-the-fly", along the MD trajectory. There are several versions of *ab initio* molecular dynamics models. The most well known among them are the Born-Oppenheimer and Car-Parrinello molecular dynamics.

Born-Oppenheimer molecular dynamics. The equations of motion can be greatly simplified by applying the adiabatic approximation that leads to a static problem for electronic subsystem at each molecular dynamics time step with the fixed nuclear positions. Consequently, the initial problem is reduced to the solution of the time-independent electronic Schrödinger equation (2.1.5) followed by propagating the nuclei via classical molecular dynamics using the forces found from the solution. The Born-Oppenheimer molecular dynamics (BOMD) is defined through the equations

$$M_I \ddot{\mathbf{R}}_I(t) = -\boldsymbol{\nabla}_I \min_{\psi_0} \{\langle \psi_0 | H_e | \psi_0 \rangle\}, \tag{2.2.30}$$

$$H_e \psi_0 = E_0 \psi_0 \tag{2.2.31}$$

for the electronic ground state wavefunction ψ_0. From the equations (2.2.30) and (2.2.31) it follows that the time-dependence of the electronic ground state wavefunction is parametric through the nuclear positions.

The disadvantage of the Born-Oppenheimer MD is that the computationally expensive electronic structure problem needs to be solved at each MD step. Since the computational burden grows as at least third power of the number of basis functions, the computational load for large systems can become very prohibitive. The advantage of the BOMD method is that, due to use of adiabatic approximation, the electronic and nuclear coordinates are completely decoupled from each other and one can use the time-step for integration of equations of motion of nuclei equal to that in classical molecular dynamics simulations of the same system (ca. 0.5 fs for systems with light hydrogen atoms). In addition, the many-electron wavefunction is exactly the ground state wavefunction at each MD time step.

Car-Parrinello molecular dynamics. In 1985 Car and Parrinello suggested a way [42] avoiding the straightforward solution of the electronic Schrödinger equation (2.2.31) at each MD time step, as it is done in the Born-Oppenheimer MD. They formulated a new efficient method, where the expansion coefficients c of the one-electron wavefunctions $\chi(\mathbf{r})$ in (2.1.77) are considered as dynamical variables. Degrees of freedom associated with the nuclei as well as these additional orbital degrees of freedom are treated at equal footing and their evolution in time is obtained by numerical integration of equations of motion.

The equations of motion in the Car-Parrinello method are derived from the extended La-

grangian that reads

$$L = \frac{1}{2}\sum_{I=1} M_I \dot{\mathbf{R}}_I^2 + \frac{1}{2}\mu \sum_{i=1} \langle \dot{\chi}_i | \dot{\chi}_i \rangle - \langle \psi_0 | H_e | \psi_0 \rangle + constraints, \tag{2.2.32}$$

where μ is a "fictitious" mass assigned to the orbital degrees of freedom. In eq. (2.2.32) first two terms correspond to the kinetic energies of nuclei and orbitals, respectively, the third term is the potential energy, and the last term includes constraints [42], such as the orthonormality conditions imposed on the one-electron wavefunctions $\chi(\mathbf{r})$

$$\int d\mathbf{r}\, \chi_i^*(\mathbf{r})\chi_j(\mathbf{r}) = \delta_{ij}. \tag{2.2.33}$$

The equations of motion for the nuclei and the orbital degrees of freedom are obtained from the Langrangian (2.2.32) by (2.2.9) and have the following form

$$M_I \ddot{\mathbf{R}}_I(t) = -\frac{\partial}{\partial \mathbf{R}_I}\langle \psi_0 | H_e | \psi_0 \rangle + \frac{\partial}{\partial \mathbf{R}_I}(constraints), \tag{2.2.34}$$

$$\mu \ddot{\chi}_i(t) = -\frac{\delta}{\delta \chi_i}\langle \psi_0 | H_e | \psi_0 \rangle + \frac{\delta}{\delta \chi_i}(constraints). \tag{2.2.35}$$

As it follows from eqs. (2.2.34) and (2.2.35) the existence of constraints leads to "constraint forces" in the equations of motion. It should also be noted that the constraints can also depend both on nuclei positions $\mathbf{R}_I$ and orbitals χ_i, see for example [43].

The second term in Langrangian (2.2.32) is a "fictitious" kinetic energy of electronic degrees of freedom. A low value of the energy means that the system is close to the Born-Oppenheimer energy surface and therefore, the wavefunction is close to the ground state wavefunction. Typically, chains of Nosé-Hoover thermostats coupled to the nuclear and electronic degrees of freedom are used to keep the system close to the Born-Oppenheimer energy surface. The "fictitious" mass μ should be taken such that the lowest frequency associated with the electronic degrees of freedom lies well above the highest phonon frequency in the system. This lowest possible electronic frequency can be estimated as

$$\omega_e^{\min} \propto \left(\frac{E_{\text{gap}}}{\mu}\right)^{1/2}, \tag{2.2.36}$$

where E_{gap} is the energy difference between the highest occupied and lowest unoccupied orbitals in the system. On the other hand, value of the mass constant μ determines the time-step Δt that can be used for the stable integration of equations of motion (2.2.34) and (2.2.35). Typical values for large-gap systems are $\mu = 500 - 1000$ a.u. with a time-step of $4 - 8$ a.u. ($\approx 0.1 - 0.2$ fs). One sees that the time-step in CPMD method is several times smaller than the time-step used in the BOMD simulations. For large systems this disadvantage is largely compensated by more efficient treatment of electronic degrees of freedom in the Car-Parrinello molecular dynamics.

Computation of characteristics.

The molecular dynamics simulations provide the access to both structural and dynamical characteristics of the system. Among structural information extracted from the results of MD simulation, the pair radial distribution function (RDF) is, perhaps, one of the most widely used characteristics. The RDF $g_{\alpha\beta}(r)$ shows the probability to find an atom of kind β at the distance r from an atom of kind α divided by that obtained for the case of uniform distribution of atoms in the system, and $g_{\alpha\beta}(r)$ is computed as

$$g_{\alpha\beta}(r) = \frac{V N_\beta(r)}{4\pi r^2 \Delta r \, N_\beta}, \tag{2.2.37}$$

where V is the volume of system, $N_\beta(r)$ is the number of atoms β in a spherical layer of thickness Δr at the distance r from the atom of kind α, and N_β is the total number of atoms of kind β in the system. For crystalline systems the RDF has relatively sharp peaks corresponding to the first, second, etc. coordination spheres, while RDFs of disordered structures typically contain a well defined first peak, while the structural disorder blurs peaks in RDF at longer distances. The radial distribution function can be compared with the results of X-ray and neutron diffraction experiments.

The calculation of the vibrational spectra in MD simulations is different from the method based on the harmonic approximation that is discussed in Section 2.2.1. The analysis of the vibrational dynamics in MD uses the time-correlation function formalism and relates the ensemble averaged time-correlation function of a dynamical variable to the corresponding vibrational spectrum. Thus, the density of vibrational states of the system $D(\omega)$ is computed as

$$D(\omega) \propto \int dt \, e^{i\omega t} \sum_k \langle \tilde{\mathbf{v}}_k(0) \cdot \tilde{\mathbf{v}}_k(t) \rangle, \tag{2.2.38}$$

where $\tilde{\mathbf{v}}_k = \sqrt{M_k}\, \mathbf{v}_k$ is the mass-weighted velocity of atom k. In similar way, one can calculate the infrared and Raman spectra by replacing $\tilde{\mathbf{v}}_k$ with the vector of dipole moment or by with the polarizability tensor [44–48]. The MD computed spectra are noway restricted to the harmonic approximation, except for the case when the potential energy surface is described by harmonic potentials.

The models and methods discussed above have been realized in computer programs that were used in the present work to obtain results presented in the subsequent chapters. These are quantum-chemical codes Gaussian03, CRYSTAL06, CPMD, and ABINIT. The former was used in studies of isolated systems, whereas the three latter programs were employed in modeling of periodic structures. In addition, a number of home-built codes were used the energy minimization and molecular dynamics calculations using the effective potential models, for the data analysis, and for computing characteristics of the systems within developed models.

2.3 Numerical methods of parameters optimization.

2.3.1 Cost function.

As it was already mentioned, a force field is a parameter-dependent model and the parameters of the model are usually defined by reproducing some reference data as closely as possible. For this purpose a cost function is defined, that measures the deviation of the target characteristics from their reference values, and the function is then minimized in the multi-dimensional space of parameters. The cost function is expected to be flexible and insensitive to a small numerical noise that allows obtaining stable solution of the minimization problem. One of the most widely used cost functions is defined as

$$\chi^2(\mathbf{p}) = \sum_i^N \frac{(y_i - f(x_i, \mathbf{p}))^2}{\sigma_i^2}, \tag{2.3.1}$$

where $\mathbf{p}$ is the vector of sought parameters, y_i are the reference values, $f(x_i, \mathbf{p})$ is a model function and σ_i are the weights given to each point. The cost function defined in such a way is always non-negative.

The methods used in the present work for the minimization of the cost function (2.3.1) can be divided into two groups: those based only on the values of the function itself and those which additionally use derivatives of the function. Note that in certain sense the problem is similar to the energy minimization discussed in section 2.2.1 and thus, the approaches for solving the problem are also to be similar.

2.3.2 Minimization methods without gradients.

Singular value decomposition (SVD).

Let us consider the model function $f(x_i, \mathbf{p})$ in (2.3.1) that has the following form

$$f(x_i, \mathbf{p}) = \sum_j^M p_j \, F_j(x_i), \tag{2.3.2}$$

where the quantities $F_j(x_i)$ can be viewed as basis functions and they can be nonlinear functions of x. The model function $f(x_i, \mathbf{p})$ defined in this way is a linear function of the parameters p_j that allows a simplification of the minimization of the cost function (2.3.1).

The minimum of (2.3.1) occurs when the derivatives of the cost function with respect to all M parameters p_j are zeros that is, in the case of model function given by (2.3.2), equivalent to

$$\sum_i^N \frac{1}{\sigma_i^2} \Big[y_i - \sum_j^M p_j \, F_j(x_i) \Big] \, F_k(x_i) = 0 \quad k = 1, \ldots, M. \tag{2.3.3}$$

Equation (2.3.3) can be rewritten in the form

$$\sum_{j}^{M} A_{kj} p_j = b_k, \tag{2.3.4}$$

where the elements A_{kj} and b_k are given by

$$A_{kj} = \sum_{i}^{N} \frac{F_k(x_i)\, F_j(x_i)}{\sigma_i^2}; \qquad b_k = \sum_{i}^{N} \frac{y_i\, F_k(x_i)}{\sigma_i^2}. \tag{2.3.5}$$

Equations (2.3.3) or (2.3.4) form a set of linear equations with respect to the parameters p_j and they can be solved by standard methods of matrix algebra. One of such methods is singular value decomposition (SVD) that efficiently deals with matrix equations and especially in the case, when the matrix is singular or close to singular. This can happen when two basis functions, which are not exactly known *a priori*, reproduce the reference data equally well (or bad).

The SVD method is based on the linear algebra theorem, which states that any $N \times M$ matrix $\mathbf{A}$ with $N \geq M$ can be written as the product of $N \times M$ column-orthogonal matrix $\mathbf{U}$, $M \times M$ diagonal matrix $\mathbf{W}$ with non-negative elements (singular values), and the transpose of an $M \times M$ orthogonal matrix $\mathbf{V}$. This can be written as

$$\mathbf{A} = \mathbf{U} \cdot \mathbf{W} \cdot \mathbf{V}^{\mathrm{T}} \tag{2.3.6}$$

with $\mathbf{U}$ and $\mathbf{V}$ being matrices that satisfy

$$\mathbf{U}\mathbf{U}^{\mathrm{T}} = \mathbf{1}, \quad \mathbf{V}\mathbf{V}^{\mathrm{T}} = \mathbf{1}. \tag{2.3.7}$$

Let us consider a matrix equation

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b}, \tag{2.3.8}$$

where $\mathbf{A}$ is a square $N \times N$ matrix. If vector $\mathbf{b}$ is in the range of matrix $\mathbf{A}$ (i.e. the vector $\mathbf{b}$ is N-dimensional one), then the solution can be constructed as

$$\mathbf{x} = \mathbf{A}^{-1} \cdot \mathbf{b} = \mathbf{V} \cdot \mathbf{W}^{-1} \cdot (\mathbf{U}^{-1} \cdot \mathbf{b}). \tag{2.3.9}$$

The main idea is that if any of the ω_i elements of the matrix $\mathbf{W}$ happens to be zero, then one should set the corresponding element in $\mathbf{W}^{-1}$ to zero too. Equation (2.3.9) then gives the solution (more exactly one of the possible solutions of (2.3.8)). In the case, when the vector $\mathbf{b}$ is not in the range of matrix $\mathbf{A}$, eq. (2.3.9) constructs the vector $\mathbf{x}$, that minimizes $|\mathbf{A} \cdot \mathbf{x} - \mathbf{b}|$.

In a general case, when the number of unknowns is smaller than the number of equations, the matrix $\mathbf{W}$ will not be singular, but some elements ω_i can, however, be very small thus leading to numerical instabilities. To obtain a stable solution, a minimum value is chosen, which is generally close to the computer precision, and the elements ω_i lesser than this threshold are zeroized (and

the corresponding ω_i^{-1} elements too).

The SVD method can also be applied for $N < M$. In this case, one should set the diagonal elements ω_i $(i = N+1, \ldots, M)$ of the matrix $\mathbf{W}$ and the corresponding columns of the matrix $\mathbf{U}$ to zero.

Downhill simplex minimization in multidimensions.

The SVD method is suitable for situations, where the model function is linear with respect to the parameters. If this is not the case, other techniques should be used to minimize the cost function (2.3.1). For this purpose, downhill simplex method, also known as Nelder-Mead method, works well, especially, when the dimensionality of the parameters space is more than one and the computation of cost function derivatives is cumbersome.

A simplex is a closed N-dimensional polyhedron defined with $N+1$ points. To start the minimization, one has to define the initial simplex through its $N+1$ vertices and the values of the cost function in the vertices, i.e. to specify the starting point $\mathbf{P}_0$. The minimization of the cost function then proceeds via moving to a new point according to the scheme

$$\mathbf{P}_i = \mathbf{P}_0 + \lambda \mathbf{e}_i, \quad i = 1, \ldots, N, \tag{2.3.10}$$

where $\mathbf{e}_i$ are N unit vectors along the axes, and λ is a characteristic scale value.

The downhill simplex minimization procedure uses a number of operations that modify the simplex size and form. The possible operations are reflection, contraction, and expansion. At the beginning, a set of reflections is performed: the highest points of the simplex are reflected to the opposite side of the simplex. This kind of geometrical operation saves the simplex volume and its nondegeneracy. If it is possible, the simplex takes bigger steps, that corresponds to the case of expansion. When the method reaches the valley, where the function minimum is located, it slows down by using the contraction operation. The convergence is usually controlled through the "simplex movement speed": if the distance the point $\mathbf{P}$ is moved at the step is smaller than a predefined value of tolerance, then the minimization is terminated.

2.3.3 Minimization methods using gradients.

The use of methods discussed previously necessitates the computation only of the function values $f(x_i, \mathbf{p})$. In many cases, the convergence of minimization procedure can significantly be accelerated, if the derivatives of the function with respect to the vector of parameters $\mathbf{p}$ are used to make a judicious choice of the direction and size of step to move to.

Let us assume that the function value and gradient at the point $\mathbf{p}_0$ are known. Then, the cost function (2.3.1) can be expanded in a Taylor series

$$\chi^2(\mathbf{p}) = \chi^2(\mathbf{p}_0) + (\mathbf{p} - \mathbf{p}_0) \cdot \boldsymbol{\nabla}\chi^2(\mathbf{p}_0) + \frac{1}{2}(\mathbf{p} - \mathbf{p}_0)^{\mathrm{T}} \cdot \mathbf{D} \cdot (\mathbf{p} - \mathbf{p}_0) + \ldots, \tag{2.3.11}$$

where $\mathbf{D}$ is the Hessian matrix of the cost function with elements defined as a second derivatives

of χ^2 with respect to parameters

$$D_{ij} = \left(\frac{\partial^2 \chi^2(\mathbf{p})}{\partial p_i \, \partial p_j} \right)_{\mathbf{p}_0} . \tag{2.3.12}$$

Two different situations can be considered:

1. The vector of parameters $\mathbf{p}$ is far from its optimum value. In this case, a steepest descent algorithm can be used to move the parameters towards the minimum of χ^2 function. Having in hands the gradient of cost function at the point $\mathbf{p}^{(n)}$, one chooses the next point as

$$\mathbf{p}^{(n+1)} = \mathbf{p}^{(n)} - \lambda \boldsymbol{\nabla} \chi^2(\mathbf{p}^{(n)}). \tag{2.3.13}$$

The parameter λ is chosen to be small enough in order to not miss the minimum, but large enough to move sufficiently fast. This is a steepest descent method of function minimization.

2. The vector of parameters is close to the optimum. In this case, the use of Hessian matrix $\mathbf{D}$ allows adapting the value of λ as a function of slope of the χ^2 function and to find the minimum in one single move. The vector of optimized parameters is then sought as

$$\mathbf{p}^{(n+1)} = \mathbf{p}^{(n)} + \mathbf{D}^{-1}(-\boldsymbol{\nabla} \chi^2(\mathbf{p}^{(n)})). \tag{2.3.14}$$

Using the expression for χ^2 function (2.3.1), one can define the gradients and Hessian matrix with respect to the parameter vector as

$$\boldsymbol{\nabla} \chi^2 = -2 \sum_i \frac{\big(y_i - f(x_i, \mathbf{p})\big)}{\sigma_i^2} \, \boldsymbol{\nabla} f(x_i, \mathbf{p}) \tag{2.3.15}$$

and

$$\mathbf{D}_{lm} = \frac{\partial^2 \chi^2}{\partial p_l \, \partial p_m} = 2 \sum_i \frac{1}{\sigma_i^2} \left[\frac{\partial f(x_i, \mathbf{p})}{\partial p_l} \, \frac{\partial f(x_i, \mathbf{p})}{\partial p_m} - \big(y_i - f(x_i, \mathbf{p})\big) \, \frac{\partial^2 f(x_i, \mathbf{p})}{\partial p_l \, \partial p_m} \right]. \tag{2.3.16}$$

As the cost function is already close to the minimum, the second term in the right-hand side of the equation (2.3.16) can be neglected and the expression for the Hessian reads

$$\mathbf{D}_{lm} = 2 \sum_i \frac{1}{\sigma_i^2} \left[\frac{\partial f(x_i, \mathbf{p})}{\partial p_l} \frac{\partial f(x_i, \mathbf{p})}{\partial p_m} \right]. \tag{2.3.17}$$

Omitting the second derivatives of model function in (2.3.16) significantly simplifies the calculations. Equation (2.3.14) is worth to be rewritten in the form

$$\sum_l d_{kl} \, \delta p_l = g_k \tag{2.3.18}$$

where

$$d_{kl} = \frac{1}{2} \, \frac{\partial^2 \chi^2}{\partial p_k \, \partial p_l} \qquad g_l = -\frac{1}{2} \, \frac{\partial \chi^2}{\partial p_k}. \tag{2.3.19}$$

and in this notation the steepest descent method is simply given as

$$\delta p_k = \lambda \times g_k. \tag{2.3.20}$$

Levenberg-Marquardt method.

Marquardt has proposed a method, based on earlier idea of Levenberg, that automatically switches between the steepest descent and Hessian approaches at different stages of the minimization procedure. The first idea of the method concerns the choice of the constant λ in (2.3.20) that has been proposed to be equal to

$$\lambda = \frac{1}{\tilde{\lambda} d_{kk}} \quad \text{leading to} \quad \delta p_k = \frac{1}{\tilde{\lambda} d_{kk}} g_k, \tag{2.3.21}$$

where $\tilde{\lambda}$ is some arbitrary dimensionless scaling factor. Note that with the definition (2.3.17), λ is always positive as it should be.

The second hint is to define an effective Hessian matrix

$$d'_{kl} = d_{kl}(1 + \tilde{\lambda}\delta_{ij}) \tag{2.3.22}$$

and to combine (2.3.18) and (2.3.21) to the unique equation

$$\sum_l d'_{kl}\delta p_l = g_k. \tag{2.3.23}$$

Note that when $\tilde{\lambda}$ is very large the matrix d'_{kl} becomes quasi-diagonal and the method switches to the steepest descent. When $\tilde{\lambda}$ is small (2.3.23) becomes (2.3.18). The Levenberg-Marquardt method is adaptive: it automatically increases the scaling factor $\tilde{\lambda}$, if the new vector of parameters fails to reduce χ^2 value, and reduces the scaling factor otherwise.

Usually, at the beginning, the $\tilde{\lambda}$ value is chosen to be small, say 10^{-3}. If new parameters vector $\mathbf{p} + \delta\mathbf{p}$ yields a bigger χ^2 value, the scaling factor is increased by an order of magnitude, and the minimization returns to the previous step. If $\chi^2(\mathbf{p} + \delta\mathbf{p}) < \chi^2(\mathbf{p})$, then the step is accepted, the scaling parameter is decreased by an order of magnitude, and the vector $\mathbf{p}$ is updated $\mathbf{p} \leftarrow \mathbf{p} + \delta\mathbf{p}$.

Levenberg-Marquardt minimization is terminated when at least one of the following conditions is satisfied:

- magnitude of the gradient $\delta\mathbf{p}$ is less than a predefined treshold value,
- the change of cost function is below a certain threshold,
- maximum number of iterations is exceeded.

The method works very well in most practical applications and has become *de facto* standard technique for nonlinear least-squares minimization.

Optimization of the vectors of parameters in the studies presented in subsequent chapters was done using all families of the methods discussed above. These methods were coded into computer programs with key subroutines borrowed from Ref. [49].

Chapter 3

Models of effective potentials in molecular modeling.

3.1 Types of effective potential model.

As it was discussed above, the main idea of effective potential model is to approximate the potential energy surface with a set of analytical functions. Such an approach greatly simplifies and accelerates the computational studies of large complex systems as compared to quantum-chemical calculations. To tackle the problem the energy $E(\{\mathbf{r}\})$ of a system is written in the form

$$E(\{\mathbf{r}\}) \approx \sum_i E_1(\mathbf{r}_i) + \frac{1}{2}\sum_{i,j} E_2(\mathbf{r}_i, \mathbf{r}_j) + \frac{1}{6}\sum_{i,j,k} E_3(\mathbf{r}_i, \mathbf{r}_j, \mathbf{r}_k) + \ldots, \tag{3.1.1}$$

where $E_n(\{\mathbf{r}\})$ ($n = 1, 2, \ldots$) denotes some analytical functions depending on coordinates of one atom ($n = 1$), two-atoms ($n = 2$), etc. These functions are the effective potentials. The first term in (3.1.1) depending on the coordinates of atoms in the space is generally omitted since the term describes the energy of the system in an external field, but not the energy of interatomic interactions. The functions E_2 account for the pairwise interactions, while terms with $n > 2$ correspond to many-body contributions to the energy.

Despite the fact that the functions E_n can, in principle, have any analytical form, it is, however, desirable that their form relies on physical grounds. The use of analytical functions based on theoretical considerations can make the design of a force field simpler and yields the model more transferable. There exist two large families of the effective potential models: ionic-type force fields and molecular mechanics force fields.

3.1.1 Ionic force fields.

The ionic force fields generally make use of the pair potential functions to describe the energy of interatomic interactions. Most of force fields for the ionic and (semi-) ionic materials employ the functional form suggested by Born [50], who has shown that a surprisingly correct description of

energetical characteristics of ionic solids can be obtained by considering these systems as consisting of an arrangement of positive and negative ions. In such a description the most part of the cohesive energy arises from the electrostatic interactions between the ions described by Coulomb law

$$E_{el} = \frac{1}{2}\sum_{i}^{N}\sum_{j\neq i}^{N}\frac{Q_i\,Q_j}{r_{ij}}, \tag{3.1.2}$$

where E_{el} is the energy of electrostatic interactions, Q_i and Q_j are charges of atoms i and j, respectively, separated by distance r_{ij}.

However, according to the Earnshaw theorem [51], a system of charges cannot be stabilized solely by electrostatic interactions and thus, other than electrostatic forces are to be present in the system. A finite compressibility of solids and liquids points to the appearance of repulsive forces, when the interatomic distance decreases. Consequently, the force field has to include an energy term that accounts for the repulsion between atoms at short distances. Born and Mayer [50] suggested that such a repulsive interaction potential should follow the exponential law and therefore, the repulsion energy contribution E_{rep} to the potential energy of a system reads

$$E_{rep} = \frac{1}{2}\sum_{i}^{N}\sum_{j\neq i}^{N}A_{ij}\,e^{-b_{ij}r_{ij}}, \tag{3.1.3}$$

with A_{ij} and b_{ij} being parameters of the potential.

Finally, another energy term has also to be taken into consideration. Thus, if one considers a rare gas crystal, the stability of such a system cannot be explained by the sum of energies (3.1.2) and (3.1.3) as the atoms possess no charges. Therefore, an attractive interaction that counterbalances the repulsion (3.1.3) should exist between the atoms. These attractive forces arise from the interactions of instantaneous multipole moments (dipoles, quadrupoles, etc.) of the electron distributions in atoms and the interactions are referred to as dispersion interactions. The second-order perturbation theory provides the functional form for the dispersion energy E_{disp}, assuming the spherical symmetry of electron distribution, as

$$E_{disp} = -\frac{1}{2}\sum_{i}^{N}\sum_{j\neq i}^{N}\Big(\frac{C_{6\,ij}}{r_{ij}^{6}} + \frac{C_{8\,ij}}{r_{ij}^{8}} + \ldots\Big), \tag{3.1.4}$$

where the parameters $C_{6\,ij}$, $C_{8\,ij}$, etc. are called dispersion constants. The sum in (3.1.4) is often cut out after the first term that accounts for the dispersion energy due to the dipole-dipole interactions.

The potential energy of interaction between two ions $E(r_{ij})$ can therefore be written as a sum of three contributions

$$E(r_{ij}) = A_{ij}\,e^{-b_{ij}r_{ij}} - \frac{C_{6\,ij}}{r_{ij}^{6}} - \frac{C_{8\,ij}}{r_{ij}^{8}} + \frac{Q_i\,Q_j}{r_{ij}} \tag{3.1.5}$$

and this potential form is known as Born-Mayer-Huggins potential. The potential is often simpli-

fied by omitting the $1/r^8$ term that results in

$$E(r_{ij}) = A_{ij}\, e^{-b_{ij} r_{ij}} - \frac{C_{6\,ij}}{r_{ij}^6} + \frac{Q_i\, Q_j}{r_{ij}}. \tag{3.1.6}$$

The potential model given by the functional form (3.1.6) is known as rigid ion potential. The rigid ion model (RIM) has been (and still is) widely used for studying inorganic materials such as oxides and salts. For example, one of the most successful force fields for silica developed by van Beest, Kramer, and van Santen [52] (BKS model) has the RIM form (3.1.6).

3.1.2 Molecular mechanics force fields.

Molecular mechanics (MM) force fields attempt to mimic both two-body and many-body interactions by recasting (3.1.1) them into the form

$$E = \sum_b E_b + \sum_a E_a + \sum_\tau E_\tau + \sum_{vdw} E_{vdw} + \sum_{el} E_{el}, \tag{3.1.7}$$

where E_b, E_a, E_τ, E_{vdw}, and E_{el} stand for the energy contributions due to the valence bonds, valence angles, dihedral angles, and nonbonded van der Waals (vdW) and electrostatic interactions, respectively. One sees that, in contrast to the ionic force field models, the MM ones explicitly use the concept of "chemical bond". The first three terms in (3.1.7) are said to account for the energy of bonded interactions, E_B. In the simplest form these three terms are given by

$$E_b = \frac{1}{2} K_b (r - r_0)^2, \tag{3.1.8}$$

$$E_a = \frac{1}{2} K_\theta (\theta - \theta_0)^2, \tag{3.1.9}$$

$$E_\tau = \sum_n A_n \big[1 + \cos(n\tau - \tau_0)\big], \tag{3.1.10}$$

where r, θ, and τ are the length of bond, valence angle, and dihedral angle, respectively, and the zero index denotes the equilibrium value (τ_0 is often referenced to as a phase factor). Parameters K_b, K_θ, and A_n are commonly called force constants. Equations (3.1.8) and (3.1.9) results from the use of harmonic approximation and they are therefore limited to small amplitudes. To allow more flexibility, the harmonic potential functions in (3.1.8) and (3.1.9) are often replaced by anharmonic ones, such as Morse potential for the bond-stretching coordinate r. Nevertheless, regardless the form of bond and angle-bending potentials, the common feature of bonded energy is that the bond-stretching (3.1.8) and angle-bending (3.1.9) potentials keep the atoms in functional groups together, e.g. in CH_3 group, while conformational changes (large relative displacements of the groups) are allowed for by the variation of the dihedral angles τ.

The last two terms in (3.1.7) account for the energy of nonbonded interactions E_{NB} and are

represented by the sum of electrostatic and van der Waals energy terms

$$E_{NB} = \frac{1}{2}\sum_{i,j}\left[\frac{Q_i\,Q_j}{r_{ij}} + 4\varepsilon_{ij}\Big(\frac{\sigma_{ij}}{r_{ij}}\Big)^{12} - \Big(\frac{\sigma_{ij}}{r_{ij}}\Big)^{6}\right] \tag{3.1.11}$$

with Q_i, ε_{ij}, and σ_{ij} being the atomic charge and the parameters of Lennard-Jones (12-6) potential, respectively. The sum in (3.1.11) generally omits the atoms involved in the bonds, bond angles, and sometimes in the dihedral angles, because the force constants of the bonded potentials (3.1.8)-(3.1.10) implicitely include the energy of electrostatic and vdW interaction between the atoms forming these internal coordinates.

Since the beginning, the MM force fields have been developed for and applied to organic molecules from simple alkanes to biological macromolecules such as proteins, enzymes, etc. These force fields also have had a success in modelling inorganic structures with covalent or partially covalent bonds. The explicit use of the bonding information, however, limits the use of MM force fields for studying the phenomena, where the bonding situation of atoms may change, e.g. phase transitions.

The choice of the functional form and of parameters of the potential functions for both the ionic and MM force fields is a complex task. In most force fields, once being determined, the parameters are then kept fixed and considered to be transferable between similar systems regardless the structure and/or environment of atoms. The assumption of transferability results in increasing the uncertainty of results of the calculations using the force field and limits their predictive power. Thus, the constancy of charges Q_i in (3.1.6) and (3.1.11) points to the inability of model to take into account the variations of electron distribution in system due to a perturbation, such as displacement of atoms from their equilibrium positions, creation of a defect, or substitution one atom by another. The issue is of special importance for the electrostatic energy (3.1.2) because of its magnitude and the long-range behaviour. However, strictly speaking, the consideration is not limited to the charge-charge interactions as a change of electron density distribution will also lead to a change of energy contributions due to higher multipoles of the multipole expansion of electrostatic energy. In most cases, these energy changes are averaged and "buried" in the parameters of effective potentials. Non-polarizable water models provide a typical example of this situation, where the increase of effective charge of O atom is necessary to mimic the polarization of molecules in condensed phase, once and for all.

Therefore, one can hope that the inclusion of the electronic polarization into a force field will result in a significant improvement of the transferability and predictive power of the effective potential model, although at the expense of increased complexity. The development of such polarizable force fields nowadays plays a central role for large-scale modeling of inorganic materials, (bio-) organic systems such as nucleid acids, peptides, etc. [53, 54]. Furthermore, one can also expect that the polarizable force field models explicitely including the electronic polarization effect can yield those system characteristics, which were unaccessable with conventional non-polarizable models [55].

3.2 Introduction of polarization effect in force field.

The polarization effect can be described as a reorganization of multipole moments of electron density distribution of a system upon a perturbation. Consequently, a polarizable force field model attempts to mimic such a reorganization. The most widely used approaches for including the polarization effects in a force field are

- induced point dipoles model,
- shell model or drude oscillators model,
- deformable and polarizable ions model, and
- chemical potential equalization model.

All these models are considered below in detail.

3.2.1 Induced point dipoles model.

The model adds point inducible dipoles on all or some of the atomic sites. The induced dipole moment $\boldsymbol{\mu}_i$ at the site i is proportional to the local electric field $\boldsymbol{\mathcal{E}}_i$ at that site with the proportionality coefficient being the polarizability tensor $\boldsymbol{\alpha}_i$ of the site. Each dipole feels the field resulting from both permanent charges and other induced dipoles. Therefore, the expression for the dipole $\boldsymbol{\mu}_i$ reads

$$\boldsymbol{\mu}_i = \boldsymbol{\alpha}_i \boldsymbol{\mathcal{E}}_i = \boldsymbol{\alpha}_i \Big[\boldsymbol{\mathcal{E}}_i^0 - \sum_{j \neq i} \mathbf{T}_{ij} \boldsymbol{\mu}_j\Big], \tag{3.2.1}$$

with $\boldsymbol{\mathcal{E}}_i^0$ being the electric field of permanent charges. The induced point dipoles interact through the dipole field tensor $\mathbf{T}_{ij}$

$$\mathbf{T}_{ij} = \frac{1}{r_{ij}^3}\mathbf{I} - \frac{3}{r_{ij}^5}\begin{pmatrix} x_{ij}^2 & x_{ij}y_{ij} & x_{ij}z_{ij} \\ y_{ij}x_{ij} & y_{ij}^2 & y_{ij}z_{ij} \\ z_{ij}x_{ij} & z_{ij}y_{ij} & z_{ij}^2 \end{pmatrix}, \tag{3.2.2}$$

where $\mathbf{I}$ is the identity matrix, r_{ij} is the distance between sites i and j, and x_{ij}, y_{ij} and z_{ij} are the Cartesian components of the vector $\mathbf{r}_{ij}$.

The energy $E_{\mu\mu}$ of the system of N interacting induced dipoles can be obtained as

$$E_{\mu\mu} = -\sum_{i=1}^{N} \boldsymbol{\mu}_i \boldsymbol{\mathcal{E}}_i + \frac{1}{2}\sum_{i=1}^{N}\sum_{j \neq i} \boldsymbol{\mu}_i \mathbf{T}_{ij} \boldsymbol{\mu}_j + \frac{1}{2}\sum_{i=1}^{N} \boldsymbol{\mu}_i \boldsymbol{\alpha}_i^{-1} \boldsymbol{\mu}_i, \tag{3.2.3}$$

where the first term corresponds the energy of dipoles in their local electric fields, the second term corrects the first one for the double counting of the dipole-dipole interactions, and the third term is the polarization energy, i. e. the energy required to create the induced dipoles [56].

In the most general case all the dipoles $\boldsymbol{\mu}_i$ interact through the dipole field tensor (3.2.2). However, the use of the expression (3.2.2) can cause the so-called "polarization catastrophe".

As it was pointed out by Applequist *et al.* [57] and Thole [58], the molecular polarization and, therefore, the induced dipole moments, may become infinite at small distances. Let us consider a diatomic molecule AB with atoms having the isotropic polarizabilities $\alpha_{\rm A}$ and $\alpha_{\rm B}$. In the frame of the induced dipoles model the molecular polarizability tensor components parallel, $\alpha_{\parallel}$, and perpendicular, $\alpha_{\perp}$, to the vector $\mathbf{r}_{\rm AB}$, respectively, can be written as [57, 58]

$$\begin{aligned} \alpha_{\parallel} &= [\alpha_{\rm A} + \alpha_{\rm B} + (4\alpha_{\rm A}\alpha_{\rm B}/r_{\rm AB}^3)]/[1 - (4\alpha_{\rm A}\alpha_{\rm B}/r_{\rm AB}^6)], \\ \alpha_{\perp} &= [\alpha_{\rm A} + \alpha_{\rm B} - (2\alpha_{\rm A}\alpha_{\rm B}/r_{\rm AB}^3)]/[1 - (\alpha_{\rm A}\alpha_{\rm B}/r_{\rm AB}^6)]. \end{aligned} \tag{3.2.4}$$

From eqs. (3.2.4) one sees that the $\alpha_{\parallel}$ and $\alpha_{\perp}$ values become infinite, when the distance between two atoms approaches $(4\alpha_{\rm A}\alpha_{\rm B})^{1/6}$ or $(\alpha_{\rm A}\alpha_{\rm B})^{1/6}$, respectively. Thus, the "polarization catastrophe" distance for O_2 molecule with $\alpha_{\rm O} = 0.77$ Å^3 is equal to 1.15 Å, that is close to the equilibrium interatomic distance of 1.21 Å.

There are two ways to avoid these singularities. The first one is to choose the polarizabilities sufficiently small such that the factor $4\alpha_{\rm A}\alpha_{\rm B}/r^6$ is always less than one at the typical distances between atoms. Unfortunately, such an approach does not remove the singularities, but only shifts them to smaller distances. Furthermore, the atomic polarizabilities $\boldsymbol{\alpha}_i$ are the fitting parameters of the induced dipole model and thus, a freedom in the choice of their values is rather limited.

The second way is to introduce a screening of the dipole-dipole interactions at small distances [58]. The reason behind such a correction is that the electronic distribution is not well represented by point dipoles tensor (3.2.2) at small distances [57–59]. Thole [58] introduced a scaling distance $s_{ij} = 1.662\,(\alpha_i\alpha_j)^{1/6}$, which leads to the change of the dipole field tensor (3.2.2) for distances $r_{ij} < s_{ij}$ as following

$$\mathbf{T}_{ij} = (4\upsilon^3 - 3\upsilon^4)\frac{1}{r_{ij}^3}\mathbf{I} - \upsilon^4\frac{3}{r_{ij}^5}\begin{pmatrix} x_{ij}^2 & x_{ij}y_{ij} & x_{ij}z_{ij} \\ y_{ij}x_{ij} & y_{ij}^2 & y_{ij}z_{ij} \\ z_{ij}x_{ij} & z_{ij}y_{ij} & z_{ij}^2 \end{pmatrix}, \tag{3.2.5}$$

where $\upsilon = r_{ij}/s_{ij}$, while $\mathbf{T}_{ij}$ is unchanged if the distance r_{ij} is greater than s_{ij}.

The computation of the dipoles $\boldsymbol{\mu}_i$ is commonly performed with the use of three methods. The first one is based on the observation that eq. (3.2.1) can be rewritten in the form of matrix equation as

$$\mathbf{AM} = \boldsymbol{\mathcal{E}}^0, \tag{3.2.6}$$

where $\mathbf{M}$ and $\boldsymbol{\mathcal{E}}^0$ are the $3N$ vectors of Cartesian components of dipoles $\boldsymbol{\mu}_i$ and of the components of electric fields $\boldsymbol{\mathcal{E}}_i^0$, respectively, and the $3N \times 3N$ matrix $\mathbf{A}$ is constructed of N 3×3 blocks. The diagonal blocks are $\boldsymbol{\alpha}_i^{-1}$ and the off-diagonal blocks are built from the $\mathbf{T}_{ij}$ dipole interaction tensors (3.2.2). Note that (3.2.6) is equivalent to the condition that the dipoles $\boldsymbol{\mu}_i$ minimize the energy $E_{\mu\mu}$ (3.2.3), i.e.

$$\frac{\partial E_{\mu\mu}}{\partial \boldsymbol{\mu}_i} = 0, \quad i = 1, \ldots, N. \tag{3.2.7}$$

The solution of (3.2.6) for the system of N point dipoles then involves the inversion of the matrix $\mathbf{A}$.

The second way of computation of dipoles consists of using an iterative procedure since each dipole feels the field of all other induced dipoles. In the iterative methods, the initial guess is made for the total field. Usually, it is taken to be the field of permanent multipoles $\boldsymbol{\mathcal{E}}^0$. The dipole moments are obtained from the eq. (3.2.1), and the iterations are performed out until the convergence is reached. This method is so far most widely used for computing the dipoles.

In molecular dynamics simulations employing force field extended with the induced dipoles model the values of induced dipoles can also be derived with the help of a predictive scheme. For instance, in the first-order prediction algorithm used by Ahlström *et al.* [60] the dipole value at the time t is determined from the values known at two previous time steps

$$\boldsymbol{\mu}_i(t) = 2\boldsymbol{\mu}_i(t - \Delta t) - \boldsymbol{\mu}_i(t - 2\Delta t), \tag{3.2.8}$$

where Δt is the integration time-step. The algorithm, however, is not stable for long MD trajectories and one can combine it with the iterative method to overcome the instabilities [60, 61].

Another scheme uses the extended Lagrangian approach in the spirit of the Car-Parrinello method [62–65]. The dipoles $\boldsymbol{\mu}$ are treated as dynamical variables characterized by a fictitious mass m_μ and the momentum $\mathbf{p}_\mu = m_\mu \dot{\boldsymbol{\mu}}$. Making use of extended Lagrangian, one can obtain the equation of motion for the dipole variable $\boldsymbol{\mu}_i$ in the form

$$m_\mu \ddot{\boldsymbol{\mu}}_i = -\frac{\partial E_{\mu\mu}}{\partial \boldsymbol{\mu}_i} = \boldsymbol{\mathcal{E}}_i - \boldsymbol{\alpha}_i^{-1}\boldsymbol{\mu}_i. \tag{3.2.9}$$

The value of m_μ parameter is usually chosen to be small in order to allow the dipoles to easily follow the dynamics of atoms (Born-Oppenheimer principle). It is also important to keep the fictitious kinetic energy of dipoles $\mathbf{p}_\mu^2/2m_\mu$ small enough such that "cold" dipole degrees of freedom match the exact solution of eq. (3.2.7) as close as possible. To make sure that the parameter is chosen properly, the trajectories obtained with the extended Lagrangian method can be compared to those yielded with the use of the iterative scheme [64].

Recently, Ponder and co-workers [66–71] developed AMOEBA force field based on a modification of the original formulations of the model given by Applequist [57] and Thole [58]. The equation (3.2.1) is modified to include the static electric field not only from permanent charges, but also from permanent multipoles

$$\boldsymbol{\mu}_i = \boldsymbol{\alpha}_i\, \boldsymbol{\mathcal{E}}_i = \boldsymbol{\alpha}_i \cdot \Big[\sum_{j\neq i} \hat{\mathbf{T}}_{ij}\mathbf{M}_j - \sum_{j\neq i} \mathbf{T}_{ij}\boldsymbol{\mu}_j\Big], \tag{3.2.10}$$

where $\mathbf{M}_i = (Q_i, \mu_{xi}, \mu_{yi}, \mu_{zi}, \Theta_{xxi}, \Theta_{xyi}, \ldots, \Theta_{zzi})$ is the vector of atomic permanent multipole components up to quadrupolar moment in the Cartesian polytensor form [72]. $\hat{\mathbf{T}}$ is the interaction matrix.

Among other important extensions of the model, one should mention approaches introducing

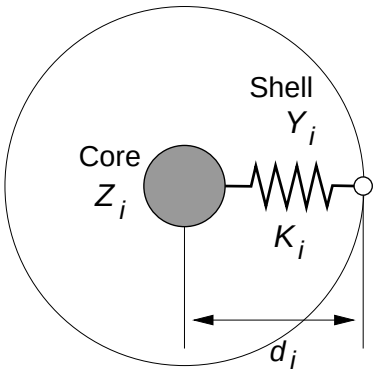

Figure 3.1: Schematic representation of atom i in the shell model. The shell particle with negative charge Y_i is attached to the core having positive charge Z_i by a harmonic spring with spring constant K_i. The core-shell separation distance is d_i and net atomic charge Q_i is equal to $Q_i = Z_i + Y_i$.

the intramolecular charge-transfer, which is modelled either with atomic capacitances [73–75] or atomic electronegativities [76–79]. The induced point dipoles model (3.2.1) was also extended for hyperpolarizabilities [80–83], optical rotation [84–87], Raman intensities [88–93], and absorption [94].

The induced point dipole model has been applied for studying polarizabilities of molecular systems including carbon nanotubes [95, 96] and fullerens [96–98], boron nitride nanotubes [99], proteins [100] and molecular clusters [101, 102]. The model was also used in the development of a large number of polarizable potentials for water [60, 65, 68, 69, 103–117].

Additionally to already mentioned AMOEBA force field, an implementation of the point induced dipole model into CHARMM force field was reported in Ref. [118]. The description for nonpolarizable part of the potential energy was based on intermolecular potential functions model [119, 120].

3.2.2 Shell model.

The shell model proposed by Dick and Overhauser [121] describes each atom as a pair of point charges, separated by a distance and attached to each other by a harmonic spring as schematically shown in Figure 3.1. The first charge is the core that generally bears a positive charge and whose position coincides with that of the nucleus. The second particle (shell) represents valence electrons of the atom and is charged negatively; the shell is generally considered to be massless. Magnitudes of the core and shell charges are usually treated as adjustable parameters of the model with the constraint that their sum is equal to the net charge of the atom.

For an atom i with a positive core charge Z_i and a negative shell charge Y_i, that are separated by core-shell distance $\mathbf{d}_i$ and hold together by a spring of strength K_i, the atom dipole is given by

$$\boldsymbol{\mu}_i = Y_i \mathbf{d}_i. \tag{3.2.11}$$

The energy $E_{\rm SM}$ of the system in the shell model consists of three terms. These are the interaction energy $E_{\rm stat}$ of atoms with an external static field $\boldsymbol{\mathcal{E}}_{\rm stat}$, electrostatic interaction energy E_{el} between atoms, and the polarization energy $E_{\rm pol}$

$$E_{\rm SM} = E_{\rm stat} + E_{el} + E_{\rm pol}. \tag{3.2.12}$$

The polarization energy accounts for the spatial separation of the core and shell point charges and is equal to

$$E_{\rm pol} = \frac{1}{2}\sum_{i=1}^{N} K_i\,\mathbf{d}_i^2. \tag{3.2.13}$$

Note that by using (3.2.11) and (3.2.13) one easily obtains the electronic polarizability α_i of atom in the shell model

$$\alpha_i = \frac{Y_i^2}{K_i} \tag{3.2.14}$$

that allows a judicious choice of the spring strength K_i.

The electrostatic energy E_{el} is written as a sum of four contributions arising from the core-core, core-shell, and shell-shell interactions

$$E_{el} = \frac{1}{2}\sum_{i=1}^{N}\sum_{j\neq i}\Big(\frac{Z_iZ_j}{|\mathbf{r}_{ij}|} - \frac{Z_iY_j}{|\mathbf{r}_{ij}+\mathbf{d}_i|} - \frac{Z_jY_i}{|\mathbf{r}_{ij}-\mathbf{d}_j|} + \frac{Y_iY_j}{|\mathbf{r}_{ij}+\mathbf{d}_i-\mathbf{d}_j|}\Big), \tag{3.2.15}$$

where the Coulombic interactions between the core and shell of the same atom are excluded. The interaction energy $E_{\rm stat}$ of the atomic dipoles with a static electric field is a sum of individual point charge interactions

$$E_{\rm stat} = -\sum_{i=1}^{N}\Big(Z_i\mathbf{r}_i\,\boldsymbol{\mathcal{E}}_i^0 + Y_i(\mathbf{r}_i+\mathbf{d}_i)\,\boldsymbol{\mathcal{E}}_i^{0\prime}\Big), \tag{3.2.16}$$

where $\boldsymbol{\mathcal{E}}_i^0$ and $\boldsymbol{\mathcal{E}}_i^{0\prime}$ is the external electric field at the core ($\mathbf{r}_i$) and shell ($\mathbf{r}_i+\mathbf{d}_i$) positions, respectively, and $\boldsymbol{\mathcal{E}}_i^0 \neq \boldsymbol{\mathcal{E}}_i^{0\prime}$ in a general case.

Therefore, the energy (3.2.12) can be written as

$$\begin{aligned} E_{\rm SM} &= \sum_{i=1}^{N}\Big(\frac{1}{2}K_i\,\mathbf{d}_i^2 + \big(Z_i\mathbf{r}_i\,\boldsymbol{\mathcal{E}}_i^0 + Y_i(\mathbf{r}_i+\mathbf{d}_i)\,\boldsymbol{\mathcal{E}}_i^{0\prime}\big)\Big) \\ &+ \frac{1}{2}\sum_{i=1}^{N}\sum_{j\neq i}\Big(\frac{Z_iZ_j}{|\mathbf{r}_{ij}|} - \frac{Z_iY_j}{|\mathbf{r}_{ij}+\mathbf{d}_i|} - \frac{Z_jY_i}{|\mathbf{r}_{ij}-\mathbf{d}_j|} + \frac{Y_iY_j}{|\mathbf{r}_{ij}+\mathbf{d}_i-\mathbf{d}_j|}\Big) \end{aligned} \tag{3.2.17}$$

and the energy (3.2.17) should be minimized with respect to the core-shell vectors $\mathbf{d}_i$. The problem is, however, complicated by the fact that the right-hand part of (3.2.17) is not a quadratic function of $\mathbf{d}_i$ and therefore, the matrix methods cannot be applied to solve the energy minimization problem and the iterative methods are used instead.

The original shell model can easily be implemented with the energy minimization technique, while its use in the molecular dynamics simulations is hampered by the fact that the shells are massless. Mitchell and Fincham [122] proposed the dynamical approach for finding the shell displacements $\mathbf{d}_i$ in MD simulations. A small mass (around 10% of the total ion mass) is assigned to the shell and the equations of motion for both cores and shells are derived from an extended Lagrangian. If the portion of the mass which is given to a shell is x_i of the total ion mass M_i, then the core-shell oscillation frequency ω_i is equal to

$$\omega_i = \sqrt{\frac{K_i}{x_i(1-x_i)M_i}}. \tag{3.2.18}$$

The frequency ω_i has to be well above the high-frequency vibration of the system to avoid coupling between the dynamics of nuclei and that of shells. If this condition is fulfilled, the behaviour of the system is independent of the mass factor x_i [122]. In practice, x_i is chosen by taking two conditions into consideration. First, the shells have to be light enough to easily follow the dynamics of nuclei (the Born-Oppenheimer principle). Second, the shells have to be heavy enough to allow the use of reasonably large time-step while integrating the equations of motion.

There exist several extensions of the basic shell model. One of such a modification [123, 124] uses anisotropic oscillators between core and shell thus allowing the description of non-spherical ions [123, 124]. Among other extensions there are "breathing shell" model [125–127] and deformable shell model [128–130]. The shell model and its extensions have been successfully used in numerous modeling studies of variety of systems. The majority of applications concerned (semi-)ionic materials [122,131,132], particularly oxides [133–152] and alkali halides [128,153–155].

3.2.3 Deformable and polarizable ions model.

The model of deformable and polarizable ions proposed by Ivanov and Maksimov [156,157] can be viewed as a generalization of Gordon and Kim model [158] with the dipole polarizability of ions taken into account. In this model an ionic crystal is treated as consisting of spherically symmetric overlapping ions. The electron density of atom i in the crystal is given by

$$\rho_i(\mathbf{r}) = \rho_i^0(\mathbf{r}) + \delta\rho_i(\mathbf{r}), \tag{3.2.19}$$

where $\rho_i^0(\mathbf{r})$ is the density of a spherical ion and $\delta\rho_i(\mathbf{r})$ is the first-order change of the density induced by a field $\boldsymbol{\mathcal{E}}$. The total electron density $\rho(\mathbf{r})$ of the system is then written as

$$\rho(\mathbf{r}) = \sum_i \rho_i(\mathbf{r} - \mathbf{R}_i), \tag{3.2.20}$$

where $\mathbf{R}_i$ is the position of ion i.

The total energy of the crystal E_{CR} can be represented in the form [156, 158]

$$E_{\mathrm{CR}} = E[\sum_i \rho_i(\mathbf{r})] - \sum_i E[\rho_i(\mathbf{r})] + E^N, \tag{3.2.21}$$

where $E[\rho(\mathbf{r})]$ is the electron density functional and E_N is the nuclei-nuclei interaction energy. The latter quantity is given by

$$E_N = \frac{1}{2}\sum_{i,j} \frac{Z_i^N Z_j^N}{|\mathbf{R}_i - \mathbf{R}_j|}, \tag{3.2.22}$$

with Z_i^N and Z_j^N being the nuclear charge of the ions i and j, respectively. The functional $E[\rho(\mathbf{r})]$ can be written as [3]

$$E[\rho(\mathbf{r})] = \int d\mathbf{r}\rho(\mathbf{r})V_{\mathrm{ext}}(\mathbf{r}) + \frac{1}{2}\int d\mathbf{r}\int d\mathbf{r}'\frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + \int d\mathbf{r}F[\rho(\mathbf{r})], \tag{3.2.23}$$

where $V_{\mathrm{ext}}(\mathbf{r})$ is an external potential and $F[\rho(\mathbf{r})]$ is a universal functional of the electron density, that accounts for the contributions of kinetic energy $T[\rho]$ and exchange-correlation energy $E_{\mathrm{xc}}[\rho]$ of electrons [158]. In a crystal the external potential $V_{\mathrm{ext}}(\mathbf{r})$) is simply the Coulombic potential of the nuclei

$$V_{\mathrm{ext}}(\mathbf{r}) = \sum_i \frac{Z_i^N}{|\mathbf{r}-\mathbf{R}_i|}, \tag{3.2.24}$$

while the explicit form of the universal functional $F[\rho(\mathbf{r})]$ is generally unknown [4].

After substitution of (3.2.19) into (3.2.23) one obtains the expression for the total energy of the crystal in the form [156]

$$\begin{aligned} E_{\mathrm{CR}} &= \frac{1}{2}\sum_{ij} Z_{\mathrm{ion}}^i U_i + \sum_{ij} V_{ij} + \sum_i S_i(U_i) \\ &+ \frac{1}{2}\sum_i \frac{\boldsymbol{\mu}_i^2}{\alpha_i} + \sum_{ij}\boldsymbol{\mu}_i\mathbf{T}_{ij}\boldsymbol{\mu}_j + \sum_{ij}\boldsymbol{\mu}_i\boldsymbol{\Gamma}_{ij}\boldsymbol{\mu}_i - \sum_{ij}\boldsymbol{\mu}_i\mathbf{S}_{ij}Z_{\mathrm{ion}}^j - \sum_i \boldsymbol{\mu}_i\mathbf{F}_i^M, \end{aligned} \tag{3.2.25}$$

where $\boldsymbol{\mu}_i$ and α_i are the dipole moment and the polarizability of ion i, respectively [157, 159].

The first term in (3.2.25) is the Madelung energy of ions, where the Madelung potential U_i is defined as

$$U_i = \sum_j \frac{Z_{\mathrm{ion}}^j}{|\mathbf{R}_i - \mathbf{R}_j|}. \tag{3.2.26}$$

The second term accounts for the short-range interaction between spherically symmetric ions and can be derived [160] following the work of Gordon and Kim [158], which uses the Thomas-Fermi model. The third sum in (3.2.25) is the contribution from the ion self-energies that depend on the Madelung potential (3.2.26) and can be calculated by integration of the exchange-correlation and kinetic energy densities. Boyer *et al.* [160] used the local-density exchange-correlation functional of Hedin and Lundqvist [161] with self-interaction corrections included explicitly [11].

The remaining energy terms in (3.2.25) represent the dipole-dipole part of the energy. The

fourth term accounts for the polarization energy, i.e. energy needed to create the dipoles $\boldsymbol{\mu}_i$ with polarizabilities α_i. The fifth and sixth terms represent the long-range and the short-range dipole-dipole interactions, respectively. The matrix $\mathbf{T}_{ij}$ is the dipole-dipole interaction tensor (3.2.2), whereas $\boldsymbol{\Gamma}_{ij}$ can be obtained from the exact dipole-dipole interaction tensor. The last two terms in (3.2.25) are the interaction energies between the electrostatic monopoles and dipoles. The next to the last term depending on the vector $\mathbf{S}_{ij}$

$$\mathbf{S}_{ij} = \mathbf{m}_{ij}\mathbf{u}_j \tag{3.2.27}$$

defines the short-range part of the energy, where the matrix $\mathbf{m}_{ij}$ is the deformability matrix, which can be calculated numerically [157], and the vector $\mathbf{u}_j$ denotes a small displacement of the ion j from its equilibrium position. Finally, the last term containing the Madelung field $\mathbf{F}_i^M$

$$\mathbf{F}_i^M = -\sum_j T_{ij} Z_{\text{ion}}^j \mathbf{u}_j \tag{3.2.28}$$

corresponds to the long-range contribution of the monopole-dipole energy.

Making use of the adiabatic condition

$$\frac{\partial E^{\text{CR}}}{\partial \boldsymbol{\mu}_i} = 0, \tag{3.2.29}$$

one obtains the equations for dipole moments $\boldsymbol{\mu}_i$

$$\boldsymbol{\mu}_i = \alpha_i \Big(\mathbf{F}_i^M - \sum_j (\mathbf{T}_{ij} + \boldsymbol{\Gamma}_{ij}) \boldsymbol{\mu}_j + \sum_j \mathbf{m}_{ij}\mathbf{u}_j \Big) \tag{3.2.30}$$

By introducing the local field as

$$\mathbf{F}_i^{\text{loc}} = \mathbf{F}_i^M - \sum_j (\mathbf{T}_{ij} + \boldsymbol{\Gamma}_{ij}) \boldsymbol{\mu}_j \tag{3.2.31}$$

one obtains from (3.2.30)

$$\boldsymbol{\mu}_i = \alpha_i \Big(\mathbf{F}_i^{\text{loc}} + \sum_j \mathbf{m}_{ij}\mathbf{u}_j \Big) \tag{3.2.32}$$

The system of equations (3.2.32) can be solved iteratively using the initial guess for the local field $\mathbf{F}_i^{\text{loc}}$ in the form of $\mathbf{F}_i^M$.

In their original paper Ivanov and Maksimov [156] calculated the transverse and longitudinal phonon frequencies for several crystals with cubic symmetry of lattice. In addition, the effective Born charges determining the infrared intensities have also been obtained for the same structures. The comparison of the results given by the model of Ivanov and Maksimov with the results obtained from the potential-induced breathing (PIB) model [160] shows significant improvements in the characteristics calculation accuracy.

All three aforementioned models introduce the polarization effect via induced dipoles. Nevertheless, there are several differences within the ideas and methods these models are based on. The main difference is at the level of treating the polarization. The shell model in its original formulation represents a mechanical analogy, whereas point dipoles and polarizable and deformable ions models use a more physically established concepts for the representation of polarization effects. It should also be mentioned that the first two models are based on parameters which are supposed to be transferable, that allows one to calculate and validate them only once. On the other hand, in the deformable and polarizable ions model one has first to perform the DFT calculations to obtain all parameters for each system of interest. Such an *ad hoc* parameterization is then followed by the computation of characteristics of interest.

3.2.4 Chemical potential equalization model.

The models presented so far concerned the atomic polarization mimicked via induced dipoles. The chemical potential equalization model (CPE) and its different flavours allow a consistent consideration of not only the dipole polarization, but also charge transfer process in the system. By the way, the most applications of this model concern the fast computation of charge distribution in large systems, while the inclusion of the dipole polarization is often left aside.

The CPE model can be derived from the density functional theory by introducing several assumptions [162, 163]. The basic idea of the CPE concept is that molecules are formed from "perturbed" neutral atoms. In the process of formation of a closed (with respect to the number of electrons) system, the electrons can flow between atoms to equalize the chemical potential everywhere in the system (note the analogy with the counterpart of classical statistical mechanics). It is noteworthy that in the frame of DFT one can rigorously show that the chemical potential is the negative of electronegativity. Consequently, the CPE concept is equivalent to the electronegativity equalization principle put forward by Sanderson in the 1950s [164].

There exist two main CPE approaches that differ in the choice of reference state for the system. In the model of Mortier *et. al.* [165] the electron density of the reference state is considered as superposed electron densities of isolated atoms, whereas the model by York and Yang [162] takes the ground state electron density of an arbitrary nuclear configuration as the reference state electron density.

Following Mortier *et. al.* [165], let us take a state, where all N atoms forming the system are at infinite distances from each other, as the reference state. The total electron density $\rho_0(\mathbf{r})$ and the total energy E_0 of the system in its reference state are then simply

$$\rho_0(\mathbf{r}) = \sum_{i=1}^{N} \rho_{0i}(\mathbf{r}), \tag{3.2.33}$$

and

$$E_0 = \sum_{i=1}^{N} E_{0i}, \tag{3.2.34}$$

where $\rho_{0i}(\mathbf{r})$ and E_{0i} are the electron density and the total energy of i-th isolated atom, respectively. When the atoms form the system the total energy change δE can be obtained by expanding the energy up to the second order with respect to the electron density change $\delta\rho_i(\mathbf{r}) = \rho_i(\mathbf{r}) - \rho_{0i}(\mathbf{r})$

$$\delta E = \sum_{i=1}^{N} \int \delta\rho_i(\mathbf{r}) \left[\frac{\delta E}{\delta\rho_i(\mathbf{r})}\right]_0 d\mathbf{r} + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} \int\int \delta\rho_i(\mathbf{r})\delta\rho_j(\mathbf{r}') \left[\frac{\delta^2 E}{\delta\rho_i(\mathbf{r})\delta\rho_j(\mathbf{r}')}\right]_0 d\mathbf{r}d\mathbf{r}'. \quad (3.2.35)$$

The functional derivative of the energy with respect to the charge density in the first term of (3.2.35) is equal to the electronegativity χ_i of the atom i [166].

$$\left[\frac{\delta E}{\delta\rho_i(\mathbf{r})}\right]_0 = \chi_i. \quad (3.2.36)$$

The second-order term in (3.2.35) can be obtained within the Kohn-Sham formalism by taking the functional derivative of the universal functional $F[\rho(\mathbf{r})]$ of the reference state density

$$F_0 = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} \int\int \frac{\rho_{0i}(\mathbf{r})\rho_{0j}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}d\mathbf{r}' + \sum_{i=1}^{N}\big(E_{\mathrm{xc}}[\rho_{0i}] + T[\rho_{0i}]\big), \quad (3.2.37)$$

where $E_{\mathrm{xc}}[\rho]$ and $T[\rho]$ are the exchange-correlation and kinetic energy functionals, respectively. In the CPE model the nonclassical contributions into the second-order term of (3.2.35), which are the second derivative of the exchange-correlation and kinetic energy functionals, are neglected. Therefore, the second-order term in (3.2.35)) can be written as

$$\left[\frac{\delta^2 E}{\delta\rho_i(\mathbf{r})\delta\rho_j(\mathbf{r}')}\right]_0 \simeq \left[\frac{\delta^2 F}{\delta\rho_i(\mathbf{r})\delta\rho_j(\mathbf{r}')}\right]_0 = \frac{1}{|\mathbf{r}-\mathbf{r}'|} \quad (3.2.38)$$

and by substituting (3.2.36) and (3.2.38) into (3.2.35) one obtains

$$\delta E = \sum_{i=1}^{N} \chi_i \int \delta\rho_i(\mathbf{r})\, d\mathbf{r} + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} \int\int \frac{\delta\rho_i(\mathbf{r})\delta\rho_j(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}d\mathbf{r}' \quad (3.2.39)$$

and the total energy of the system E is then equal to

$$E = \sum_{i=1}^{N} E_{0i} + \delta E. \quad (3.2.40)$$

At this point one should construct the total differential atomic densities $\delta\rho_i(\mathbf{r})$. In the general case these densities can be built as a sum of spherical and dipolar contributions

$$\delta\rho_i(\mathbf{r}) = q_i\, s_i(\mathbf{r}) + \boldsymbol{\mu}_i \cdot \mathbf{p}_i(\mathbf{r}), \quad (3.2.41)$$

where $s_i(\mathbf{r})$ and $\mathbf{p}_i(\mathbf{r})$ are orthonormalized spherical and dipolar form factors, respectively. Because of the orthonormalization condition, the q_i and $\boldsymbol{\mu}_i$ quantities correspond to the atomic charge and

atomic dipole, respectively. Considering only spherical contribution to the density change results in *s*-model, while taking both the density contributions in (3.2.41) into account corresponds to *sp*-model.

By substituting (3.2.41) into (3.2.39) one obtains the CPE energy expression in the following form

$$\begin{aligned} E &= E_0^* + \sum_{i=1}^{N} \chi_i \, q_i + \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} J(r_{ij}) \, q_i \, q_j \\ &+ \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{\beta=x,y,z} F_\beta(r_{ij}) \, q_i \mu_{\beta i} + \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{\beta,\gamma=x,y,z} P_{\beta\gamma}(r_{ij}) \, \mu_{\beta i} \mu_{\gamma j}, \end{aligned} \tag{3.2.42}$$

where E_0^* is an unknown energy term that includes exchange-correlation, kinetic and valence energy contributions. The $J(r_{ij})$, $F_\beta(r_{ij})$ and $P_{\beta\gamma}(r_{ij})$ are the charge-charge, charge-dipole, and dipole-dipole interaction potentials, respectively, defined as

$$J(r_{ij}) = \int\int \frac{s_i(\mathbf{r}) \, s_j(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \, d\mathbf{r} d\mathbf{r}', \tag{3.2.43}$$

$$F_\beta(r_{ij}) = \int\int \frac{s_i(\mathbf{r}) \, p_{\beta j}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \, d\mathbf{r} d\mathbf{r}', \tag{3.2.44}$$

$$P_{\beta\gamma}(r_{ij}) = \int\int \frac{p_{\beta i}(\mathbf{r}) \, p_{\gamma j}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \, d\mathbf{r} d\mathbf{r}'. \tag{3.2.45}$$

The full expressions for the potentials (3.2.43), (3.2.44) and (3.2.45) obtained for Gaussian form of the factors $s_i(\mathbf{r})$ and $p_{\beta i}(\mathbf{r})$ are given in Appendix A

The atomic charges and dipoles are sought such that they minimize the energy (3.2.42) under the condition that the total charge Q_0 of the system is conserved

$$\delta \left[E - \tilde{\chi} \left(\sum_{i=1}^{N} \int \Big(q_i \, s_i(\mathbf{r}) + \boldsymbol{\mu}_i \cdot \mathbf{p}_i(\mathbf{r}) \Big) \, d\mathbf{r} - Q_0 \right) \right] = 0, \tag{3.2.46}$$

where $\tilde{\chi}$ is an undetermined Lagrange multiplier equal to a global electronegativity of the system. From (3.2.46) one can easily obtain the derivatives of the energy with respect to the variables q_i and $\mu_{\beta i}$ as

$$\frac{\partial E}{\partial q_i} = \tilde{\chi}, \tag{3.2.47}$$

$$\frac{\partial E}{\partial \mu_{\beta i}} = 0. \tag{3.2.48}$$

Equation (3.2.42) can be recast in a matrix form as

$$E = E_0 + \boldsymbol{\chi}^T \mathbf{q} + \frac{1}{2} \mathbf{q}^T \mathbf{H} \mathbf{q} \tag{3.2.49}$$

that allows (3.2.47) and (3.2.48) to be written as

$$\mathbf{H}\mathbf{q} = \tilde{\chi}\mathbf{e} - \boldsymbol{\chi}, \tag{3.2.50}$$

where the vectors $\mathbf{q}$, $\boldsymbol{\chi}$, and $\mathbf{e}$ in (3.2.49) and (3.2.50) are defined as follows

$$\begin{aligned}
\mathbf{q} &= (q_1, q_2, \ldots, q_N, \mu_{x1}, \mu_{y1}, \mu_{z1}, \mu_{x2}, \mu_{y2}, \mu_{z2}, \ldots \mu_{xN}, \mu_{yN}, \mu_{zN}), && (3.2.51)\\
\boldsymbol{\chi} &= (\chi_1, \chi_2, \ldots, \chi_N, \underbrace{0, 0, 0, 0, 0, 0, \ldots, 0, 0, 0}_{3N}) && (3.2.52)\\
\mathbf{e} &= (\underbrace{1, 1, \ldots, 1}_{N}, \underbrace{0, 0, 0, 0, 0, 0, \ldots, 0, 0, 0}_{3N}) && (3.2.53)
\end{aligned}$$

and the matrix $\mathbf{H}$, called hardness matrix, is constructed of the $J(r_{ij})$, $F_\beta(r_{ij})$ and $P_{\beta\gamma}(r_{ij})$ pair potentials defined with eqs.(3.2.43), (3.2.44) and (3.2.45). Solving eqn. (3.2.50) for $\mathbf{q}$ gives

$$\mathbf{q} = \tilde{\chi}\,\mathbf{H}^{-1}\mathbf{e} - \mathbf{H}^{-1}\boldsymbol{\chi}, \tag{3.2.54}$$

whereas the global electronegativity $\tilde{\chi}$ can be obtained from eq. (3.2.54) as follows

$$\tilde{\chi} = \frac{Q_0 + \mathbf{e}^T\mathbf{H}^{-1}\boldsymbol{\chi}}{\mathbf{e}^T\mathbf{H}^{-1}\mathbf{e}} \tag{3.2.55}$$

with the use of the condition of conservation of total charge Q_0

$$\mathbf{e}^T\mathbf{q} = Q_0. \tag{3.2.56}$$

There exist two different representations of the CPE approach [167]. The first one uses atomic charges as charge variables and the representation is often referred to as atomic representation. The most known methods based on the atomic representation are the electronegativity equalization method (EEM) proposed by Mortier *et al.* [165] and the charge equilibration (QEq) method by Rappé and Goddard [168]. The second representation employs charge variables related to the charge transfer between pairs of atoms, e.g. atoms connected by chemical bonds. Consequently, the representation is called bond representation. The methods based on the bond representation are the atom-atom charge transfer (AACT) method [169] and split-charge equilibration (SQE) method [167].

Force fields including the chemical potential equalization principle were applied to a variety of systems such as water [76,170–177], ions in aqueous solutions [178,179], organic liquids [180–185], aqueus solvation of amides [186], vapor-liquid equilibrium [187], organic molecules [163,167,188–196], heterogeneous systems [197,198], organic solids [199] and microporous inorganic solids [200–202].

3.3 Polarizable force field: model and approaches.

Before proceeding to the description of the approach used for the development of the force field model developed in the present work, it might be useful to formulate the desired qualities of a force field. Such a list will not be specific to polarizable models, but it will rather reflect ultimate properties of an effective potentials model. In our opinion, qualities of a successful force field might be listed as follows:

(a) Functional form of effective potentials should be based on solid theoretical grounds. This should allow a transparent interpretation of different contributions to the total energy and a judicious choice of parameters.

(b) Force field should be flexible enough, so that an extension based on the same theoretical model could be added "on top" of the existing force field. Furthermore, it should also be adaptable to new systems (atomic kinds) without complete reworking the model.

(c) Force field should be neither complex to program nor computationally expensive so that it could be employed in simulations of large systems and in a long time-scale.

(d) Force field should mimic the energetic, structural, and dynamical characteristics of different conformations (polymorphs) of the same chemical compound. However, the reproduction of the correct behaviour and of the right trends in properties is more important than exact coincidence of values. Both, however, would be perfect.

While the three first in the list items are the qualities that need to be taken into account while designing the model, the last quality reflects the issues concerning the transferability and predictive power of the force field. The remaining part of this section presents the approaches that are expected to allow the developed model to meet the points in the above checklist.

Like many other force fields, the envisaged model splits the total energy E_{TOT} into energy contributions based on the distance criterion

$$E_{\mathrm{TOT}} = E_{\mathrm{LR}} + \left[E^{C}_{\mathrm{SR}} + E_{\mathrm{D}}\right] \tag{3.3.1}$$

where E_{LR} is the energy term corresponding to long-range interatomic interactions and the two terms in the square brackets stand for the contributions of short-range interactions. The long-range and short-range do not mean that the corresponding interactions exist only at long or short distances, respectively. This notation is used to indicates the rate of decay of the interaction potentials with the distance.

Long-range interactions. The previous sections of this chapter clearly show that the inclusion of electronic polarization into a force field goes via the model related to the long-range interactions. In the present work, the long-range interactions in the systems of interest are chosen to be treated with the chemical potential equalization approach. This choice is dictated by two reasons. First,

the CPE model has its origin in the density functional theory (cf. item (a) in the checklist above). Second, CPE models of successive levels of complexity can be designed (item (b)) by taking different contributions to the differential atomic densities (3.2.41). The increase of complexity level is expected to lead to an improvement of the force field quality.

To obtain the long-range energy term, an extensive parameterization of CPE models is carried out using results of quantum-chemical calculations of isolated and periodic systems. As an outcome, the models yield a charge distribution in the system. The atomic charges obtained with the model are to be used as a basic quantity linking the long-range and short-range potentials of the force field. Next chapter provides a thorough presentation of the CPE model used and discusses results of the use of the model for computing the charges and electrostatic potential in the systems of interest.

Short-range interactions. Short-range energy contribution to the total energy consists of two terms. The first term E^C_{SR} includes the repulsion energy due to the overlap of atomic electron densities at short distances and the attraction energy of exchange and short-range electrostatic interactions. These two "classical" contributions to the short-range energy can directly be obtained from quantum-chemical calculations using either the HF method or (semi-) local XC functionals. The parameterization of this classical part will be done with the use of the force-matching technique. This method first suggested by Ercolessi and Adams [203], has nowadays become a standard routine of force filed parameterization using the results of quantum-chemical calculations. Chapter 6 is devoted to the parameterization of the E^C_{SR} energy term by using different models of effective potentials.

The second term E_{D} corresponds to the dispersion energy. Quantum-chemical calculations at the HF level or those using (semi-) local functionals lack this energy, which is due to the interactions of instantaneous multipole moments of electron density distribution. Consequently, the contribution of dispersion energy is to be obtained by using another approach. The method of Silvestrelli based on the use of maximally localized Wannier functions was chosen for this purpose [204]. The method has been successfully employed for computing the dispersion interactions in different systems. The adaptation and application of the Silvesterelli method for the systems of interest is discussed in Chapter 5.

In addition to the polarizable long-range part of the force field, an attempt has been undertaken to make short-range potentials depending on the structure of system and on the environment of atoms. As it was mentioned above, the atomic charges obtained within the CPE model were used as a parameter relating the instantaneous charge state of the atoms to atomic parameters of short-range interaction potentials. Importance of such a model extension is considered in the corresponding chapters.

Chapter 4

Modelling long-range interactions in silicates.*

4.1 Introduction

The knowledge of the electronic distribution in a molecular system and the variation of the distribution upon a perturbation is a key-point to the understanding of the system behavior at the atomic level. This issue is commonly addressed by quantum-chemical computations, which despite the significant progress in both the computer hardware and computational algorithms, remain an expensive task, especially when studying complex disordered systems. Consequently, significant attention is paid to the development of simplified parametric models that allow the quantities related to the electronic distribution to be computed with a modest cost, while keeping transferability and/or precision of quantum-chemical calculations. Besides the proper utility, such models can further be used as a part of polarizable force fields, thus permitting realistic studies of large complex systems (yet) inaccessible to quantum-chemical calculations. An additional advantage of parametric models is that they provide much more transparent relation between obtained results and the underlying physics, whereas the outcome of a quantum-mechanical model is harder to interpret due to the complexity of the electronic wavefunction.

The concept of atomic charges provides a tangible means of relating the electronic distribution in the system with its structure and reactivity at the atomic level. Consequently, a wealth of efforts was and is concentrated on the development of models and methods capable of reliable computation of atomic charges and by making use of this information, of getting new insight into the reactivity or structural and dynamical characteristics. Models based on the chemical potential equalization (CPE) or electronegativity equalization (EE) concept appear to be among the most promising candidates to predict the charge distribution in molecules and solids. A number of CPE-based models have been proposed over the past two decades [165, 167–169, 205,

*T. Verstraelen, S. V. Sukhomlinov, V. Van Speybroeck, M. Waroquier, and K.S. Smirnov. Computation of charge distribution and electrostatic potential in silicates with the use of chemical potential equalization models. *J. Phys. Chem. C*, 116, 490-504, (2012)

206], in which atomic charges are variational degrees of freedom and the *electronic* ground state corresponds to an energy minimum under a total charge constraint. The latter is equivalent to the condition that all atomic chemical potentials (or electronegativities) must be equal. This concept was generalized to s- and p-type density basis functions [162], effectively extending the CPE approach with atomic inducible dipoles. CPE models were applied to different systems such as dense and microporous inorganic solids [170, 200, 201], water [76, 171, 175–177, 207], ionic liquids [208], organic liquids [180, 182, 184], other organic molecules [163, 188, 191, 194–196, 209], biomolecular systems [189, 193], and heterogeneous systems [183, 197, 198]. Furthermore, CPE models were also used to study various properties of electronic systems, including exchange-polarization coupling, [175, 207] intermolecular charge transfer, [184] charge transfer during bond dissociation, [194, 196] non-linear polarizability [176, 209] and optical linear response properties such as IR and Rahman intensities. [163]

The earliest CPE model is the Electronegativity Equalization Method (EEM) [165]. Although this model was later modified with more realistic electrostatic interactions [168, 210] or by adding more interaction sites [205], EEM and these early extensions predict that the dipole polarizability scales cubicly with system size, while in the macroscopic limit, one observes a linear relation for dielectric materials. [169, 211] The latter weakness is problematic when one tries to apply parameters based on small molecules to larger (dielectric) systems [212]. A few ad-hoc improvements were proposed to overcome this fundamental error. [163, 167] We are mainly interested in a recent extension of the EEM, the Split Charge Equilibration (SQE) [167], which fixes this polarizability scaling issue [211] and is capable of properly describing electronic dielectric screening effects in condensed matter. [213] It is especially encouraging that the SQE model exhibits a transition between the EEM-like polarizability scaling for small systems to the linear scaling for extended systems, [211, 213] as is often observed in computational studies [207, 214–217].

Since the primary goal of all CPE-based models is the fast and reliable computation of charge distribution in large systems, the parameterization of such a model is done by fitting atomic charges produced by the model to those derived from quantum-chemical calculations. The problem is that the atomic charge is not a quantum-chemical observable [218] and consequently, many electronic density partitioning schemes can be used to produce the reference charge values. The first parametrizations employed atomic charges resulting from the Mulliken population analysis [165]. More recently, atomic charges fitted to the electrostatic potential (ESP) fits [210, 219, 220], Natural charges [195, 219], Bader's Atom In Molecule (AIM) charges [193] or charges produced by Stockholder partitioning (Hirshfeld [219] or Hirshfeld-I [195]) were used for this purpose.

Since the beginning of the development of CPE-based models, oxide materials ((alumino)silicates, in particular) were one of important application fields of the corresponding methods. [170, 200, 201] A CPE-based model is also an essential component of a polarizable force field for silicates. Nevertheless, to our knowledge, an extensive calibration and validation of the EEM [165] or the SQE [167] for oxide crystals based on a large set of representative cluster models is not yet carried out. (For organic systems, such extensive calibrations can be found in the literature [182, 191, 195].) In general, it is not clear yet whether parameters for the EEM or SQE

model are simply transferable from small clusters to extended or periodic systems. In this paper, we derive parameters for both models based on DFT computations on a large number of silicate clusters, and use the results to answer the following research questions:

(i) How well do different CPE-based models reproduce the electronic density distribution, the electrostatic potential, and electronic linear response properties of oxide systems?

(ii) Which types of input data are needed in a reliable parameter calibration protocol for two frequently used CPE models (EEM and SQE)?

(iii) To what extent are the parameters derived from cluster computations transferable to systems of different size and density, in the view of the development of a polarizable force field for condensed matter simulations? Especially the transferability of the parameters to periodic systems is extensively tested.

The Chapter is organized as follows. Section 4.2 briefly discusses the EEM and SQE models and presents parameters entering the basic EEM and SQE equations. It goes on to describe the systems used for the calibration and testing of parameters and provides details of quantum-chemical calculations used to obtain reference quantities employed in the parameterization. The final part of the second section deals with the criteria and strategy of the parameter calibration. Results of the calibration of different CPE-based models are reported in the first part of Section 4.3 which also illustrates the performance of the models in reproducing different characteristics of molecular systems. The second part of the third section deals with the transferability of EEM and SQE parameterizations to periodic systems. The last part of the Chapter provides the conclusions of the work and gives answers to the questions listed in the above paragraph. Additional information that might be of interest for the reader is given in Appendices and as Supplementary Information to the paper.

4.2 Theoretical models and computational details

4.2.1 EEM and SQE models

In the framework of CPE models, a general expression for the charge-dependent energy $E(\mathbf{u})$ of a system can be written in a matrix form

$$E(\mathbf{u}) = \mathbf{x}^T\mathbf{u} + \frac{1}{2}\mathbf{u}^T\mathbf{H}\mathbf{u}, \tag{4.2.1}$$

where $\mathbf{u}$ is a vector of charge variables, $\mathbf{x}$ a vector with first-order parameters, and $\mathbf{H}$ is the so-called hardness matrix built of second-order parameters and containing information about the electrostatic interactions in the system. Following the variational principle the charge variables u_i can be obtained as a solution of equation

$$\mathbf{u} = -\mathbf{H}^{-1}\mathbf{x}. \tag{4.2.2}$$

The dependence of elements of the matrix **H** on the interatomic distances makes the charge variables dependent on the geometry and on the environment of atoms in the system.

The first model used in the present study, the electronegativity equalization model by Mortier and co-workers [165], employs atomic charges, q_i, as the charge variables in 4.2.1. The explicit EEM form for the energy reads

$$E_{\mathrm{EEM}}(\mathbf{q}) = \sum_i \left[\chi_i q_i + \frac{1}{2}\eta_i q_i^2\right] + \frac{1}{2}\sum_i \sum_{j\neq i} q_i q_j \, J(r_{ji}), \tag{4.2.3}$$

where the sums run over atoms in the system, $J(r_{ji})$ is the electrostatic potential between the atoms i and j, and χ_i and η_i are the first- and second-order parameters known as effective atomic electronegativity and hardness, respectively. Minimization of this energy under a total charge constraint yields the equilibrium charge distribution. It is shown that the EEM energy in 4.2.3 can be deduced from the density functional theory [165] and the model is of great help for the prediction of the charge distribution in large systems and for qualitative understanding of system's reactivity on the basis of reactivity indices. Furthermore, it is often used as an underlying numerical model of the conceptual density functional theory [221].

In its original formulation the EEM was found to suffer from a number of deficiencies. The most crucial one limiting the use of EEM in the development of polarizable force fields, is the strongly non-linear behavior of the dipole polarizability with the system's size. The SQE model [167] circumvents this problem by employing split-charges as the charge variables in 4.2.1. A split-charge or charge transfer parameter (CTP), p_{ij}, is an amount of charge transfered from atom j to atom i, when the electronegativities of atoms constituting the system get equalized. The CTPs satisfy the following condition

$$p_{ij} = -p_{ji} \tag{4.2.4}$$

and the net atomic charge q_i on atom i is then equal

$$q_i = \sum_j^{(i)} p_{ij}, \tag{4.2.5}$$

where the superscript (i) is used to denote that the sum runs over atoms j to which the charge transfer from atom i is allowed. It is noteworthy that the condition, 4.2.4, constrains the total charge of system to be equal to zero.

In terms of the split-charges the expression for the energy reads

$$\begin{aligned} E_{\mathrm{SQE}}(\mathbf{p}) = &\sum_i \sum_k^{(i)} \xi_{ik} p_{ik} + \frac{1}{2}\sum_i \sum_k^{(i)} \kappa_{ik} p_{ik}^2 + \frac{1}{2}\sum_i \eta_i \sum_k^{(i)} \sum_l^{(i)} p_{ik} p_{il} \\ &+ \frac{1}{2}\sum_i \sum_{j\neq i} J(r_{ij}) \sum_k^{(i)} \sum_l^{(j)} p_{ik} p_{jl}, \end{aligned} \tag{4.2.6}$$

where the sums run over the CTPs and ξ_{ik}, κ_{ik}, and η_i are the first-order and second-order

parameters, respectively. The parameters ξ_{ik} and κ_{ik} are the bond electronegativity and bond hardness from the atom-atom charge transfer (AACT) model [169], respectively. It is worthy of noting that 4.2.6 is given in a redundant set of charge variables p_{ij}. For the practical applications of SQE model one has to write 4.2.6 in the unique set of CTPs using the condition 4.2.4 and the implicit constraints $\xi_{ji} = -\xi_{ij}$ and $\kappa_{ji} = \kappa_{ij}$ [167].

Although several subtly different forms of the SQE model are possible, it is essential that the bond hardness term (second term in 4.2.6) is present. [167] In the most general case, when the charge transfers are allowed between all atoms in the system and all the bond hardness parameters are set to zero, the SQE model is isomorphous to the EEM and the parameters in 4.2.6 can then be represented via those of 4.2.3. One can, however, limit the number of CTPs by permitting the charge transfers only between certain atoms. Thus, in the present work only the CTPs for the nearest neighbors (i.e. bonded atoms) are non-zero. Given the parameters in 4.2.6 are known, these CTPs can be found by minimizing the energy $E_{\mathrm{SQE}}(\mathbf{p})$ with respect to the split charges p_{ij}. 4.2.3 and 4.2.6 give the charge-dependent energy in the EEM and SQE representations, respectively, for isolated systems. The modification of the equations as well as of the hardness matrix elements for periodic systems are given in Appendix B. It is worthy of note that the condition of chemical potential (electronegativity) equalization never appears in the SQE equations because the total charge is implicitly constrained to be zero. It can, however, be shown that the minimization of energy, 4.2.6, with respect to CTPs is equivalent to the use of electronegativity equalization principle in the EEM [167].

In 4.2.3 and 4.2.6 $J(r_{ij})$ stands for an electrostatic interaction potential that is taken as Coulombic potential $1/r_{ij}$ between the two point charges in the simplest case. However, a realistic electrostatic interaction between a pair of atoms significantly deviates from the $1/r$ dependence at short distances. Hence, to mimic the interactions between bonded atoms, it is more realistic to use a potential $J(r_{ij})$ that describes the interaction energy between distributed charges [168,210]. In the present work the $J(r_{ij})$ potential function was taken in the form

$$J(r_{ij}) = \frac{\mathrm{erf}(\alpha_{ij} r_{ij})}{r_{ij}}, \tag{4.2.7}$$

which corresponds to the interaction potential between two Gaussian charge distributions $g(\mathbf{r})$ given by

$$g_i(\mathbf{r}) = \left(\frac{1}{2\pi R_i^2}\right)^{3/2} \exp(-(\mathbf{r} - \mathbf{r}_i)^2 / 2R_i^2), \tag{4.2.8}$$

where $\mathbf{r}_i$ is the position of atom i and the parameter R_i (the standard deviation) can be viewed as an effective radius of atom i. With this expression the parameter α_{ij} in 4.2.7 is obtained as

$$\alpha_{ij} = \left(\frac{1}{2R_i^2 + 2R_j^2}\right)^{1/2}. \tag{4.2.9}$$

The quantities χ_i, η_i, ξ_{ik}, κ_{ik}, and R_i in 4.2.3, 4.2.6, and 4.2.8 are the model parameters whose

values were obtained in a fitting procedure as described below.

4.2.2 Computational schemes

The EEM and SQE models are calibrated on two different reference quantities derived from quantum-chemical calculations. The first calibration scheme follows a common way of parametrizing CPE models and uses atomic charges as reference quantities to obtain the model parameters. In the present work the reference charges are computed with the iterative Hirshfeld (HI) method [222]. Since the atomic charge is not an observable, the choice of population analysis scheme is not unique and other charge schemes could have been used [219, 220]. The choice of the HI partitioning scheme is based (i) on the fact that the computation of HI charges relies on the partitioning of the electronic density, i.e. a quantum-chemical observable, (ii) on a weak dependence of the HI charges on the basis set used in the quantum-chemical calculations [223], (iii) on the quality of the ESP predicted by HI charges [224], and (iv) on the robustness of HI charges with respect to conformational changes [212, 225]. The second parameterization scheme is based on the reference values of the electrostatic potential (ESP) in the view of the application of the EEM and SQE models to the development of polarizable force fields. The combination of model/quantity finally resulted in four computational schemes employed in the parameterization EEM/HI, EEM/ESP, SQE/HI, SQE/ESP. The reference data used for the parameter calibration are derived from quantum-chemical calculations on isolated molecules. Following procedures outlined below, the isolated systems are divided into training and validation sets that allowed assessment of the transferability of parameters: the training set is used for the actual calibration, while the validation set is used to test how well the parameters would work for molecules not present in the training set. In addition to the validation set containing isolated molecules, quantum-mechanical computations on periodic systems are used to test the transferability to crystalline oxides.

4.2.3 Training and validation sets

The database of isolated molecule structures consists of 207 oxide clusters containing three-coordinated aluminum cations, and four-coordinated silicon and zirconium cations. The finite clusters are mostly hydrogen-terminated. Some hydroxyl terminations are present, but none of them form internal hydrogen bonds that distort the cluster geometry. The chemical formula of these structures can be written as $\mathrm{Si}_x\mathrm{Al}_y\mathrm{Zr}_z\mathrm{O}_{\frac{1}{2}(4x+3y+4z-n-m)}\mathrm{H}_n(\mathrm{OH})_m$. The molecular structures are distributed over a training set (103 molecules) and validation set (104 molecules), such that they are statistically representative for each other. Table 4.1 gathers the relevant information on each set.

The database of periodic systems contains crystalline modifications of silicon oxide, tetragonal modification of zirconia (t-ZrO_2), and zirconium silicate ($ZrSiO_4$, zircon polymorph). Silica structures chosen for these calculations were α-quartz, α-cristobalite, and four all-silica zeolite structures of the JBW, DFT, SOD, and NPO topologies [226]. This selection of periodic systems,

Table 4.1: Number of atoms of each element in training and validation sets for molecular systems.

Atom	Training	Validation
H	747	747
O	534	533
Al	57	57
Zr	65	65
Si	230	231
Total	1633	1633

Table 4.2: Structural characteristics of periodic systems.

Structure	Space group	No. of atoms per unit cell	V_0 (Å^3)^a	FD^b
α-quartz	$P3_221$	9	113.007	26.6
α-cristobalite	$P4_12_12$	12	171.104	23.4
JBW (SiO_2)	$Pmma$	18	321.279	18.7
DFT (SiO_2)	$P4_2/mmc$	24	456.609	17.5
NPO (SiO_2)	$P\bar{6}2c$	18	350.960	17.1
SOD (SiO_2)	$Im3m$	36	725.049	16.6
t-ZrO_2	$P4_2/nmc$	6	67.810	
$ZrSiO_4$ (zircon)	$I4_1/amd$	24	269.620	

^{a} Unit cell volume after geometry optimization. ^{b} Framework density (number of Si atoms per 1000 Å^3).

which are not used for the calibration of the parameters, permits to test the transferability of parameters on both dense and microporous materials. Table 4.2 provides some pertinent characteristics of the structures; CIF files and images of the optimized structures can be found in Supporting Information to the paper.

4.2.4 Computational details

Isolated Molecule calculations The calculations on the isolated systems were performed at the DFT level with the B3LYP exchange-correlation functional [227]. The 6-311+(d,p) [228, 229] all-electron basis set was used for the H, O, Al, and Si atoms, while the Zr atoms were described with the LANL2DZ effective core potentials [230]. The geometry of molecules was first optimized without any symmetry constraints and was followed by the calculation of the electronic density and electrostatic potential values on a three-dimensional grid of points (cubes files). The calculations were done with Gaussian03 code [231].

Periodic calculations The calculations of periodic systems were carried at the same level as the isolated molecule calculations, i.e. within DFT using the B3LYP [227] functional, and they were done with the CRYSTAL06 code [232]. Atoms were described by all-electron basis sets: 976-31d621G for Zr atoms, 86-311G* for Si atoms, 85-11G* for Al atoms, 6-31G* for O atoms and 8-21G* for H atoms [233–236]. The Brillouin zone integration was carried out over a Monkhorst-Pack grid of $6 \times 6 \times 6$ k-points [23]. Model structures were optimized under the constraints imposed by the symmetry of crystalline lattices. These structural optimizations were followed by the computation of cube files with the electronic density and electrostatic potential.

Computation of Hirshfeld-I charges The computation of the Hirshfeld-I charges is carried out with HiPart [237]. For the isolated molecules, the Hirshfeld-I charges are derived using the procedure outlined in the work of Bultinck *et al* [222]. For the periodic systems, a slightly modified procedure is used for technical reasons. Due to the presence of the cusps in the electron density, one should use spherical atom-centered grids to carry out the numerical integrations required for the Hirshfeld-I partitioning [238]. For the isolated systems, this approach is feasible because the formatted checkpoint files from Gaussian03 provide sufficient information to evaluate the electron density on arbitrary grids. However, it is technically infeasible to extract similar grid data from CRYSTAL06 computations in a post-processing analysis. Therefore we used the cubic grids from the density cube files to approximate the Hirshfeld-I charges as follows: (i) first the contribution to the density from the core orbitals is subtracted, using spherically averaged core-densities from atomic computations, and (ii) the remainder (which is a smooth function that can be integrated on cubic grids) is partitioned using Hirshfeld-I algorithm.

4.2.5 Parametrization strategy

Cost functions In line with the calibration of EEM and SQE parameters for organic systems in earlier work, [195] two types of cost functions (X) are used in the calibration procedure: static (X_{S}) and response (X_{R}) ones. The former is based on the values of the reference quantity itself (HI charge or ESP), whereas the latter corresponds to the derivatives of the values with respect to an external perturbation taken in the form of an uniform electric field. For each of the four computational schemes (EEM/HI, etc.) a series of parameters is calibrated with the cost function

$$X_\lambda = (X_{\mathrm{S}} + \lambda X_{\mathrm{R}})/(1+\lambda) \tag{4.2.10}$$

where the weight λ is scanned over several orders of magnitude. When λ approaches zero, the total cost approaches the static cost (X_{S}), while the response cost has a minimal effect. The lower bound of the λ-scan was chosen such that the static cost function converges to a constant value. Similar considerations were used to determine the upper bound of the λ-scan. Some testing revealed that the transition from a purely static cost function to a pure response cost function is found for λ going from 10^{-6} to 10^{0}. The final choice of λ (after performing the scans) is based on the following criteria:

- Small values of both the X_{S} and X_{R} cost functions,
- Minimum number of constraints (*vide infra*),
- Small value of the condition number of the Hessian matrix of the cost function in 4.2.10. This guarantees that the parameters have a minimal sensitivity to "noise" in the reference characteristics. [239]

The static X_{S} and response X_{R} cost functions for each CPE model (EEM or SQE) and for each reference quantity (HI charges or ESP) are discussed in Appendix C in detail.

Given that the response cost function (X_{R}) only depends on the second-order parameters, one could try to reduce the number of independent parameters by splitting the calibration into two steps: (i) first determine the second-order parameters that minimize X_{R}, (ii) consequently keep these second-order parameters fixed and find the remaimder of the parameters by minimizing the static cost function (X_{S}). However, this was not helpful because the second-order parameters found in the first step were systematically poorly determined and reached unreasonable values, making it impossible to find useful first-order parameters to complete the model. Only when a linear combination of both cost functions was used, the calibrations became well-conditioned and reasonable parameters could be obtained.

Constraints There are a series of inequality constraints for the second-order parameters to keep the matrix of second order coefficient in both the EEM and SQE positive definite: [195, 211]

- The radii of the charge distributions have to be larger than 0.1 Bohr; we do not take zero as the lower bound due to the next inequality constraint.
- The atomic hardness has to be larger than the self-interaction potential of atomic Gaussian charge distribution: $\eta_i \geq (\sqrt{\pi}R_i)^{-1}$
- The bond hardness κ_{ik} has to be positive.

Two additional constraints on the parameters must be introduced because the cost function, 4.2.10, is not sensitive to deviations from these two constraints. [239] Thus, in case of EEM the effective electronegativity of hydrogen atom is fixed at $\chi^*_{\mathrm{H}} = 0.0$ eV. In case of the SQE model the effective atomic hardness of hydrogen is kept at the lower bound corresponding the radius R_i of the charge distribution in 4.2.8, i. e. $\eta^*_{\mathrm{H}} = (\sqrt{\pi}R_{\mathrm{H}})^{-1}$.

Minimizer algorithm Due to the non-linear dependence of the EEM or SQE charges on the parameters, the calibration is a non-linear least-squares (NLLSQ) problem. A conjugate gradient optimizer with a diagonal preconditioner is used to optimize the parameters [240]. An active set algorithm is used to implement the inequality constraints. A inequality constraint only becomes active when the optimizer tries to push the parameters into the infeasible region. As soon as the dot product of the gradient of the cost function and the normal of a constraint becomes negative, the corresponding constraint is deactivated. The initial values for the parameters χ^*_i, η^*_i, ξ^*_{ik}, κ^*_{ik},

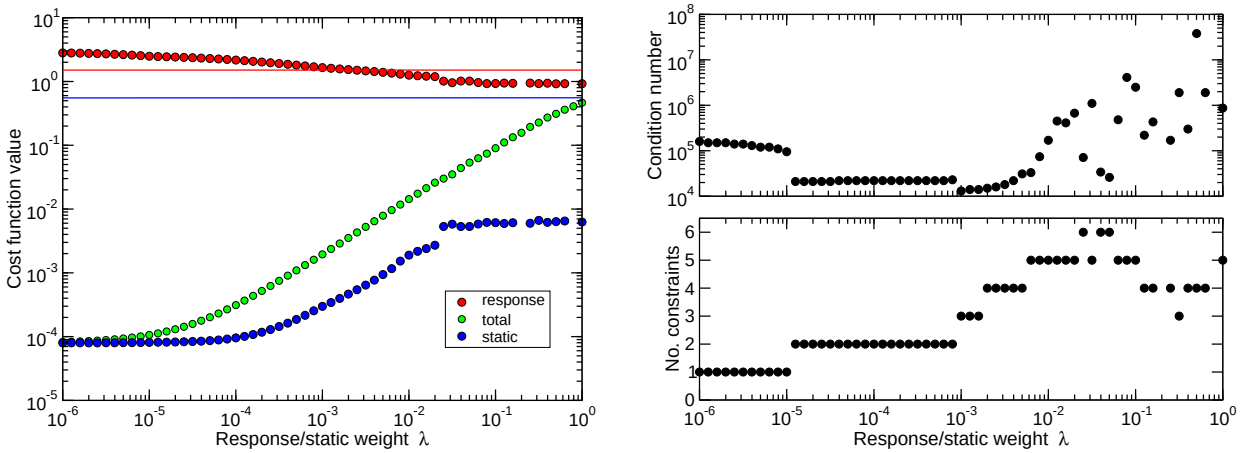

Figure 4.1: Dependence of EEM/HI total X_λ, static X_S, and response X_R cost functions on the weight λ, 4.2.10. Horizontal blue and red lines indicate "worst-case" values of static and response cost functions, respectively (see text for discussion).

and R_i, are 0 eV, 20 eV, 0 eV, 5 eV and 1 Å, respectively. The optimal parameters are not sensitive to the initial values. For parameters that have a lower bound, the initial values are set such that all inequality constraints are satisfied with a considerable margin.

4.3 Results and discussion

4.3.1 Choice of the best set of parameters

In order to find an optimal set of parameters for each computational scheme (model/quantity combination) the coefficient λ in 4.2.10 was varied over several orders of magnitude in the [0, 1] interval and a set of parameters minimizing X_λ was found in a NLLSQ fitting procedure for each value of λ. The best set of parameters was then chosen using the above criteria, i.e. the values of the static and response cost functions, the condition number, and the number of constraints.

4.1 presents the dependence of the total (X_λ), static (X_S), and response (X_R) cost functions (left panel), the number of constraints, and the condition number (right panel) on the weight λ for the EEM/HI scheme. The blue and red horizontal lines indicate "worst-case" values of the static and response cost functions, respectively. This "worst-case" value corresponds to the value of the cost function when the model would predict zero values for the target quantities, i.e. when there are no electrostatic interactions or the charges are not sensitive to an external electric field. In principle, the cost function may become higher when the predicted values have the wrong sign. However, in the latter case it is more "accurate" not to model electrostatic interactions at all, i.e. use the "worst-case" limit. Therefore the "worst-case" values of the cost function indicate a threshold below which the use of the corresponding quantity in the calibration procedure starts to make sense.

4.1 shows that at low values of λ (pure static case) the response cost function is above the

threshold level. An increase of λ does not lead to an improvement of the situation: when the response cost function drops below the "worst-case" value, the static cost function has already increased by more than an order of magnitude. At even larger values of λ, the total cost function becomes ill-defined and a large number of constraints needs to be used to get the minimization of 4.2.10 converged (see 4.1, right panel). It is also noteworthy that regardless the λ value the response cost function remains close to the "worst-case" value. Such a behavior points to the fact that the EEM with HI charges only makes sense when the static cost function is used. No improvement is possible by including response information. The EEM/ESP scheme reveals similar behavior. (Figure of λ-scan is included in the Supporting Information.) The main difference between the EEM/HI and the EEM/ESP calibrations, is that in the latter case the optimal value of the static cost function is merely 0.18 times the "worst case" value, while for EEM/HI this ratio is as low as 1.4×10^{-4}. Because of this poor result, we do not consider the EEM/ESP model for the validation in the remainder of the paper.

The above results allow to conclude that the electronegativity equalization model is only useful for computing static characteristics, such as the charge distribution and (to lesser extent) the electrostatic potential, whereas it fails to mimic the response of these quantities to an applied external perturbation. Consequently, sets of parameters corresponding to pure static cost functions in 4.2.10 were chosen for the EEM-based schemes.

The SQE-based schemes behave differently. The best parameter set for the SQE/HI scheme can be found at $\lambda = 10^{-3.4}$, where both the static and response cost functions increase by 60 % above their minimum values, whereas the total cost function is still close to its optimum value. (Figure of the λ-scan is included in the Supporting Information.) It should also be noted that, in contrast to the EEM/HI scheme, the parameters obtained in the limit of high values of λ can be useful because both the partial cost functions remain significantly below their respective "worst-case" values, while a small value of the condition number points to the stability of the NLLSQ solution with respect to "noise" in the training set.

4.2 displays the values of the total, static, and response cost functions (left panel), the number of constraints and value of the condition number (right panel) as a function of the weight λ in 4.2.10 for the SQE/ESP scheme. Again, in contrast to the EEM/HI model (4.1), both the static and response SQE/ESP cost functions lie always below their respective "worst-case" values and a compromise between the static and response characteristics can be found for the value of $\lambda = 10^{-5}$ with only one half-open constraint active ($\eta_{\mathrm{O}} = (\sqrt{\pi} R_{\mathrm{O}})^{-1}$) and a small value of the condition number. Note that an increase of λ value leads to errors in the static characteristics that cannot be compensated by the decrease of the error in the response data. Furthermore a larger number of constraints needs to be used to converge the function, 4.2.10, to the minimum.

We can conclude that the SQE parameters can be calibrated to reproduce both the static and response data, either using atomic charges or the ESP as target data. This is a distinct improvement compared to the EEM where only the EEM/HI calibration in the static limit behaves satisfactory.

4.3 gathers the best sets of EEM and SQE model parameters that were chosen according to

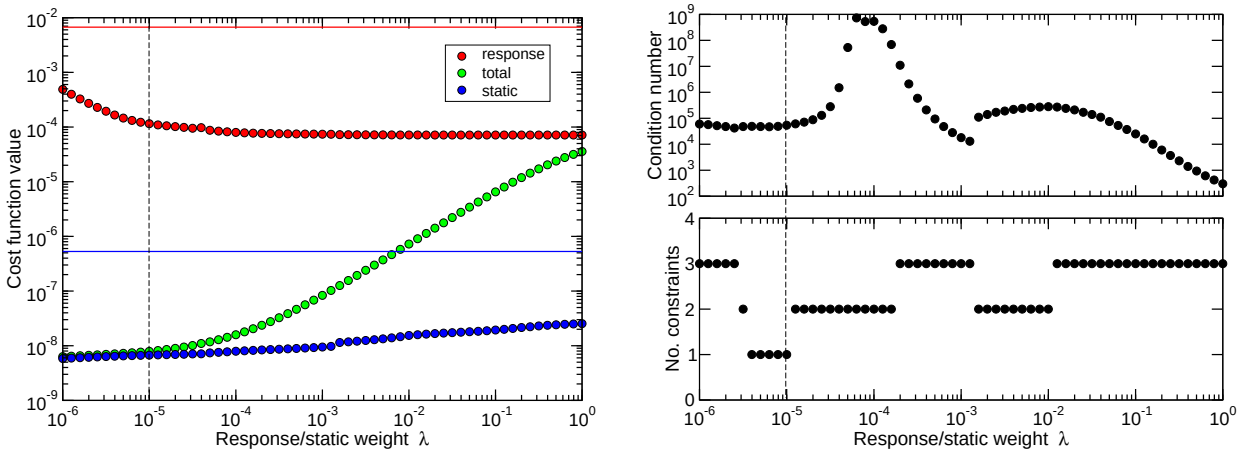

Figure 4.2: Dependence of SQE/ESP total X_λ, static X_S, and response X_R cost functions on weight λ, 4.2.10. Horizontal blue and red lines indicate "worst-case" values of static and response cost functions, respectively (see text for discussion; vertical dashed lines indicate the selected point.

the criteria mentioned above after the analysis of each of four model/quantity combinations.

4.3.2 Charge calculation with EEM and SQE models

Isolated systems. Despite the ambiguity of the atomic charge and the existence of many charge partitioning schemes, atomic charges provide an extremely useful guidance for the understanding of many properties of system at the atomic level. Consequently, models permitting a fast and reliable computation of charge distribution are of significant interest. The EEM was such a scheme from the very beginning of its development.

4.3 shows the correlation of charges computed by the EEM or SQE models with the reference iterative Hirshfeld charges derived from the results of DFT calculations on molecules of the validation set. Both models perform remarkably well. The root mean square deviation (RMSD) between the SQE/HI and DFT/HI charges, 0.0213 e, is slightly lower than the RMSD between the EEM/HI and DFT/HI charges, 0.0284 e. These relatively small RMSD values testify the transferability of both the CPE models for the computation of charge distribution in isolated systems of modest (up to 20-30 atoms) size.

Periodic systems. In what follows, we will validate to what extent the EEM/HI and SQE/HI calibrations can reliably predict the atomic charges in periodic systems. After all, the parameters are based on a training set containing only isolated systems. The left-hand part of 4.4 reports the atomic charges computed with EEM/HI and SQE/HI schemes for the periodic structures of the validation set and compares the charges with those obtained by applying the HI procedure to the electronic density from periodic DFT computations (DFT/HI column). The performance of the schemes in the prediction of the charge distribution in periodic systems is not that good as one

Table 4.3: EEM and SQE parameters obtained in the NLLSQ fits.

Parameter	EEM-HI	EEM-ESP	SQE-HI	SQE-ESP
χ_{H} (eV)	0.0000	0.0000		
χ_{O} (eV)	4.1441	3.7764		
χ_{Al} (eV)	−7.0722	−5.9772		
χ_{Si} (eV)	−5.4968	−0.1682		
χ_{Zr} (eV)	−17.3596	−221.4566		
η_{H} (eV)	13.7726	16.7508	11.9294	14.4867
η_{O} (eV)	15.0954	17.5301	11.9075	10.6047
η_{Al} (eV)	12.1315	15.0322	10.4054	10.0619
η_{Si} (eV)	11.8331	9.4161	10.2542	9.2006
η_{Zr} (eV)	13.7565	153.5238	10.7434	8.5684
$\xi_{\mathrm{H\text{-}O}}$ (eV)			−2.7666	−1.9923
$\xi_{\mathrm{H\text{-}Al}}$ (eV)			3.1308	2.3323
$\xi_{\mathrm{H\text{-}Si}}$ (eV)			2.0820	0.4558
$\xi_{\mathrm{H\text{-}Zr}}$ (eV)			5.9027	1.6812
$\xi_{\mathrm{O\text{-}Al}}$ (eV)			7.7564	4.6836
$\xi_{\mathrm{O\text{-}Si}}$ (eV)			5.8791	1.6479
$\xi_{\mathrm{O\text{-}Zr}}$ (eV)			8.3020	2.3925
$\kappa_{\mathrm{H\text{-}O}}$ (eV)			5.1124	3.1167
$\kappa_{\mathrm{H\text{-}Al}}$ (eV)			3.1262	1.0450
$\kappa_{\mathrm{H\text{-}Si}}$ (eV)			3.7648	1.3472
$\kappa_{\mathrm{H\text{-}Zr}}$ (eV)			2.5118	0.1420
$\kappa_{\mathrm{O\text{-}Al}}$ (eV)			6.1704	5.8499
$\kappa_{\mathrm{O\text{-}Si}}$ (eV)			6.1868	4.1075
$\kappa_{\mathrm{O\text{-}Zr}}$ (eV)			4.0558	2.9940
R_{H} (Å)	0.7724	0.4850	0.6810	0.5608
R_{O} (Å)	0.5382	0.4634	0.6823	0.7661
R_{Al} (Å)	0.7230	0.5405	0.7833	1.0523
R_{Si} (Å)	0.8082	0.8913	0.7923	0.8924
R_{Zr} (Å)	0.8551	0.0529	0.9414	1.0305

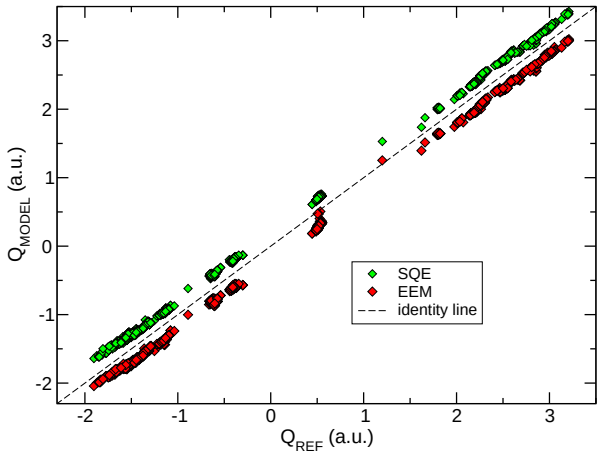

Figure 4.3: Correlation between the reference DFT HI charges and those predicted by the EEM and SQE models for molecules in the validation set. For the sake of clarity the EEM and SQE data were equally shifted down and up, respectively, along the y-axis.

might expect. One sees that the semi-empirical models largely overestimate charge transfer from metallic cations to the oxygen atoms making the systems more ionic than they are according to the results of iterative Hirshfeld analysis of quantum-chemical data.

Results of the SQE/HI model for the zirconium silicate and zirconia are worth a special remark. The model yields a charge for the Zr cation in the zirconium silicate that is larger than the formal ionic charge. The reason for such an artifact is the following. The Zr cations in the structure are eight-fold coordinated with four O atoms at a distance of 2.163 Å and four O atoms at a distance of 2.287 Å, whereas the SQE/HI parameters were calibrated on isolated systems having four-fold coordinated Zr cations. Consequently, the use of the bonding-specific parameters together with the explicit charge-transfer channels in the SQE scheme leads to an overestimation of the charge flow from the zirconium to oxygen atoms. In the case of zirconia the cations have four oxygens at a distance of 2.096 Å and four at 2.371 Å. The coordination number of the cations used in the SQE calculation was then taken to be four and the computed value of Zr atomic charge is below the formal ionic charge. Allowing the charge transfer also to the four next-nearest oxygens at the 2.371 Å distance leads to the zirconium atoms charge $q_{Zr} = 4.3212\ |e|$, that is, like in the case of $ZrSiO_4$, larger than the formal ionic charge of the cation. Note that the problem does not occur with the EEM scheme because it has only atom-based first-order terms in 4.2.3. A way of improving the SQE model is the use of distance-dependent bond electronegativities $\xi_{ji}(r_{ji})$, as it was suggested by Chen and Martínez [206]. Indeed, an attempt of introducing such a parameter for the Zr-O bond in the form

$$\xi_{ij}(r_{ij}) = \frac{\xi_{ij}}{1 + \exp(a(r_{ij} - b))} \tag{4.3.1}$$

with *ad hoc* parameters $a = 18.0$ Å^{-1} and $b = 2.38$ Å resulted in a decrease of charge of eight-fold

Table 4.4: Atomic charges in periodic systems obtained with different schemes.

Structure	EEM/HI	SQE/HI	DFT/HI	EEM/ESP	SQE/ESP	DFT/ESP
			O atoms			
α-quartz	−1.6417	−1.5750	−1.2840	−0.9800	−0.6833	−0.7741
α-cristobalite	−1.6281	−1.5715	−1.2762	−0.9634	−0.6839	−1.0512
JBW	−1.6605	−1.5902	−1.2933	−0.9681	−0.6924	−0.8167
	−1.6894	−1.6122	−1.2974	−0.9801	−0.7047	−0.7978
	−1.6114	−1.5620	−1.2846	−0.9588	−0.6742	−0.7622
	−1.6337	−1.5825	−1.3001	−0.9811	−0.6858	−0.8705
DFT	−1.6273	−1.5708	−1.2805	−0.9581	−0.6814	−0.8818
	−1.6941	−1.6184	−1.3081	−0.9877	−0.7093	−0.9121
	−1.5624	−1.5245	−1.2615	−0.9284	−0.6531	−0.7805
SOD	−1.6526	−1.5912	−1.2966	−0.9733	−0.6911	−0.8667
NPO	−1.6060	−1.5549	−1.2594	−0.9401	−0.6759	−0.6574
	−1.5641	−1.5239	−1.2591	−0.9204	−0.6549	−0.5393
t-ZrO2	−1.7224	−1.7095	−1.5542	−0.7721	−0.9574	−1.7357
ZrSiO4	−1.6363	−1.8660	−1.3914	−0.8453	−1.0227	−1.2971
			Si atoms			
α-quartz	3.2834	3.1500	2.5680	1.9600	1.3667	1.5482
α-cristobalite	3.2562	3.1430	2.5524	1.9268	1.3678	2.1024
JBW	3.2648	3.1522	2.5707	1.9020	1.3684	1.5716
	3.3373	3.1945	2.6118	2.0110	1.3869	1.6829
DFT	3.2555	3.1422	2.5654	1.9161	1.3626	1.7281
SOD	3.3052	3.1824	2.5932	1.9466	1.3822	1.7335
NPO	3.1701	3.0788	2.5185	1.8605	1.3308	1.1967
ZrSiO4	3.0416	3.1342	2.4133	1.8404	1.5285	2.1799
			Zr atoms			
t-ZrO2	3.4447	3.4190	3.1084	1.5442	1.9148	3.4714
ZrSiO4	3.5037	4.3298	3.1525	1.5409	2.5622	3.0086

coordinated Zr atoms in $ZrSiO_4$ to $q_{Zr} = 3.9821\ |e|$ without a significant change of charge of the four-fold coordinated zirconium in the tetragonal zirconia.

For the crystalline silica polymorphs both semi-empirical schemes give similar precision in the estimation of Hirshfeld-I charges in the periodic systems with the mean relative errors of 27 % and 22 % for the EEM/HI and SQE/HI models, respectively. It is not immediately clear why the transferability of the EEM/HI and SQE/HI to the periodic systems is only qualitative and not quantitative. There are several plausible explanations, some of which listed below:

1. The B3LYP XC functional in Gaussian03 package is slightly different from the B3LYP used in CRYSTAL06 code: the programs employ VWN3 and VWN5 LDA correlation functionals, respectively. However, we think that the difference can hardly account for the discrepancy between the semi-empirical and DFT HI charges.

2. Different basis sets are used for the periodic and cluster computations. Given a weak dependence of Hirshfeld-I charges on the basis set, this effect is expected to be small, although it may still contribute to the discrepancy.

3. The molecules in the training set are yet too small to capture the characteristics of atoms in the solid state. This explanation is the most plausible, as we show below, despite the fact that no notable sensitivity of the iterative Hirshfeld charges to the size of isolated molecules can be inferred from 4.3. The difference between the reference and model charges are so small that it is impossible to deduce some dependence of the errors on molecule size.

Corrections for periodic systems. Below we show how the parameters derived from the isolated systems in the training set can be corrected to accurately reproduce the DFT/HI charges in condensed phase, and how this correction can be rationalized. Let us consider the EEM model and write two equations (cf. 4.2.2)

$$\mathbf{q}_{\mathrm{EEM}} = -\mathbf{H}^{-1}\mathbf{x} \tag{4.3.2}$$

$$\mathbf{q}_{\mathrm{DFT/HI}} = -\mathbf{H}^{*-1}\mathbf{x}^{*}, \tag{4.3.3}$$

where we introduced $\mathbf{H}^*$ and $\mathbf{x}^*$ as (hypothetical) more accurate descriptions of the hardness matrix and the electronegativity vector, respectively, in periodic systems, which result exactly in the DFT Hirshfeld-I charges. The last equation can be rewritten so that the $\mathbf{H}^*$ matrix also includes the difference between the $\mathbf{x}$ and $\mathbf{x}^*$ vectors, [239] i.e.

$$\mathbf{q}_{\mathrm{DFT/HI}} = -\mathbf{H}^{*-1}\mathbf{x}. \tag{4.3.4}$$

It needs to be validated how the hardness matrix needs to be adapted to get more accurate predictions for the charges in periodic systems. Let us first assume that the main difference between the $\mathbf{H}$ and $\mathbf{H}^*$ matrices is due to the diagonal elements. We will further test if this

Table 4.5: Elements of **G** matrix (cf. 4.3.7) computed for the oxygen and silicon atoms in crystalline silica structures using the parameters of EEM/HI scheme.

Structure	G_{O} (eV)	G_{Si} (eV)
α-quartz	0.8827	0.3761
α-cristobalite	0.8196	0.4215
NPO	0.6264	0.4729
SOD	0.7388	0.4313

assumption is consistent with the observed errors. 4.3.2 and 4.3.3 then become

$$\mathbf{q}_{\mathrm{EEM}} = -\mathbf{H}^{-1}\mathbf{x} \tag{4.3.5}$$

$$\mathbf{q}_{\mathrm{DFT/HI}} = -(\mathbf{H}+\mathbf{G})^{-1}\mathbf{x}, \tag{4.3.6}$$

where **G** is a diagonal matrix of corrections to the hardness matrix that should reduce the observed errors. Straightforward elimination of **x** from both equations can be used to estimate the corrections to the diagonal elements of the EEM hardness matrix **H**:

$$\mathbf{H}(\mathbf{q}_{\mathrm{EEM}} - \mathbf{q}_{\mathrm{DFT/HI}}) = \mathbf{G}\mathbf{q}_{\mathrm{DFT/HI}}. \tag{4.3.7}$$

Table 4.5 reports elements of the **G** matrix for the Si and O atoms in some of silica structures of Table 4.2. Clearly, the values are structure dependent, but the difference is not very large and we can compute mean values which are $G_O = 0.7668$ (0.0958) eV and $G_{Si} = 0.4216$ (0.0354) eV. Making use of these values, we now calculate the HI charges of atoms in all our silica polymorphs by modifying the diagonal elements of the hardness matrix. The results are gathered in Table 4.6 that shows a nice performance of EEM/HI model modified for the solid state computations ($\mathrm{EEM_S}$/HI column): the mean relative error for the charges on oxygens drops from 27 % to 1 %. Interestingly, the use of the same G_{ii} parameters in the SQE/HI scheme leads to similar spectacular improvement of HI charges whose mean relative deviation from the reference DFT values decreases to 1 % (Table 4.6, $\mathrm{SQE_S}$/HI column).

The above results can be rationalized in the following way. Table 4.5 shows that G_{ii} parameters are *positive*, which represents an increase of the effective hardness of atoms in solids as compared to atoms in molecules. The inverse of the hardness is the softness, which is related to the polarizability [221, 241]. Therefore the increase of atomic hardness needed to mimic HI charges in solids, can be interpreted as an evidence of *decrease of polarizability* of atoms, when going from molecules to solids. Indeed, the crystalline field in solids confines the electrons, which then have a reduced ability to respond to an external perturbation in comparison to a molecule. A generalization of the above correction scheme can provide a way of obtaining EEM and SQE charge-based models transferable among a large palette of systems of different size and densities.

Although this correction scheme is very effective, it is clear that this problem must be analyzed

Table 4.6: Hirshfeld-I atomic charges computed with the molecular EEM/HI and SQE/HI models, and with the models corrected for solid state calculations (EEM_S/HI and SQE_S/HI); the right-hand column reports the reference Hirshfeld-I charges obtained in periodic DFT computations (DFT/HI).

Structure	EEM/HI	EEM_S/HI	SQE/HI	SQE_S/HI	DFT/HI
			O atoms		
α-quartz	−1.6417	−1.2739	−1.5750	−1.2956	−1.2840
α-cristobalite	−1.6281	−1.2881	−1.5715	−1.2932	−1.2762
JBW	−1.6605	−1.2869	−1.5902	−1.3068	−1.2933
	−1.6894	−1.3081	−1.6122	−1.3239	−1.2974
	−1.6114	−1.2518	−1.5620	−1.2846	−1.2846
	−1.6337	−1.2663	−1.5825	−1.2998	−1.3001
DFT	−1.6273	−1.2644	−1.5708	−1.2927	−1.2805
	−1.6941	−1.3141	−1.6184	−1.3293	−1.3081
	−1.5624	−1.2185	−1.5245	−1.2570	−1.2615
SOD	−1.6526	−1.3153	−1.5912	−1.3065	−1.2966
NPO	−1.6060	−1.2544	−1.5549	−1.2833	−1.2594
	−1.5641	−1.2263	−1.5239	−1.2595	−1.2591
			Si atoms		
α-quartz	3.2834	2.5478	3.1500	2.5912	2.5680
α-cristobalite	3.2562	2.5762	3.1430	2.5864	2.5524
JBW	3.2648	2.5319	3.1522	2.5908	2.5707
	3.3373	2.5880	3.1945	2.6249	2.6118
DFT	3.2555	2.5307	3.1422	2.5858	2.5654
SOD	3.3052	2.6306	3.1824	2.6131	2.5932
NPO	3.1701	2.4807	3.0788	2.5428	2.5185

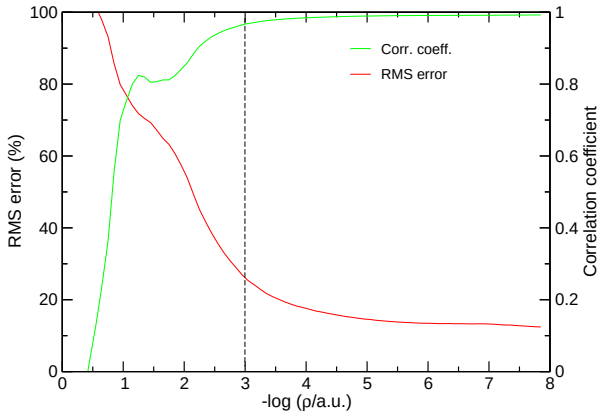

Figure 4.4: Relative error in ESP values and correlation coefficient between reference and model ESP values *vs* electronic density for molecules in validation set; results obtained with SQE/ESP parameters of 4.3. Vertical dashed line corresponds to the threshold density value.

in more detail in future work. For example, it is not yet clear how this correction scales from zero to the values given above when going from the isolated molecules in the training set to infinite periodic systems, i.e. starting from which size the system cannot be considered anymore as isolated.

4.3.3 Electrostatic potential

Isolated systems As it was mentioned above, the EEM/ESP scheme is useful only for computing the static characteristics and therefore, the following presentation and discussion are focused on the ESP calibrated SQE model that is capable of modeling both static and response ESP values.

Initially we investigate how well the SQE/ESP model reproduces the reference ESP values from the DFT computations on the isolated systems. 4.4 shows the relative RMS error and the correlation coefficient between the SQE/ESP predictions and DFT/ESP reference data as function of the electronic density computed in the same point. One sees that the model reproduces the ESP around the molecules with the minimum mean error of ca 13 % and with the value of 0.99 for correlation coefficient. One observes a significant increase of the error accompanied by a decrease of the correlation coefficient for the values of the density above 10^{-3} a.u. This value can therefore be taken as a threshold beyond which the results obtained with the SQE/ESP model become unreliable. These results can be understood by analyzing the Poisson equation (in atomic units),

$$\nabla^2 V(\mathbf{r}) = -4\pi\rho_{\text{total}}(\mathbf{r}), \tag{4.3.8}$$

where V is the electrostatic potential and ρ_{total} represents the *complete* charge density. In the case of a DFT computation, ρ_{total} consists of the nuclear plus electronic charge density. In the charge model used for the ESP fit, ρ_{total} just contains the Gaussian charge densities from 4.2.8. Only

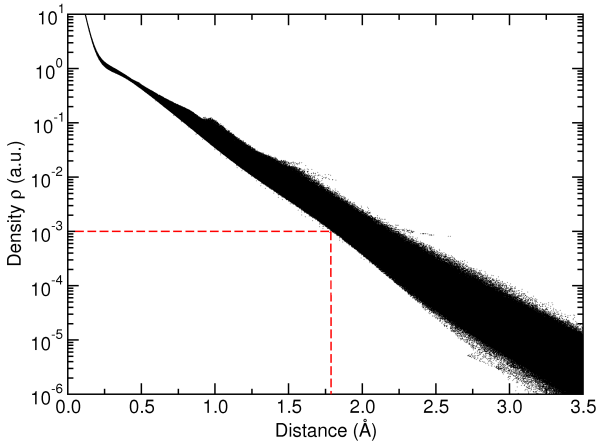

Figure 4.5: Plot of electronic density value in a given point as a function of distance from closest oxygen atom. Dashed lines indicate coordinates for $\rho = 10^{-3}$ a.u.

at the points **r** where the right-hand side of the Poisson equation is almost the same in both the DFT description and the charge model, one can expect the charge model to work. This condition is fulfilled at those distances from the nuclei, where the nuclear charge screened by the electronic density is well approximated by the spherical Gaussian charge distribution, 4.2.8, i.e. sufficiently far from the nuclei. For that reason, one can only use grid points "outside" the molecule, when computing ESP-fitted charges.

Singh and Kollman [242] found that grid points for fitting ESP charges need to be chosen at shells of at least 1.2 times of atomic Van der Waals radii, which is a quick method for selecting points where the electron density is low. The correlation between the distance from the oxygen atoms and the electron density in a given point is shown in 4.5. This plot is based on data from all molecules in the training and validation set. Interestingly, the threshold density value of 10^{-3} roughly corresponds to the distance of 1.8 Å that is ca. 1.2 times the Van der Waals radius of an O atom [243], in agreement with the finding by Kollman *et al.* [242]. The spread of the electron density for a given distance, however, shows that it is safer to rely on the actual electron density (instead of fixed radii) to determine the grid points for the ESP fitting procedure. It is also questionable to what extent such radii are transferable between different oxidation states of a given element. The volume of an atom, and hence also its radius, depends on the population, which can only be deduced once the charges are fitted.

It is also of interest to test the ability of the SQE/ESP model to mimic other properties than the ones used in the calibration procedure, i.e charges and the electrostatic potentials. Two such properties were chosen: the dipole moment and the dipole polarizability [162]. 4.6 and 4.7 present the comparison of the dipole moment components and of the principal components of the dipole polarizability tensor computed with SQE/ESP model with those obtained in the quantum-chemical calculations for molecules of the validation set. The agreement between the model predictions and the reference data is very satisfactory. Again, the current SQE/ESP model

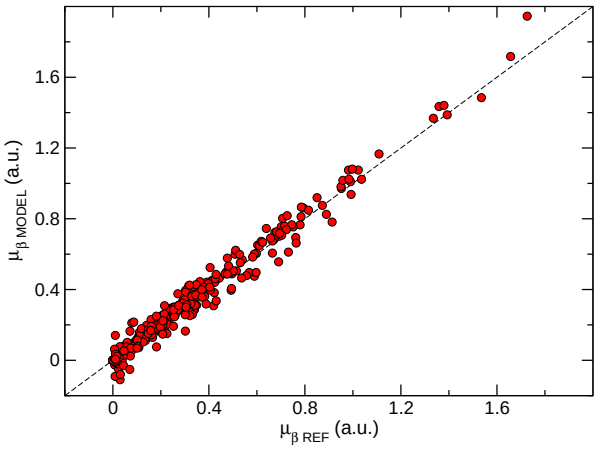

Figure 4.6: Comparison of the dipole moment components for the molecules in the validation set, obtained with the reference DFT method and with the SQE/ESP model.

uses spherically symmetric atomic density basis and consequently, the model is not capable of predicting the out-of-plane polarizability component of plane molecules (points on the x-axis in 4.7).

Periodic models. The transferability of the SQE/ESP parameters to periodic systems was tested by comparing the electrostatic potential on a cubic grid computed with the SQE/ESP model and the DFT computations for the structures listed in Table 4.2. According to the results reported above, the ESP values should be comparable in those points of the 3D grid where the value of the electronic density does not exceed the threshold 10^{-3} a.u. This issue, however, complicates the comparison in dense structures because only a relatively small number of points fulfill the criterion. 4.8 (left pannel) shows those volumes in the crystallographic unit cell of α-quartz structure, where the value of the electronic density is below the threshold value $\rho = 10^{-3}$ a.u. One sees that only a small part of the volume (ca. 5 %) can be used to compute the values of the ESP for the structure. Consequently, all dense structures (t-ZrO_2, $ZrSiO_4$, α-cristiobalite) provide a hard test for the predictive power of the model. The situation is markedly different for zeolitic structures. Thus, the electronic density is below the threshold value in ca. 36 % of grid points in the sodalite structure (4.8, right pannel). Therefore, to asses the transferability of the SQE/ESP model we have chosen to present two extreme cases of the most and least dense structures: α-quartz and sodalite (SOD), respectively (cf. Table Table 4.2).

4.9 and 4.10 present the relative error and the correlation coefficient for the ESP values computed as a function of the electronic density in the α-quartz and sodalite structures. Despite the difference in the number of points employed in the electrostatic potential calculations, the ESP is reproduced in both structures equally well. The relative error does not exceed 20 % with the correlation coefficient above 0.98, at the grid points, where the electronic density value remains below the threshold 10^{-3}. These confidence values are close to those obtained for isolated molecules and

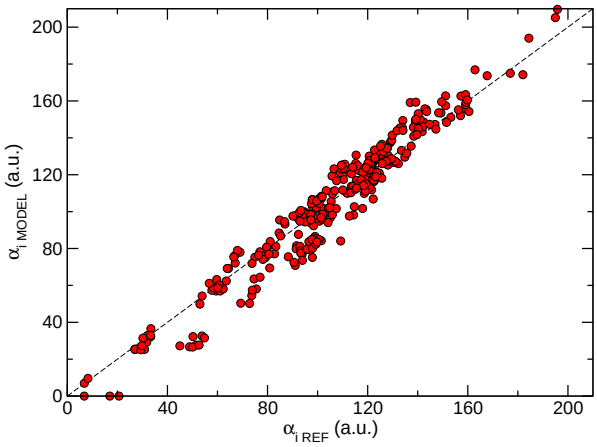

Figure 4.7: Comparison of all principal components α_i of the polarizability tensors for the molecules in the validation set, obtained with the reference DFT method and with the SQE/ESP model.

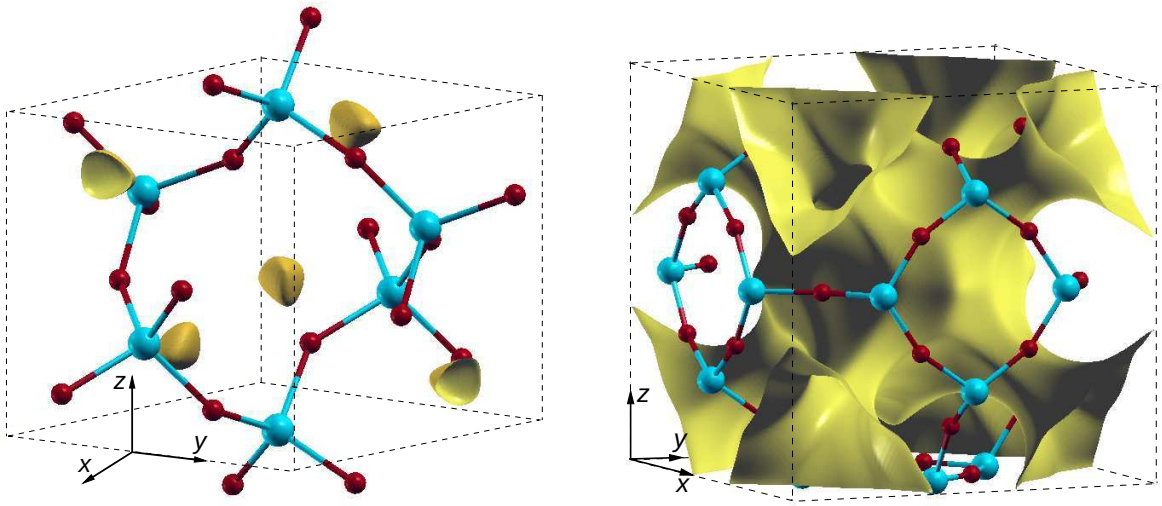

Figure 4.8: Left: image of α-quartz structure showing the region, where electronic density value is below 10^{-3} a.u. (in yellow color); silicon and oxygen atoms are shown as cyan and red balls, respectively. Right: image of sodalite unit cell with the isodensity surface corresponding to $\rho = 10^{-3}$ a.u. The images were generated with XCrySDen program [244].

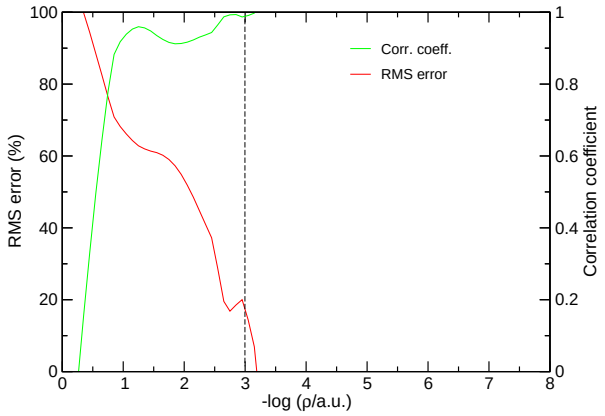

Figure 4.9: Relative error in ESP values and correlation coefficient between reference and model ESP values *vs* electronic density in α-quartz structure; results obtained with SQE/ESP parameters of 4.3. Vertical dashed line corresponds to the threshold density value $\rho = 10^{-3}$ a.u.

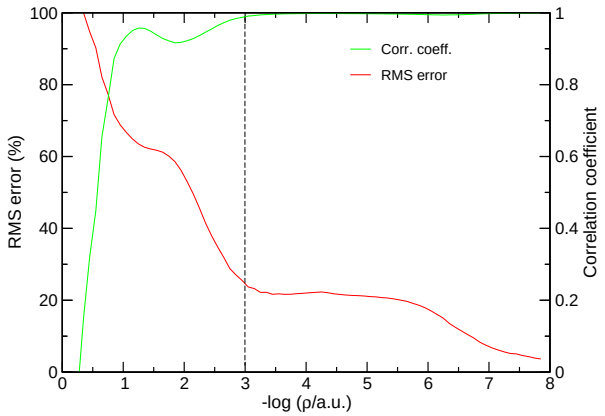

Figure 4.10: Relative error in ESP values and correlation coefficient between reference and model ESP values *vs* electronic density in sodalite structure; results obtained with SQE/ESP parameters of 4.3. Vertical dashed line corresponds to the threshold density value $\rho = 10^{-3}$ a.u.

indicate a good transferability of the SQE/ESP parameters from isolated to periodic systems.

4.11 presents the ESP and the ESP gradient obtained for the sodalite structure along the $<100>$ direction (lines connecting the centers of opposite 4R rings) in the DFT computation and with the SQE/ESP model. The agreement between the semi-empirical and *ab initio* data is remarkable. It is noteworthy that the electronic density in the regions with the coordinate less than 1 Å and more than 8 Å is above the threshold value 10^{-3} and thus, the SQE results in this regions can not be considered as reliable.

Because the SQE/ESP model successfully reproduces the ESP derived from the DFT computations, one would expect that the SQE/ESP charges correlate with ESP fitted charges from the DFT computations. The right-hand part of 4.4 reports ESP charges for the periodic structures predicted with the EEM/ESP and SQE/ESP schemes and compares the values with ESP charges

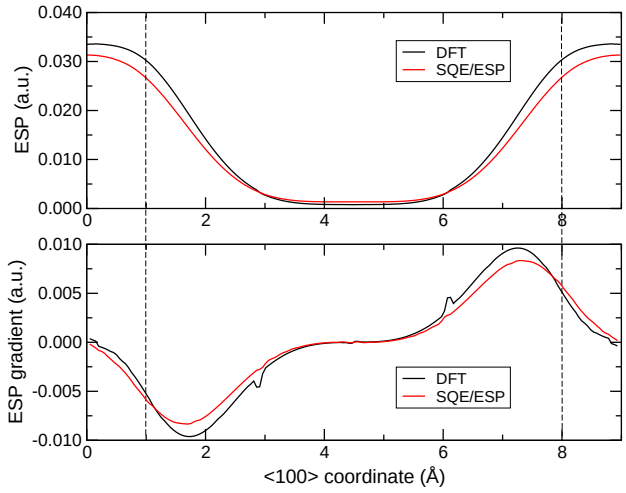

Figure 4.11: ESP (upper panel) and ESP gradient (lower panel) obtained in periodic DFT and SQE/ESP computations along the $<100>$ direction in sodalite structure (lines connecting the centers of opposite 4R rings). The electron density in the region between the dashed lines is below the threshold value $\rho = 10^{-3}$ a.u.

derived from the DFT calculations. The EEM model generally overestimates, while the SQE one underestimates the ESP charges. It is notable that both models do reproduce the trends in the magnitude of charges from one structure to another. There are, however, a few difficult cases. Thus, the variation of ESP charges predicted by SQE model between the atoms in the α-quartz and α-cristobalite structures is significantly smaller than that obtained from the DFT results. Furthermore, one can note that the EEM scheme fails to reproduce the ESP charge distribution in the zircon polymorph of $ZrSiO_4$: the EEM/ESP charge on the Si atom is larger than that on Zr one, whereas the DFT/ESP charges show the inverse trend. Morover, the decrease of the oxygen charge as compared to the pure silica is absent in the EEM/ESP model, while the SQE/ESP predicts this trend qualitatively.

There are large discrepancies in the comparison of ESP fitted charges and the corresponding values computed with the SQE/ESP and EEM/ESP models. However, such differences do not mean that the results of ESP-based models are unreliable. It is well known that large changes in atomic charges can lead to only small changes in the ESP surrounding the atoms. This also means that two sets of charges that give a fair reproduction of the ESP may be manifestly different [245]. This lack of sensitivity of ESP fitted charges results in unpredictable contributions, i.e. apparent as noise on the charges, which makes it hard to use them for a direct comparison [212]. The right-hand part of 4.4 contains three sets of charges that attempt to give a good description of the electrostatic potential in the interstitial regions of crystalline structures, but that does not imply these charges should be equal or show the same subtle trends.

Recently, Campanã and co-workers proposed a method for generating ESP charges for periodic systems from results of periodic-quantum-chemical calculations [246]. The authors noted that in

the calculations employing the plane-wave basis set and pseudo-potentials, the ESP values are defined up to a constant offset and thus the cost function should be based on the difference of potential values in the grid points rather than on the values themselves. Making use of their approach and of different plane-wave codes (VASP, CPMD, and SIESTA), Campanã *et al.* reported ESP charges for the Si atoms of sodalite structure in the range 1.151 to 1.389 $|e|$. These values have the same order of magnitude as $q_{\mathrm{Si}} = 1.7335\ |e|$ obtained in our all-electron DFT calculations with localized basis set; probably by chance the SQE/ESP predicted charge $q_{\mathrm{Si}} = 1.3824\ |e|$ nicely fits the data of ref. [246].

Magnitude of Hirshfeld-I charges. A general observation in the Hirshfeld-I results is that the absolute values of HI charges are significantly larger than the corresponding absolute values of the ESP fitted charges. This finding is markedly different from earlier studies where a reasonable correspondence between HI and ESP fitted charges was found for a set of organic molecules [224]. We suspect that this is a deficiency inherent to the Hirshfeld-I procedure when applied to (nearly) ionic systems. The oxygen charges in the oxide clusters are mostly between -1.0 and -2.0 $|e|$, which means that the corresponding proatoms in the Hirshfeld-I scheme are a linear interpolation between the isolated oxygen anion and dianion. The density profiles of these anionic pro-atoms are mostly determined by the limitations of the 6-311G+(d,p) basis set, which may lead to artifacts in the Hirshfeld-I partitioning. In contrast, the charges on oxygen atoms in organic molecules are the range from 0 to -1.0 $|e|$ and therefore the basis set limitations are expected to be less important for constructing the pro-atoms [195]. This issue should be analyzed in detail in future work, potentially leading to an improved Hirshfeld-I scheme that also gives ESP-quality charges for oxides and ionic systems.

4.4 Conclusions

An extensive parametrization of the electronegativity equalization model (EEM) and split-charge equilibration (SQE) model was performed for silicate materials on the basis of quantum-chemical calculations of oxide clusters containing aluminum, silicon, and zirconium cations. The calibration of the parameters in these models was done using the iterative Hirshfeld (HI) charges and the electrostatic potential (ESP) as the reference quantities. The total cost function used in the non-linear least squares minimization procedure was a linear combination of a static cost function based on HI charges or ESP grid data, and a response cost function that included changes of these reference quantities upon an applied external electric field. The transferability of parameters was assessed by a comparison with HI charges and ESP grid data for a validation set of isolated molecules and a number of crystalline structures.

The outcome of the parametrization procedure allows us to conclude that the EEM model is capable of mimicking static characteristics only, while it fails to reproduce the response of the electronic distribution and ESP to an external electric field. The SQE model performs well for both static and response properties and also provides correct results for properties that were not

explicitly included in the parameterization procedure, i.e. the dipole moment and the dipole polarizability.

Both the EEM and SQE calibrations can be used for a fast and reliable calculation of HI charges in isolated molecules that were not included in the training set, but they reveal a limited transferability to periodic systems. It is however possible to propose a correction to the atomic hardness parameters of each element based on the differences between the EEM/HI and the DFT/HI charges for the periodic systems. This correction amounts to an increase of the atomic hardness in the solid state, which can be related to decrease of atomic polarizability due to the confinement by the crystal field. These corrections were found to be transferable between the EEM and SQE models. The corrected EEM/HI and SQE/HI schemes reproduce the reference DFT HI charges in the periodic systems with the mean relative error of less than 2 %.

The ESP-based parametrizations were found to provide reliable results only in those regions of space, where the electronic density values are lower than a threshold value $\rho = 10^{-3}$ a.u. The density criterion was found to be consistent with the distance criterion used in the Merz-Kollman ESP charge fitting scheme for a quick selection of grid points. The SQE/ESP model shows a good transferability from molecular to periodic structures, if the regions for computing ESP were chosen according to the criterion above. The EEM/ESP scheme was found to perform the worst among all models studied and hence the use of EEM method for computing the ESP-related characteristics should be avoided.

As an indirect result of this work, we observed very large differences between ESP fitted charges and iterative Hirshfeld charges. Future work should clarify the origin of the overestimation of the charge transfer in partially ionic systems by the iterative Hirshfeld scheme. Ideally, such work could lead to an improved iterative Hirshfeld scheme, whose charges satisfactorily approximate the electrostatic potential of silica clusters.

Chapter 5

Pairwise dispersion interaction in oxides with maximally localized Wannier functions.*

5.1 Introduction.

Despite a spectacular progress in the atomistic simulations of complex systems from the first principles during the last two decades, the models of effective potentials (force fields) still remain an indispensable tool for the understanding of the behaviour of many complex systems at the atomic level. Nowadays, the determination of parameters of potential functions is commonly done with the use of *ab initio* quantum-chemical calculations [203, 208, 247–249]. Such calculations can readily provide the data necessary for the parametrization of short-range bonding potentials and also yield, by using specific electron density partitioning scheme [250], the effective atomic charges for describing the long-range electrostatic interactions. On the other hand, the parameterization of dispersion interactions remains a challenging task and is often performed on an empirical basis.

The problem stems from the fact that the dispersion interactions are due to the instantaneous fluctuations of the electron density and are a part of the long-range correlation energy. The treatment of the correlation energy with post Hartree-Fock methods is very prohibitive, even for relatively small systems. This energy is also lacking in the calculations with DFT methods, which are capable of treating large systems at modest computational cost, using local and semi-local exchange-correlation functionals. To remedy the problem several approaches have been proposed [204, 251–257]. Some of them use (semi)empirical corrections to the conventional DFT functionals [251,252], while other methods employ the self-consistent electron density to obtain the dispersion energy contribution [204,253–257]. These and other methods as well as the perspectives of the inclusion of the dispersion interactions in the DFT calculations have recently been reviewed in [258–260].

*S. V. Sukhomlinov and K.S. Smirnov. Structure-dependent interatomic dispersion coefficients in oxides with maximally localized Wannier functions. *J. Phys.: Condens. Matter* 24, 475501 (2012).

Regardless the used approach the dispersion energy E_D between two interacting entities (atoms, molecules) separated by distance r is commonly represented in the form

$$E_D = -\frac{C_6}{r^6} - \frac{C_8}{r^8} + \ldots, \tag{5.1.1}$$

where C_6, C_8, etc. are the dispersion coefficients whose values, and in particular that of the leading r^{-6} term, just need to be determined. From the viewpoint of force field development, the method proposed by Silvestrelli [204,255] for computing the C_6 coefficients is of particular interest. The approach uses maximally localized Wannier functions (MLWFs) to decompose the electron density of system into localized fragments and then it applies the model by Andersson *et al.* [261] to compute the pairwise fragment-fragment dispersion interaction energies. Thus, the Silvestrelli approach effectively recasts the many-body dispersion energy into a pairwise form and since it makes use of the self-consistent electron density, the resulting pairwise energy contributions reflect the environment of the interacting density fragments and depend on the structure of the system. The method was successfully applied to various systems, where the dispersion interactions play an important role [262–264], and it was shown to be simple, efficient, and accurate.

The use of the Silvestrelli approach for the calibration of the dispersion parameters of a force field model is, however, hampered by the fact that the method operates the MLWFs as the interacting entities, whereas the force fields commonly employ interaction centers localized on the atoms. Nonetheless, for some systems, where the Wannier function centers are close to the atomic positions, one might attempt to assign the MLWFs to the nearest atom and then to compute the dispersion interactions in an *atom pairwise* manner. This way was used by Salanne and co-workers in their development of force field for alkali halides and aqueous ions [265, 266].

Force field models for ionic and semi-ionic solids commonly employ the C_6/r^6 term to represent the dispersion interaction energy in the systems [247, 267]. The importance of the dispersion energy for the proper description of the relative stability of oxide polymorphs has recently been highlighted by Conesa [268]. The present paper reports results of using the Silvestrelli method for the computation, and of analysis of the interatomic C_6 dispersion coefficients in oxide materials and the study aims at answering the following questions:

(i) Can the original Silvestrelli model be modified to operate with atom-related parameters rather than with the characteristics of MLWFs?

(ii) How do the dispersion coefficients vary in different oxide systems? Which system characteristics are of importance for the coefficient values?

(iii) How are the quantities, which determine the dispersion interactions, related to other parameters of force field model, in particular to the atomic charges playing a special role in the polarizable force fields.

The Chapter is organized as follows. The next section provides the theoretical background of the Silvestrelli approach and describes a modification of the method that allows, for the systems

of interest, computing the dispersion coefficients with the help of atomic parameters. Section 5.3 presents the systems studied in the work and details of DFT calculations. The subsequent section reports and discusses the obtained results. It first presents findings that validate the modification suggested in Section 5.2 and then discusses values of computed dispersion coefficients and their relation to the structural characteristics systems, such as the mean nearest-neighbour distance and the coordination number of atoms. The last part of Section 5.4 presents relation between the quantities determining coefficients and atomic charges. The final part of the manuscript provides conclusions of the work. Additional information on the studied systems that might be of interest for the reader is given in Appendix D. Atomic units are used throughout the Chapter if otherwise is not stated explicitly.

5.2 Theory.

5.2.1 Maximally localized Wannier functions.

To partition the total electron density of the system into localized fragments the Silvestrelli approach uses the maximally localized Wannier functions $\{w_n\}$, that are constructed by applying an unitary transformation $\mathbf{U}$ to the occupied Kohn-Sham orbitals $\{\phi_m\}$

$$|w_n\rangle = \sum_{m=1}^{N_{occ}} U_{nm}|\phi_m\rangle \tag{5.2.1}$$

such that the MLWFs minimize the functional Ω defined as

$$\Omega = \sum_{n}^{N_{occ}} \left(\langle w_n|r^2|w_n\rangle - (\langle w_n|\mathbf{r}|w_n\rangle)^2\right) = \sum_{n}^{N_{occ}} \left(\langle r_n^2\rangle - \bar{\mathbf{r}}_n^2\right) = \sum_{n}^{N_{occ}} S_n^2 \tag{5.2.2}$$

with S_n being the spread of the n-th Wannier function. It is commonly assumed that the MLWFs exponentially decay in the real space and that each Wannier function can be represented by a hydrogen-like normalized orbital

$$w_n(\mathbf{r} - \bar{\mathbf{r}}_n) = \frac{3^{3/4}}{\sqrt{\pi}\, S_n^{3/2}} e^{-\sqrt{3}|\mathbf{r}-\bar{\mathbf{r}}_n|/S_n} \tag{5.2.3}$$

with the Cartesian component ξ_n ($\xi_n = x_n,\, y_n,\, z_n$) of the Wannier function center (WFC) $\bar{\mathbf{r}}_n$

$$\xi_n = -\frac{L}{2\pi}\mathrm{Im}\ln\langle w_n|e^{-\imath 2\pi\xi/L}|w_n\rangle. \tag{5.2.4}$$

In the calculations of isolated systems using the localized basis set the Wannier function localization procedure corresponds to determining Boys orbitals [269]. In the calculations of periodic systems, which commonly employ the periodic boundary conditions and the delocalized plane-wave basis set, the MLWFs allow the spatial partitioning the electron density in an unambiguous

way. We refer to [270] for a recent review on the theoretical aspects and applications of MLWFs.

5.2.2 Dispersion coefficient and MLWFs.

As it was mentioned in the Introduction, the leading term of the dispersion energy E_{Dnl} for the interaction of two spherically symmetric entities n and l separated by distance r_{nl} is given by

$$E_{Dnl} = -\frac{C_{6nl}}{r_{nl}^6}, \tag{5.2.5}$$

where C_{6nl} stands for the dispersion coefficient. The right-hand part of the expression (5.2.5) is often multiplied by a damping function $f(r_{nl})$ that prevents the dispersion energy from the unphysical divergence at short distances [204, 251–255, 271].

Andersson, Langreth, and Lundqvist [261] have shown that the dispersion coefficient C_{6nl} of two non-overlapping density fragments $\rho_n(\mathbf{r})$ and $\rho_l(\mathbf{r})$ can be calculated as

$$C_{6nl} = \frac{3}{32\pi^{3/2}} \int_V d\mathbf{r} \int_{V'} d\mathbf{r}' \frac{\sqrt{\rho_n(\mathbf{r})\rho_l(\mathbf{r}')}}{\sqrt{\rho_n(\mathbf{r})} + \sqrt{\rho_l(\mathbf{r}')}}. \tag{5.2.6}$$

Based on the result by Andersson *et al.*, Silvestrelli [204] has suggested to use the MLWFs as such density fragments and, with $\rho_n(\mathbf{r}) = |w_n(\mathbf{r})|^2$, the formula (5.2.6) for the dispersion coefficient can then be rewritten as

$$C_{6nl} = \frac{3}{32\pi^{3/2}} \int_{|\mathbf{r}|<r_c} d\mathbf{r} \int_{|\mathbf{r}'|<r_c'} d\mathbf{r}' \frac{|w_n(\mathbf{r})||w_l(\mathbf{r}')|}{|w_n(\mathbf{r})| + |w_l(\mathbf{r}')|}, \tag{5.2.7}$$

where

$$r_c = \sqrt{3}\, S_n\big(0.769 + \frac{1}{2}\ln S_n\big) \quad \text{and} \quad r_c' = \sqrt{3}\, S_l\big(0.769 + \frac{1}{2}\ln S_l\big) \tag{5.2.8}$$

are the cutoff radii chosen by equating the length scale for density change to the electron gas screening length [204]. By substituting the analytical representation of MLWFs (5.2.3) into (5.2.7) one obtains the computationally tractable expression for the C_{6nl} coefficient [204, 255]

$$C_{6nl} = \frac{S_n^3 S_l^3}{2 \cdot 3^{5/4}} \int_0^{x_c} dx\, x^2 e^{-x} \int_0^{y_c} dy\, \frac{y^2 e^{-y}}{S_l^{3/2} e^{-x} + S_n^{3/2} e^{-y}}, \tag{5.2.9}$$

where

$$x_c = \sqrt{3} r_c / S_n, \qquad y_c = \sqrt{3} r_c' / S_l. \tag{5.2.10}$$

If each MLWF is occupied by two electrons, as in the case of spin-degenerate systems, the density of each fragment in (5.2.6) should be multiplied by 2 and consequently, the value of C_{6nl} coefficient in (5.2.7) and (5.2.9) must then be scaled by $\sqrt{2}$.

5.2.3 Localization of MLWFs and character of chemical bond in system.

It is instructive to analyse the position of the Wannier function centers with respect to the nuclear positions in the systems with different character of chemical bond. Figure 5.1 shows the positions of WFCs in two such systems: crystalline silicon and silicon dioxide. In the silicon structure (Figure 5.1a) the WFCs are situated between the Si atoms, thus reflecting the sharing of valence electrons by atoms in the covalent Si-Si bonds. The situation becomes different for the systems with (partially) ionic character of bonding, such as silica (Figure 5.1b), where the positions of the WFCs are close to the positions of the oxygen atoms that corresponds to the picture of (partially) ionic character of Si-O bonds with the electron density transferred toward the electronegative oxygen atoms.

Figure 5.1: Position of Wannier function centers in crystalline silicon (a) and in amorphous silica (b) (fragments of structures are shown). The silicon and oxygen atoms, and Wannier function centers are the wheat, red, and small green balls, respectively.

Let us now analyse Figure 5.1 from the perspective of computing the dispersion energy in an atom pairwise manner. The position of the WFCs midway between the atoms in the covalent systems (Figure 5.1a) does not allow assigning the MLWFs to either of two bonded atoms. Such a situation hampers the use of the MLWFs in designing an atom-based scheme for the dispersion interactions, whereas it is, of course, possible to compute the total dispersion energy of the system by the approach [204]. The situation could, perhaps, be improved by using partially occupied MLWFs [272, 273]. In this approach, the set of occupied Kohn-Sham orbitals in (5.2.1) is complemented by suitably chosen unoccupied states and an extended set of MLWFs is obtained by using the disentanglement method [274]. The scheme generally improves the localization and symmetry properties of the Wannier functions and, as the number of the MLWFs is more than N_{occ}, the functions have partial occupancies. To take the partial occupancies of the MLWFs into account while computing the dispersion coefficient by (5.2.9), the equation has to be modified as suggested by Andrinopoulos and co-workers [264]. However, despite the improved localization and symmetry, the use of the partially occupied MLWFs does not necessarily lead to shifting the WFCs to the atoms. Thus, the comparison of Figure 3 in Ref. [273] with Figure 5.1a shows that the method does not displace MLWFs accounting for the bonding σ orbitals centered on the Si-Si

bonds. The same conclusion can be drawn from the analysis of Figures 6, 10 and 12 of Ref. [264].

On the other hand, Figure 5.1b shows that the positions of the Wannier function centers in the ionic systems are close to the atoms. This feature can be used to represent the dispersion energy of the system in the pairwise form

$$E_D = -\sum_i \sum_{j>i} \frac{C_6^{ij}}{r_{ij}^6}, \tag{5.2.11}$$

where the indexes i and j denote the interacting atoms and C_6^{ij} are *effective interatomic* dispersion coefficients defined as [265, 266]

$$C_6^{ij} = \sum_{n\in i} \sum_{l\in j} C_{6nl}, \tag{5.2.12}$$

with n and l being the Wannier functions localized on the atoms i and j, respectively. Note that the coefficient C_6^{ij} in (5.2.12) depends only on the spreads of MLWFs localized on the atoms i and j, but not on the WFC positions. In the subsequent part, we suggest a way permitting to condense the dependence to atom-related parameters.

5.2.4 Effective atomic orbitals.

In an analogy with the C_{6nl} coefficients (5.2.9) computed via the density fragments described by the MLWFs, one can consider the atomic dispersion coefficients C_6^{ij} as due to the interaction of two electron densities $\rho_i(\mathbf{r})$ and $\rho_j(\mathbf{r})$ centred at the atoms i and j, respectively. The density $\rho_i(\mathbf{r})$ is given by $\rho_i(\mathbf{r}) = |\varphi_i(\mathbf{r})|^2$ and the effective atomic orbital $\varphi_i(\mathbf{r})$ is supposed to have the same form as the Wannier functions $w_n(\mathbf{r})$, i.e. the hydrogen-like normalized orbital (5.2.3). The effective orbital $\varphi_i(\mathbf{r})$ is completely characterized by its spread S_i^* and this quantity can be obtained from the numerical solution of (5.2.9) letting $l = n$. Indeed, the dependence of C_6^{nn} on the spread S_n can be fitted (correlation coefficient of 1.000) with the function

$$C_6^{nn} = p_0 \left(S_n - p_1\right)^q \tag{5.2.13}$$

with the parameters $p_0 = 2.05009$, $p_1 = 0.34521$, and $q = 5.03082$. With the use of (5.2.13), the spread S_i^* of the effective orbital $\varphi_i(\mathbf{r})$ is given by

$$S_i^* = p_1 + \left(S_6^{ii}/p_0\right)^{1/q}, \tag{5.2.14}$$

where S_6^{ii} stands for a self-atom dispersion coefficient computed by (5.2.12) for $j = i$. The dispersion coefficient C_6^{ij} for two different atoms i and j can then be calculated by feeding the spreads S_i^* and S_j^* into (5.2.9).

Provided that the effective orbital model proves to be viable, the model is expected to simplify the computation of the dispersion coefficients C_6^{ij} and it should allow the analysis of factors influencing values of the coefficients with the use of the *atomic* parameters S_i^* and S_6^{ii}.

Combination rules. Alternatively, given the known values of the self-atom dispersion coefficients S_6^{ii} the interatomic coefficients C_6^{ij} can be derived using combination rules. Thus, one of widely used combination rules for unlike dispersion coefficients reads [275]

$$C_6^{ij} = 2\frac{\alpha_i \alpha_j \, S_6^{ii} S_6^{jj}}{S_6^{ii}\alpha_j^2 + S_6^{jj}\alpha_i^2} \tag{5.2.15}$$

with α_i and α_j being the static dipole polarizability of atoms i and j, respectively. Alternatively, a simple geometric mean rule

$$C_6^{ij} = \left(S_6^{ii} S_6^{jj}\right)^{1/2}. \tag{5.2.16}$$

can be employed. It can be derived from (5.2.15) under a certain condition [276] and the geometric mean rule (5.2.16) was used by Grimme to compute C_6^{ij} coefficients in an early version of the DFT-D method [251].

The atomic polarizabilities in (5.2.15) can be considered as independent parameters or they can be computed with the help of the Slater-Kirkwood formula for the dispersion coefficient as

$$\alpha_i = \sqrt[3]{\frac{16}{9N_i^*}}\left(S_6^{ii}\right)^{2/3}, \tag{5.2.17}$$

where N_i^* is an effective number of electrons contributing to the polarizability of atom i. In the present work this quantity was obtained by the formula proposed by Cambi *et al.* [277]

$$N^* = N_{\text{ext}}\Big(1 + \big(1 - \frac{N_{\text{ext}}}{N_{\text{int}}}\big)\big(\frac{N_{\text{int}}}{N_{\text{ext}} + N_{\text{int}}}\big)^2\Big) \tag{5.2.18}$$

with N_{ext} and N_{int} being the number of "external" and "internal" electrons, respectively. These were taken to be equal to the number of electrons in the valence and core shells of the atomic reference state used in the pseudopotential generation (*vide infra*). Thus, N^* values for the oxygen and zirconium atoms obtained by (5.2.18) are equal to 5.25 and 13.49, respectively.

Note that by substituting (5.2.13) into (5.2.17) one obtains the dependence of the atomic polarizability on the spread S_i^* and, as the value of the exponent q in (5.2.13) is more than 3/2, the polarizability α_i rapidly grows with the increase of the spread S_i^*.

5.3 Calculations.

Systems studied in the work include both crystalline and amorphous structures of silica and zirconia. The crystalline systems are α-quartz (α-QRZ) and α-cristobalite (α-CRB), and the tetragonal modification of ZrO_2 (t-ZrO_2). Amorphous a-SiO_2 and a-ZrO_2 structures were generated by classical molecular dynamics simulations using the simulated annealing technique. The number of atoms, shape, and size of simulation cell for each system are summarized in Table 5.1. Two configurations were employed in the computations of MLWFs in the studied systems. The

Table 5.1: Number of atoms N, shape, and size of simulation cell for studied systems.

System	N	Supercell shape	Supercell parameters (Å)
α-QRZ	72	orthorhombic	8.5103 × 9.8268 × 10.8104
α-CRB	96	tetragonal	9.9418 × 9.9418 × 13.8556
a-SiO_2	192	cubic	14.2640 × 14.2640 × 14.2640
t-ZrO_2	108	tetragonal	10.7382 × 10.7382 × 10.3294
LD a-ZrO_2	192	cubic	13.9690 × 13.9690 × 13.9690
HD a-ZrO_2	192	cubic	13.5020 × 13.5020 × 13.5020

first configuration denoted hereafter as "ini", corresponded to the initial structure. The second one ("cpmd") was the final configuration of a short (0.3 ps) *ab initio* Car-Parrinello molecular dynamics run at an elevated temperature. The temperature was taken to be 500 K for crystalline SiO_2 modifications (below the α to β phase transition) and 1000 K for the other systems. Two structures with a low and high densities, which denoted hereafter as LD and HD, respectively, were considered for the amorphous zirconia. Details on the construction of the structural models can be found in Appendix.

The ground-state DFT plane-wave calculations were carried out with the CPMD code [278]. The calculations used the PBE exchange-correlation functional [13] and the norm-conserving pseudopotentials generated according to the Troullier-Martins scheme [33]. The reference configuration for the O and Si atoms corresponded to a neutral atom, while the Zr^{2+} ion was used as the reference configuration in the generation of Zr pseudopotential with 4s, 4p, and 4d orbitals as semicore states. The pseudopotentials were taken from the CPMD pseudopotential library [278]. The plane-wave kinetic energy cutoff was equal to 60 Ha and the Brillouin zone integration was limited to the Γ-point. The calculation of the MLWFs was done using the built-in capabilities of the CPMD code. The wavefunction was optimized with the convergence criterion 10^{-6} prior to the computation of the MLWFs. The Car-Parrinello molecular dynamics runs used the fictious orbital mass of 700 a.u. and the integration time-step of 4 a.u. A Nosé-Hoover thermostat was applied to both the nuclear and electronic degrees of freedom.

5.4 Results and discussion.

5.4.1 Effective atomic orbitals method and combination rules.

The dispersion C_6^{ij} coefficients for the interatomic interactions in the studied systems were computed by using three methods:

1. the original Silvestrelli approach, *i.e.* by eqns. (5.2.9) and (5.2.12),

2. effective atomic orbital model. The self-atom S_6^{ii} dispersion coefficients were calculated with (5.2.12) and the coefficients were fed into (5.2.14) to obtain the values of spreads S_i^*. The

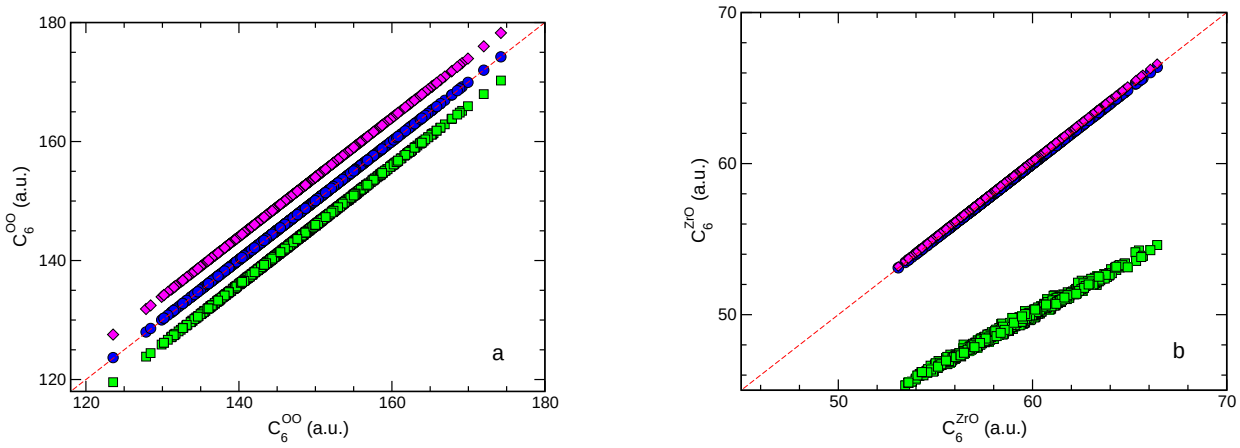

Figure 5.2: The correlation of the C_6^{ij} coefficients computed by the effective orbital method (blue circles) and by the combination rules (5.2.15) and (5.2.16) (green squares and magenta diamonds, respectively) with the reference C_6^{ij} values obtained by the original Sivestrelli approach (x axis) for the oxygen-oxygen (a) and zirconium-oxygen (b) interactions in the LD structure of a-ZrO_2. For the sake of clarity, in figure (a) the C_6^{OO} values computed by the combination rules (5.2.15) and (5.2.16) were equally shifted up and down, respectively, along the y axis. Red dashed line is the identity line.

spreads of effective orbitals of atoms i and j were then used to compute the C_6^{ij} coefficients by (5.2.9).

3. By the combination rules (5.2.15) and (5.2.16) with the atomic polarizabilities computed by (5.2.17).

Figure 5.2 shows the correlations between the reference C_6^{ij} coefficients computed by the original Sivestrelli approach (i) for the oxygen-oxygen and zirconium-oxygen interactions in the LD structure of a-ZrO_2 with the coefficients obtained by the effective orbitals model (ii) and by the combination rules (5.2.15) and (5.2.16) (iii). One sees that the effective orbitals method yields the C_6^{ij} values in almost perfect agreement with those obtained by the original approach. The combination rule (5.2.15) performs very well for the dispersion coefficients of like atoms (Figure 5.2a), while it underestimates the coefficients for unlike atoms (Figure 5.2b). In contrast, the geometric mean combination rule (5.2.16) gives both the like and unlike coefficient values in a very good agreement with those computed by the Silvestrelli method (Figures 5.2a and 5.2b). The reason for the difference is obviously due the use of the atomic polarizabilities computed by (5.2.17) in (5.2.15), but not considering them as independent parameters. In such a case, as it was shown in [268], the rule (5.2.15) predicts values of the C_6^{ij} coefficients for unlike atoms smaller than those derived by the geometric mean combination rule (5.2.16).

The relative root mean squared error between the C_6^{ij} values obtained with the method (i) and method (ii) does not exceed 0.1 % that justifies the use of the effective orbitals model for computing the dispersion coefficients in the studied systems. Such a good performance of the method seems to be surprising if one takes into account that the method replaces the MLWFs

Table 5.2: C_6^{ij} dispersion coefficients (in a.u.) in MO_2 oxide structures (M=Si, Zr), values in parentheses are the standard deviations.

System		C_6^{OO}	C_6^{MM}	C_6^{MO}
α-QRZ	ini	48.06		
	cpmd	60.16 (2.13)		
α-CRB	ini	49.53		
	cpmd	61.52 (2.47)		
a-SiO_2	ini	56.82 (3.10)		
	cpmd	67.50 (3.62)		
t-ZrO_2	ini	110.12	21.49	48.56
	cpmd	112.46 (4.73)	22.18 (0.25)	49.86 (1.54)
LD a-ZrO_2	cpmd	148.14 (8.18)	23.62 (0.27)	59.02 (2.35)
HD a-ZrO_2	cpmd	143.50 (9.93)	23.61 (0.27)	58.07 (2.88)

around the atom by a single atom-centered s-like function. This fact can be rationalized by noticing that the centers of MLWFs identified with a particular atom form a tetrahedron around the atom with the tetrahedron center on the atom. Consequently, these four MLWFs are well approximated by a single atom-centered s-function. The situation can, however, change if the spatial arrangement of the WFCs around the atom will significantly deviate from tetrahedral, as it can happen in systems with directional bonding. In this case, the effective orbitals method may fail because the center of the Wannier functions distribution around the atom will not coincide with the atom position thus leading to a deviation of the C_6^{ij} coefficients obtained with the method from those computed using the original Silvestrelli approach.

The above results allow us to conclude that both the effective orbitals method and the combination rule (5.2.16) can be used for computing the dispersion coefficients with the help of the atomic parameters: spreads S_i^* and self-atom coefficients S_6^{ii}, respectively. The accuracy of the two methods is virtually identical, while the use of geometric mean combination rule (5.2.16) is notably simpler.

5.4.2 Atomic dispersion coefficients.

Mean values of the C_6^{ij} coefficients in the studied systems are given in Table 5.2. One sees that the values of the coefficients are not constant. The maximum variation is observed for the oxygen-oxygen dispersion coefficient that also has largest standard deviations, especially in the amorphous structures. Figure 5.3 shows the distribution of the C_6^{OO} coefficient values for the atoms in the "cpmd" configurations of the studied systems. The C_6^{OO} coefficient value is more than twice larger in zirconium oxide structures as compared to the SiO_2 ones and the change of the coefficient value occurs not only from silicon to zirconium oxide, but also for different phases of the same compound. The data in Table 5.2 show that C_6^{OO} increases by ca. 20 % from crystalline to amorphous structures. It is noteworthy that the dispersion coefficient of zirconium atoms is

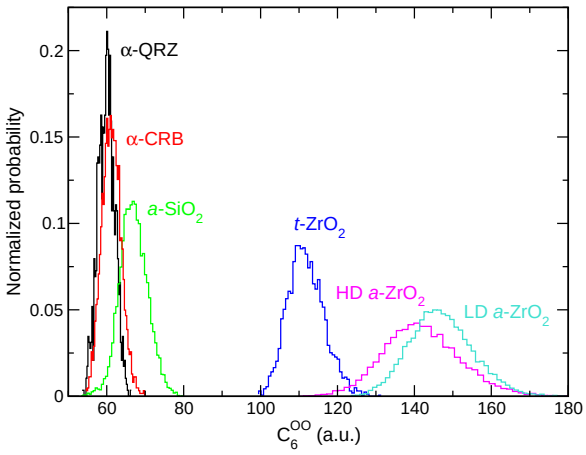

Figure 5.3: Distribution of the C_6^{OO} dispersion coefficient values in silica and zirconia "cpmd" structures.

less system dependent than the oxygen-oxygen one.

Values of the oxygen-oxygen dispersion coefficients reported in Table 5.2 can be compared with values given in [268,279] and which were obtained specifically for the oxygen atoms in oxides. The C_6^{OO} dispersion constants derived from the experimental data [280] by Conesa [268] are equal to 61.3 a.u., 144.1 a.u., and 154.3 a.u. for the α-Al_2O_3, TiO_2 rutile, and TiO_2 anatase structures, respectively. These values are in a good agreement with the data of Table 5.2. Thus, the C_6^{OO} dispersion coefficient in the corundum structure is close to the values computed in silica, while the coefficients for the oxygen atoms in the titanium dioxide polymorphs have the same order of magnitude as the values obtained in the ZrO_2 structures. Wilson and co-workers [279] making use of the results of *ab initio* calculations, obtained the C_6^{OO} coefficients in the MgO and CaO oxides equal to 101.2 a.u. and 171.1 a.u., respectively. These values also agree with the C_6^{OO} coefficients in Table 5.2 taking the M-O bond lengths in these oxides (2.11 and 2.41 Å, respectively) into account (*vide infra*).

Since the value of the interatomic C_6^{OO} coefficient can be computed via the self-atom coefficients S_6^{ii} and S_6^{jj}, the following discussion will mainly be focused on the behaviour of these and related quantities. Table 5.3 reports mean values of the self-atom dispersion coefficients S_6^{ii} and of the atomic polarizabilities α_i obtained by (5.2.17); the last column of Table 5.3 gathers mean values of the M-O bond length in the systems. The behaviour of the S_6^{ii} values accounts for the trends in the C_6^{ij} coefficients (Table 5.2). The S_6^{OO} coefficient reveals the largest variations and the change in the value can be related to the M-O bond length in the structures. Indeed, as the analysis of Table 5.3 shows, longer M-O bond corresponds to larger values of the S_6^{OO} coefficient and of the O^{2-} polarizability. This trend is more apparent for the SiO_2 structures, probably because of more a homogeneous environment of the oxygen atoms in these systems.

Dispersion coefficients and structural characteristics. To get a better insight into the influence of structural characteristics on the S_6^{OO} coefficients in systems, the coefficient values

Table 5.3: Self-atom dispersion coefficients S_6^{ii} (in a.u.), atomic polarizabilities α_i (in Å^3), and mean values of bond lengths $R_{\mathrm{M-O}}$ (in Å) in MO_2 oxide structures (M=Si, Zr). Values in parentheses are the standard deviations.

System		S_6^{OO}	$\alpha_{\mathrm{O}^{2-}}$	S_6^{MM}	$\alpha_{\mathrm{M}^{4+}}$	$R_{\mathrm{M-O}}$
α-QRZ	ini	48.07	1.365			1.607
	cpmd	60.20 (3.04)	1.585 (0.053)			1.710
α-CRB	ini	49.53	1.393			1.600
	cpmd	61.57 (3.52)	1.610 (0.061)			1.693
a-SiO_2	ini	56.91 (4.46)	1.527 (0.079)			1.617
	cpmd	67.60 (5.14)	1.713 (0.087)			1.681
t-ZrO_2	ini	110.12	2.373	21.49	0.583	2.214
	cpmd	112.57 (6.79)	2.407 (0.096)	22.19 (0.36)	0.595 (0.007)	2.145
LD a-ZrO_2	cpmd	148.37 (11.68)	2.892 (0.151)	23.63 (0.38)	0.621 (0.007)	2.124
HD a-ZrO_2	cpmd	143.85 (14.19)	2.833 (0.186)	23.61 (0.38)	0.621 (0.007)	2.129

were related to the mean nearest-neighbour M-O distance R_{MO}. The cut-off distances at which the atoms were considered to be the neighbours were equal to 1.94 Å and 2.39 Å for the SiO_2 and ZrO_2 systems, respectively. These values are by *ca.* 10 % larger than the sum of the single-bond covalent atomic radii [281] that allows taking longer M-O bonds in amorphous structures into consideration. In addition, values of the S_6^{OO} coefficients as a function of the R_{MO} distance were also computed in the ideal crystalline structures of α-quartz, and of the tetragonal and cubic (c) ZrO_2 polymorphs subjected to the isotropic compression/expansion.

Figures 5.4a and 5.4b show the obtained dependences for the silica and zirconia structures, respectively. The general trend is that the S_6^{OO} coefficient increases when R_{MO} rises in all the studied systems, as it could already be inferred from Table 5.3. The behaviour is more apparent for the silica structures, where the coefficient values closely follow the dependence obtained in the compressed/expanded α-quartz structure in a wide range of the nearest-neighbour M-O distances. It is worthy of note that for a given R_{MO} distance, the S_6^{OO} coefficient in the amorphous systems is larger than those in the crystalline ones.

In the ideal t-ZrO_2 and c-ZrO_2 structures the dependences of S_6^{OO} coefficient on the R_{MO} distance are almost identical to each other (Figure 5.4b), despite the fact that the oxygen atoms have different environment in these systems: in c-ZrO_2 the oxygen atoms reside in the center of ideal tetrahedra leading to four equivalent Zr-O bonds ($R_{\mathrm{M-O}}$=2.202 Å), while in the tetragonal zirconia these tetrahedra have the C_{2v} symmetry resulting in two short ($R_{\mathrm{M-O}}$=2.071 Å) and two long ($R_{\mathrm{M-O}}$=2.361 Å) Zr-O bonds. The thermal motion and the structural disorder in the zirconia structures lead to a much larger scatter of the S_6^{OO} values in these systems (Figure 5.4b) as compared to the silica ones. The complication comes from the fact that the oxygen atoms in the ZrO_2 models have different number of nearest-neighbours (coordination number, CN) in contrast to the SiO_2 structures, where all oxygens were found to be two-fold coordinated. The

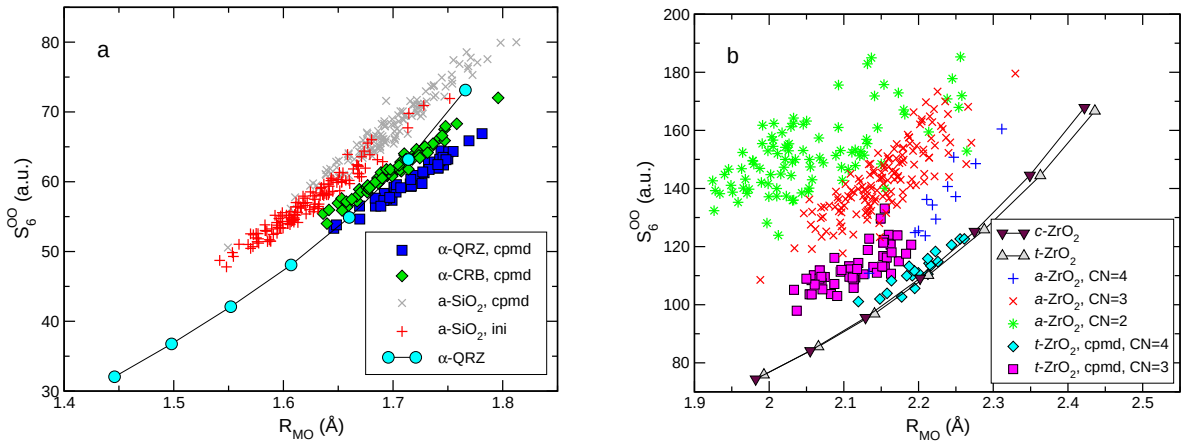

Figure 5.4: Self-atom dispersion coefficients S_6^{OO} in silica (a) and zirconia (b) structures as a function of mean nearest-neighbour distance R_{MO} (M=Si, Zr). In figure (b) CN stands for the coordination number of oxygen atoms.

coordination number of oxygen affects the value of S_6^{OO} coefficient so that the atoms with smaller CN have larger dispersion coefficients. Thus, for a given R_{MO} distance the oxygen atoms with CN=3 in the "cpmd" configuration of t-ZrO_2 are characterized by S_6^{OO} values that are up to 20 % larger than those of the four-fold coordinated atoms. In the amorphous structures the difference is amplified by the structural disorder that leads to further increase of the S_6^{OO} values, as it was obtained in silica (Figure 5.4a).

It is noteworthy that the same analysis applies to the values of oxide ion polarizability whose value increases as the R_{MO} distance rises and the oxygen coordination number decreases. An increase of the polarizability of the oxide and fluorine anions with the cation-anion distance has been previously reported by Heaton *et al.* [282] and by Salanne *et al.* [283], respectively. The authors of the former paper also showed the increase of the O^{2-} polarizability with the decrease of the coordination number of oxygen atom in the crystalline MgO phases (*cf.* Figure 6 of [282]). It should be noted that the dependence of the C_6^{ij} dispersion coefficient on the coordination number of interacting atoms was included by Grimme and co-workers in a recent version of the DFT-D method [252].

The polarizability values given in Table 5.3 can be compared with the literature data. The oxygen anion polarizability in the silica structures well agrees with the theoretical value of $\alpha_{O^{2-}} = 1.48$ Å^3 by Salanne *et al.* [283] for amorphous SiO_2 and with the value $\alpha_{O^{2-}} = 1.401 - 1.454$ Å^3 derived by Dimitrov and Sakka [280] from the experimental data on α-quartz. The computed oxide ion polarizability in t-ZrO_2 is by *ca.* 20 % larger than the corresponding values $\alpha_{O^{2-}} = 1.897 - 2.054$ and $\alpha_{O^{2-}} = 2.204$ Å^3 extracted from the experimental data in [280] and [284], respectively. The polarizability of the zirconium cation can be compared with the values $\alpha_{Zr^{4+}} = 0.357$ Å^3 and $\alpha_{Zr^{4+}} = 0.409$ Å^3 based on the experimental data [280, 284].

The variation of the self-atom dispersion coefficients S_6^{ii} upon the change of the M-O bond length and the oxygen atom coordination number, and the related changes of the O^{2-} polarizability

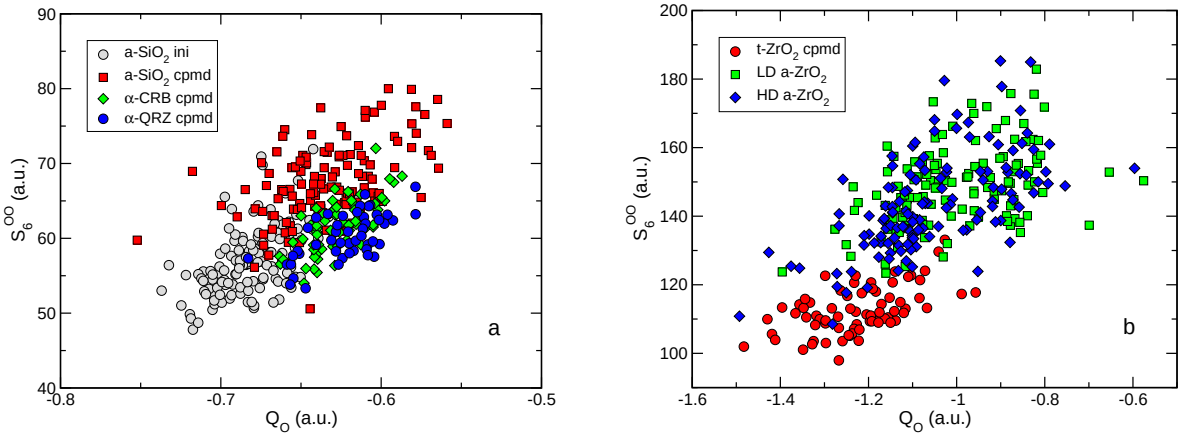

Figure 5.5: Correlation of self-atom dispersion coefficients S_6^{OO} with the oxygen atom charge Q_{O} computed with SQE/ESP method [285].

can be associated with the action of two factors [282, 283]. The first one is a change of Coulombic potential confining the electron density around the atoms and the second effect is a change of volume accessible for the anion electrons, when the cation size changes. These factors are expected to act in opposite directions: the increase of $R_{\mathrm{M-O}}$ distance weakens the confining potential, while the increase of cation size diminishes the volume accessible for the electron density of oxide ion. The behaviour of the S_6^{OO} coefficient and the O^{2-} polarizability indicates that the former effect plays more important role. Furthermore, the decrease of the coordination number also results in the increasing values of the S_6^{ii} dispersion coefficient and of the oxide ion polarizability because of weakening the confining potential.

Dispersion coefficients and atomic charges. The fact that the S_6^{OO} coefficient and the $\alpha_{\mathrm{O}^{2-}}$ polarizability values depend on the confining potential suggests that there should exist a relation between these quantities and other atomic characteristics related to the electrostatic potential, *e.g.* atomic charge. Despite the fact that the charge is not a quantum-chemical observable and there exist many charge schemes [250], the atomic charge concept provides a useful and tangible means of representing the electronic distribution in the system. Furthermore, the atomic charges are important parameters of most force field models of oxide materials. In the present work, we have attempted to relate the self-atom S_6^{ii} coefficient to the atomic charges computed by the split-charge equilibration model (SQE) [167] using a recent set of parameters calibrated on the electrostatic potential (ESP) as the reference quantity [285]. These SQE/ESP charges reproduce the ESP values in both dense and microporous silicate structures and they depend on the structure of system and on the environment of atoms, *i.e.* on the factors playing an important role for the values of the dispersion coefficients.

Figure 5.5 shows values of the self-atom S_6^{OO} coefficient *versus* the SQE/ESP oxygen charges Q_{O} in the silica and zirconia structures. Despite a marked scatter of points, one clearly sees that larger dispersion coefficient corresponds to smaller (in absolute value) charge of the oxygen atom;

note that the same trend can be deduced for the oxide ion polarizability using (5.2.17). At the first glance such a dependence of the S_6^{OO} coefficients and of the polarizability on the charge is surprising since more negative ions are expected to have larger polarizability, whereas Figure 5.5 indicates that the polarizability of ions in solids behaves in an opposite way. Given that the absolute value of SQE/ESP charge is related to the strength of the electrostatic potential at the atom, it becomes clear that larger absolute value of charge reflects stronger confining potential, which, in its turn, results in lower S_6^{OO} coefficient and $\alpha_{\mathrm{O}^{2-}}$ polarizability.

It is noteworthy that the variation of the charges and of the dispersion coefficient is notably larger in the ZrO_2 structures than in the silica ones. As it was mentioned above, this feature is attributed to a more homogeneous environment of the O atoms in the silicon dioxide structures, and in particular to the fact that all oxygen atoms in the studied silica models are two-fold coordinated. The coordination number factor plays an important role for the SQE/ESP charges because the SQE model uses a rigid distance based bonding criterion to allow the charge transfer between atoms. Consequently, a small variation of the interatomic distance might result in a large change of charge if the distance satisfies (or does not) the bonding criterion. In particular, few points with small absolute values of the oxygen charge in Figure 5.5b are due to the one-fold coordinated oxygen atoms in the a-ZrO_2 structures. The use of a more flexible bonding criterion, *e.g.* based in the bond-order SQE model [196, 286] is expected to improve the quality of the SQE/ESP charges and thus, might allow a better description of the relation between the atomic charges and the self-atom dispersion coefficient.

Figures 5.3 and 5.5a show a less sensitivity of the oxygen-oxygen dispersion coefficient and of the oxygen charge to the structural changes in different SiO_2 phases. This feature can be just the reason why a transferable ionic force field model, such as the BKS one [52], could be constructed for silicon dioxide, while no such a successful model has been put forward for zirconia. This remark is in a line with a recent result by Vuilleumier and co-workers [287], who showed that the oxygen electronic polarization in silicate melts only weakly depends on the melt composition.

Geometry-dependent dispersion coefficients. The results presented above point to the eventual necessity of taking the dependence of the dispersion coefficients on the geometry into account while constructing a transferable force field model for oxides. The issue is, however, complicated by the fact that the S_6^{ii} values depend not only on the nearest-neighbour distance, but also on the number of such neighbours. In the case of the silica structures with the two-fold coordinated oxygen atoms the latter factor is excluded and such a geometry dependence can be incorporated into a force field using the data of Figure 5.4a. Combining all the data in the figure into one set, the S_6^{OO} on R_{SiO} dependence can be fitted by the function

$$S_6^{\mathrm{OO}} = a_0 + a_1\, R_{\mathrm{SiO}}^b, \tag{5.4.1}$$

where the parameters are $a_0 = 0.00$ a.u., $a_1 = 2.347$ a.u, and $b = 2.837$. To test the performance of (5.4.1), the S_6^{OO} coefficients were computed by the effective orbital method and by (5.4.1) in

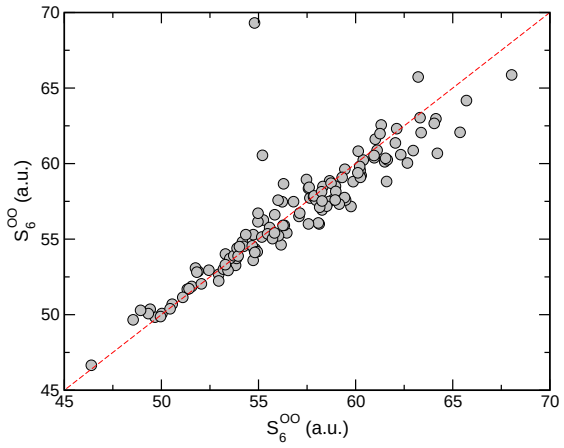

Figure 5.6: Correlation of S_6^{OO} coefficients computed by (5.4.1) (y-axis) with reference S_6^{OO} values (x-axis) in a-SiO_2 structure.

a snapshot configuration of a-SiO_2 obtained in a classical MD run. The correlation between the two set of coefficients is displayed in Figure 5.6. Few points drop out of a very good correlation and the inspection of the structure shows that these points belong to the oxygen atoms having a particular environment. Thus, the point with the coordinates (55, 69) corresponds to an oxygen that could be viewed as three-fold coordinated and is characterized by a large nearest-neighbour distance. Consequently, such a formally two-fold coordinated oxygen has too large S_6^{OO} value predicted by (5.4.1) as compared to the value obtained from the MLWFs analysis. Nevertheless, even with such "bad" points the mean root square relative error (RMSE) of values predicted by (5.4.1) does not exceed 3.5 %.

In the case of the zirconia structures, where the oxygen atoms can have different number of nearest neighbours, a simple relationship between S_6^{OO} and R_{MO} in the form of (5.4.1) does not suffice since for these systems one has also to take the dependence on the coordination number into account. Analysis of Figure 5.4b suggests that the S_6^{OO} on R_{MO} dependences for different coordination numbers follow the same trend, but shifted with respect to each other along the y-axis such that the atoms with lower CN value have larger dispersion coefficient. The following simple modification of (5.4.1) permits to take the behaviour into consideration:

$$S_6^{\mathrm{OO}} = a_0 + a_1\, R_{\mathrm{ZrO}}^b + (4 - W_{\mathrm{O}})\Delta_S, \tag{5.4.2}$$

where W_{O} and Δ_S are the coordination number and the shift parameter, respecively. The former quantity is obtained according to Grimme [252] as a sum of counting function f_{ij} defined as

$$f_{ij} = \frac{1}{1 + \exp(k(r_{ij}/R_{ij}^0 - 1))}, \tag{5.4.3}$$

with k being a scale factor and R_{ij}^0 is a distance parameter proportional to the sum of single-bond covalent radii of atoms i and j. The fractional coordination number W_i of atom i and and the

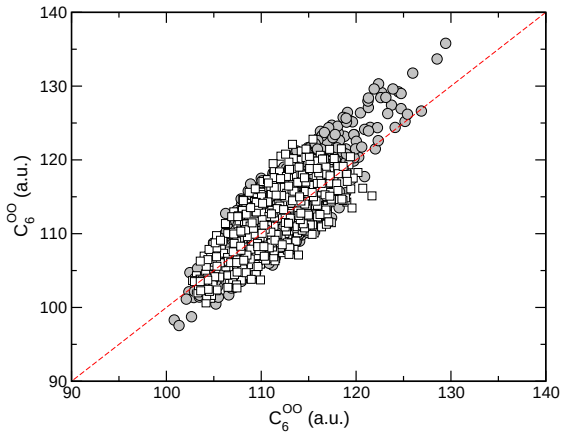

Figure 5.7: Correlation of interatomic C_6^{OO} coefficients computed by (5.4.2) and (5.2.16) (y-axis) with reference C_6^{OO} values (x-axis) in "cpmd" structures of tetragonal (grey circles) and cubic (open squares) ZrO_2.

mean nearest neighbour distance R_{ij} are given by

$$W_i = \sum_{j \neq i} f_{ij} \tag{5.4.4}$$

and

$$R_{\mathrm{ij}} = \frac{1}{W_i} \sum_{j \neq i} f_{ij} r_{ij}. \tag{5.4.5}$$

The parameter R_{ij}^0 in (5.4.3) was taken to be equal to the nearest-neighbour distance criterion (2.39 Å), whereas the scale factor $k = 90$ was chosen such that the fractional coordination numbers of the Zr and O atoms in the cubic ZrO_2 structure were equal to their formal coordination numbers 8 and 4, respectively.

The parameters a_0, a_1, and b in (5.4.2) were obtained by fitting the S_6^{OO} *versus* R_{MO} dependence (Figure 5.4b) for the ideal t- and c-ZrO_2 structures with (5.4.1) and they are equal to $a_0 = 36.5$ a.u., $a_1 = 0.013$ a.u., and $b = 6.045$. The shift parameter $\Delta_S = 33.1$ a.u. was then derived by a least-square fit of S_6^{OO} coefficients with (5.4.2) in all zirconia modifications. The RMS error of fitted S_6^{OO} values is equal to 7.6 % with the maximum relative error (MxRE) of *ca.* 30 %. The errors can be significantly reduced by applying a separate fit to the data for the crystalline and amorphous zirconia phases. Thus, the RMSE and MxRE drop to 3.5 % and 11.8 %, respectively, with $\Delta_S = 19.4$ a.u. in a separate fit for the "cpmd" configuration of t-ZrO_2. In the case of the amorphous structures such a fit yields $\Delta_S = 35.6$ a.u. and the RMSE and MxRE of 5.9 % and 26.5 %, respectively.

To test the applicability of (5.4.2) the interatomic C_6 dispersion coefficients were computed for the oxygen atoms in the final configurations of the tetragonal and cubic ZrO_2 structures after 1 ps CPMD run at 1000 K. Figure 5.7 shows the correlation between the reference values obtained with the MLWFs analysis with those computed using (5.4.2) and the combination rule (5.2.16).

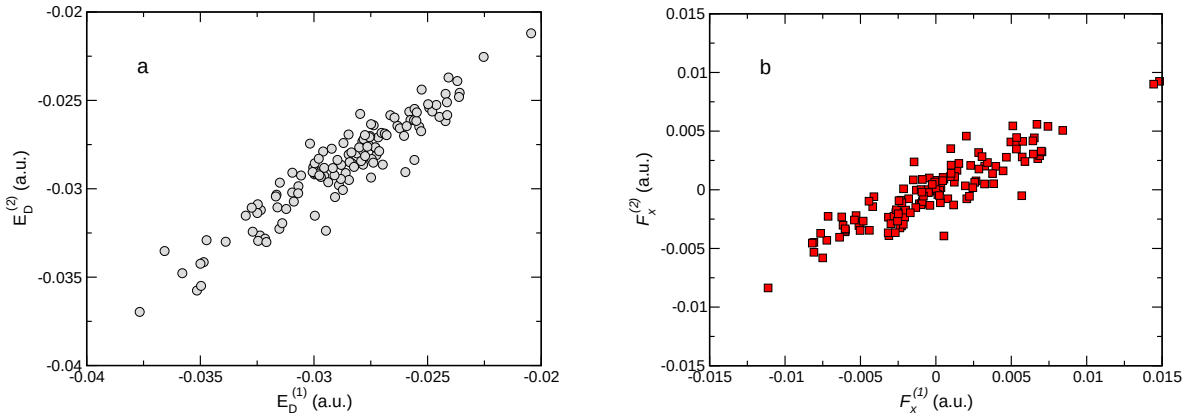

Figure 5.8: Correlation of dispersion energies (a) and x-component of Cartesian force (b) computed with fixed (x-axis) and structure-dependent (y-axis) dispersion coefficients for oxygen atoms in amorphous silica structure. The superscript of the axis labels denotes the method used to compute the corresponding quantity (see text for details).

The agreement is good and the RMSE and MxRE between the reference and predicted values of dispersion coefficients are equal to 3 % and 10.3 %, respectively. The results shown in Figures 5.6 and 5.7 indicate that a force field model with geometry-dependent dispersion coefficients can be constructed for a specific oxide using the above discussed schemes.

5.4.3 Implication for energy and forces.

In order to find out how the structure-dependent dispersion coefficients affect the energies and forces of atoms in the system, the dispersion energies and the Cartesian forces were calculated for the oxygen atoms in a structure of a-SiO_2 taken from a snapshot of classical MD run at 1000 K. In the first calculation the fixed value of the dispersion coefficient $C_6^{OO} = 56.91$ a.u. was attributed to the oxygen atoms regardless their environment. The second calculation used the interatomic C_6^{OO} coefficients computed by the geometric mean rule (5.2.16) with the use of the S_6^{OO} coefficients obtained by (5.4.1), i.e. by taking the environment of the atoms into account. The correlation between the energies and the force components derived with the two calculations was estimated by using the following criterion ΔX

$$\Delta X = \frac{\sum_i |X(i) - X_0(i)|}{\sum_i |X_0(i)|}, \tag{5.4.6}$$

where $X(i)$ denotes either the dispersion energy E_D of atom i or the x-component of the force on the atom, and subscript 0 signifies the reference value.

Figure 5.8 shows the interaction energy and the x-component of the Cartesian force on the oxygen atoms in the silica structure. For the both quantities, one observes a good correlation between the values computed with the two methods with the correlation coefficient $R^{>}0.9$. However, while the error (5.4.6) in the energies ΔE^D does not exceed 3 %, that is very acceptable value, the error in the forces ΔF_x amounts to 49 %. Figure 5.8b indeed shows that the use of

the constant dispersion coefficient generally leads to larger dispersion forces that in the case of geometry-dependent S_6^{OO} coefficients.

5.5 Conclusions.

Interatomic C_6^{ij} dispersion coefficients in crystalline and amorphous structures of silicon and zirconium dioxides were obtained with the approach proposed by Silvestrelli [204, 255] for computing the dispersion interaction energy from results of DFT calculations and based on the use of maximally-localized Wannier functions (MLWFs) for partitioning electronic density. The localization of MLWFs close to the nuclei in these systems makes it possible to assign the functions to the atoms in an unambiguous way, thus allowing the calculation of dispersion coefficients in an atom-wise manner. It is found that the MLWFs can be condensed to effective orbitals centred on atoms and the spread of these effective orbitals can then be used for computing the interatomic dispersion coefficients using the formalism by Andersson *et al.* [261]. The effective orbitals method yields the results in a very good agreement with those obtained by the original approach.

The dispersion coefficient C_6^{OO} for oxygen atoms was found to have the largest variation. Its value changes not only from one oxide to another, but also for the different modifications of the same compound. On the other hand, the dispersion coefficient for cations is almost insensitive to structural changes. Results of the study allowed to test two widely used combination rules for calculation of unlike dispersion coefficients. The coefficients for unlike interactions were found to be best reproduced with the geometric mean combination rule, in agreement with finding by Grimme [252].

The dipole polarizability of ions in the systems was computed making use of the Slater-Kirkwood formula and the values of self-atom dispersion coefficient appearing in the effective orbitals model. The oxide ion polarizability in silica structures is in a good agreement with the values obtained in previous calculations and from experimental data [280], whereas in the ZrO_2 phases the polarizability is slightly larger than the $\alpha_{\mathrm{O}^{2-}}$ values derived from the experimental data. The values of the self-atom dispersion coefficient S_6^{OO} and of the oxide ion polarizability were found to correlate with the nearest-neighbours distance and with the coordination number of oxygen atoms in the structures. This feature is attributed to the effect of confinement by electrostatic potential. The values of the coefficient and $\alpha_{\mathrm{O}^{2-}}$ quantity were related to the oxygen charges computed with SQE/ESP method [285]. In all studied systems the atoms having larger absolute values of charge were found to have a smaller S_6^{OO} value and to be less polarizable because of stronger confinement of electrons in a more deep potential wall. Simple functional forms for the dependence of the S_6^{OO} coefficients on the nearest-neighbour distance and coordination number are suggested. A test calculation of the dispersion energies and forces for the oxygen atoms in an amorphous silica have shown that the use of structure-dependent dispersion coefficients results in smaller values of the forces, than in the case of constant C_6 coefficient. On the other hand, the values of atomic dispersion energies derived by both ways agree with each other within few percent.

Chapter 6

Short-range potentials by force-matching procedure.

6.1 Introduction.

The general strategy of the polarizable force field model development was discussed in Section 3.3 of Chapter 3 and is based on the assumption that the total potential energy of the system can be described as a sum of long-range and short-range energy contributions with the same additivity supposed to be valid for any system characteristics derived from the energy. Consequently, the long-range and short-range energy terms can be analyzed separately and then combined into one model. This model can then be extended by the "nonclassical" contribution due to the dispersion interactions, as discussed in Chapter 5, thus leading to the complete force field.

The development of the long-range part of the force field was presented in Chapter 4 in details. The long-range interactions in the model are envisaged to be treated with the SQE method using either HI- or ESP-based parameterization. Despite the fact that the long-range energy has a pairwise form given by eqn. (3.2.42), the resulting atomic charges are, in fact, many-body quantities because of the many-body character of the inverse of the hardness matrix $\mathbf{H}^{-1}$. Since the magnitude of the electrostatic energy is large, one can expect that these *effectively* many-body charges capture the most important part of the many-body interactions in the system. It is worthy of noticing that the dispersion energy has also *effectively* many-body character since the dispersion coefficients are computed with the help of the maximally localized Wannier functions based on the self-consistent electron density.

The present chapter deals with the development of the remaining part of the polarizable force field, namely the model of short-range interactions. The short-range interaction potentials are to be obtained by using the force-matching method suggested by Ercolessi and Adams in 1994 [203]. The key idea of the method is to replace a limited set of data available for the potential fitting from the experiments by (potentially) unlimited number of quantities resulting from the quantum-chemical calculations, notably the atomic forces. Indeed, $3N$ Cartesian forces in a system of N atoms provide a wealth of information that can be used to obtain reliable and

transferable effective potentials. Besides the forces, the quantities which are typically used in the force-matching procedure are the differences in the energies of configurations and the components of the stress tensor (for the periodic systems). These "observables" together with the Cartesian forces play the most important role for the structural and dynamical properties of systems, and for the relative stabilities of different conformations (polymorphs).

As it was discussed in Section 3.3, silicon dioxide (SiO_2) is chosen as a model system in the present study. This compound has a rich phase diagram with numerous crystalline polymorphs with a wide range of densities. Thus, the α-quartz structure has the density $\rho = 2.66$ g cm^{-3}, while the ITQ-40 zeolite is characterized by the density as low as $\rho = 1.01$ g cm^{-3}, i.e. just above the density of water. Clearly, design of a transferable polarizable force field for silica is a challenging task. The dense crystalline silica polymorphs, such as quartz, cristobalite, tridimite, coesite undergo a sequence of phase transitions upon the change of temperature and pressure. Furthermore, the Si-O bonds are neither purely ionic nor purely covalent that provides additional difficulties for the choice of the model interaction potentials. Hence, a really "good" and transferable potential model should allow describing all these features, at least in a semi-quantitative manner.

As the exact form of interatomic potential function is unknown, one has to rely on some approximations. In the present work, we assume that the short-range interactions in silica can be described by effective *pairwise* potentials. The choice of such a simple model relies on the fact that there exist successful force fields for silica based on the pairwise approximation. Thus, one of the most widely used silica force fields suggested by van Beest, Kramer and van Santen [52] describes the short-range potential by the rigid ion potential (3.1.6). The same form, but with different values of the parameters, is used in the model developed by Carré *et al.* [288]. The variety of computer simulation studies have shown that such simple force fields are capable of mimicking the structural and some of dynamical characteristics. On the other hand, such models often fail to correctly represent the vibrational dynamics [289, 290]. Nevertheless, as a first approximation, the present work makes use of pairwise effective potentials to describe the short-range interactions in silicon dioxide.

Two ways of representing the energy with analytical functions were employed in the study. The first employs of cubic splines to describe the potential curve of pair interactions extracted from the *ab initio* data with the help of the force-matching procedure. No physical significance is assigned to the model parameters (spline coefficients) and the only requirement is that the spline representation of the effective potentials should reproduce the reference data as closely as possible. The second way uses a set of suitably chosen analytical functions. In this case the model potentials as well as their parameters are expected (and it is highly desirable) to have the physical meaning. The analytical form of potential functions is chosen from the theoretical considerations regarding the distance dependence of different contributions to the total energy of the system, e.g. short-range repulsion, induction energies, etc.

Results of the force-matching procedure using the two ways are presented in subsequent sections of the chapter.

6.2 Minimization criteria.

The force-matching procedure is a minimization procedure and as such, it has to use a function whose minimum in the multidimensional space of parameters needs to be found. Following the previous studies on the use of the force-matching method, the present work employs the cost function based on the atomic forces and on the elements of the stress tensor in the system.

Cost functions and merit criteria.

The cost function related to the atomic forces is defined as

$$\chi_F^2 = \frac{1}{3n_c N_a} \sum_{i_c=1}^{n_c} \sum_{i_a=1}^{N_a} \sum_{\alpha=x,y,z} (F^{\alpha}_{M,i_a,i_c} - F^{\alpha}_{AI,i_a,i_c})^2 \tag{6.2.1}$$

and that of the stress tensor reads

$$\chi_S^2 = \frac{1}{3n_c} \sum_{i_c=1}^{n_c} \sum_{\alpha\geq\beta} (S^{\alpha\beta}_{M,i_c} - S^{\alpha\beta}_{AI,i_c})^2, \tag{6.2.2}$$

where

n_c is the number of atomic configurations used in the force-matching procedure;

N_a is the number of atoms in the system;

F^{α}_{M,i_a,i_c} is the α-component of force acting on atom i_a in the atomic configuration i_c computed with the model;

F^{α}_{AI,i_a,i_c} is the α component of the reference *ab initio* force acting on atom i_a in the atomic configuration i_c;

$S^{\alpha\beta}_{M,i_c}$ is the $\alpha\beta$ component of the stress tensor for the atomic configuration i_c computed with the model;

$S^{\alpha\beta}_{AI,i_c}$ is the $\alpha\beta$ component of the *ab initio* stress tensor for the configuration i_c.

It can be seen that both χ_F^2 and χ_S^2 quantities are the absolute mean errors. The use of the absolute errors instead of the relative ones is dictated by the need of excluding possible undesirable effects that can occur because of a large relative error of reference quantities having small absolute values. Indeed, despite the fact that these small values can be of little importance (or even negligible), their corresponding relative errors can be so large that they will heavily contaminate the cost function.

Similarly to the way employed in the parameterization of the long-range potential model (cf. Appendix C), the function to minimize (total cost function) was constructed as

$$\chi^2_{\mathrm{TOT}} = \frac{\chi_F^2 + \lambda\chi_S^2}{1+\lambda}, \tag{6.2.3}$$

where the parameter λ is a weight parameter whose value was varied in the interval $[10^{-4}, 10^{0}]$. One sees that when λ approaches its lower bound, the stress cost function has a minimal effect on the total cost function. In the opposite case, the force and stress cost functions have equal contributions to χ^2_{TOT}.

It should be noted that in our experience $\chi^2_F \gg \chi^2_S$. Consequently, the χ^2_S function was scaled by a factor of $\langle {F^{\alpha}_{AI,i_a,i_c}}^2 \rangle / \langle {S^{\alpha\beta}_{AI,i_c}}^2 \rangle$, where $\langle \ldots \rangle$ states for the value averaged over all configurations used in the force-matching.

Merit criteria.

Besides the values of the cost functions, the quality of obtained effective potentials was tested against the following merit criteria [248]

$$\Delta F = \frac{\sqrt{\sum_{i_c=1}^{n_c} \sum_{i_a=1}^{N_a} \sum_{\alpha=x,y,z} (F^{\alpha}_{M,i_a} - F^{\alpha}_{AI,i_a})^2}}{\sqrt{\sum_{i_c=1}^{n_c} \sum_{i_a=1}^{N_a} \sum_{\alpha=x,y,z} (F^{\alpha}_{AI,i_a})^2}} \tag{6.2.4}$$

$$\Delta S = \frac{\sqrt{\sum_{i_c=1}^{n_c} \sum_{\alpha,\beta} (S^{\alpha\beta}_{M} - S^{\alpha\beta}_{AI})^2}}{\sqrt{\sum_{i_c=1}^{n_c} \sum_{\alpha,\beta} (S^{\alpha\beta}_{AI})^2}} \tag{6.2.5}$$

$$\Delta E = \frac{\sqrt{\sum_{i_c,j_c}^{n_c} \left((E^{M}_{i_c} - E^{M}_{j_c}) - (E^{AI}_{i_c} - E^{AI}_{j_c}) \right)^2}}{\sqrt{\sum_{i_c,j_c}^{n_c} (E^{AI}_{i_c} - E^{AI}_{j_c})^2}}, \tag{6.2.6}$$

where

$E^{M}_{i_c}$ and $E^{M}_{j_c}$ are the energies of the configurations i_c and j_c, respectively, computed with the model;

$E^{AI}_{i_c}$ and $E^{AI}_{j_c}$ are the *ab initio* energies of the configurations i_c and j_c, respectively.

These criteria are related to the mean relative error in the computation of the forces, the stress, and the energy, respectively, but they are not contaminated by possibly large relative errors discussed above. Note that the difference in the energy of configurations is not used in the minimization procedure and therefore, ΔE quantity serves an additional validation criterion for the quality of the fit.

6.3 Creating the database of reference values.

Design of a transferable potential necessitates sampling of a large part of the conformational space to obtain the reference quantities corresponding to as large as possible interval of interatomic distances. This aim cannot be achieved by studying the crystalline structures in which the distances are found within narrow intervals corresponding to the first, second, and further coordination

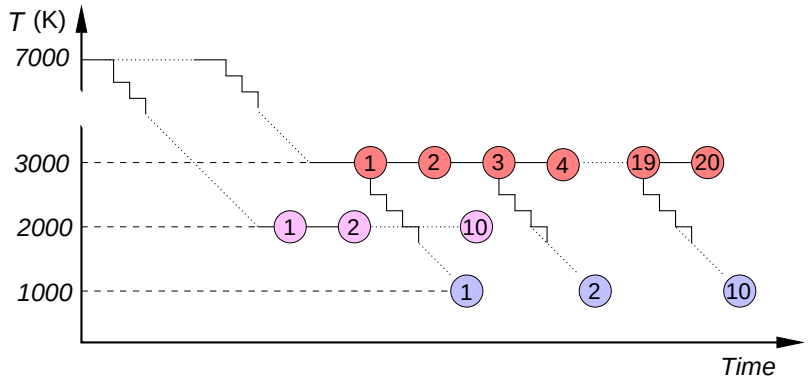

Figure 6.1: Scheme showing the preparation of amorphous silica configurations starting from the crystalline structure of β-cristobalite.

spheres. Consequently, the structures used to create the databases for the force-matching procedure are commonly taken for amorphous modifications of the compound, where a large scatter of the interatomic distances occurs.

In the present work the reference structures of amorphous silica were selected according to the procedure described below. A system consisting of $2 \times 2 \times 2$ unit cells of β-cristobalite (64 silicon and 128 oxygen atoms) was constructed using the data of Ref. [291]. In order to obtain an amorphous silica structure, classical molecular dynamics simulations were carried out with the use of BKS potential [52]. The classical equations of motion of atoms were integrated with the velocity Verlet algorithm using the time-step $\Delta t = 1$ fs. The electrostatic interactions were treated with the Ewald sum method. Since the MD simulations of silica in NPT ensemble with the BKS potential tend to produce too dense structures, all MD calculations were carried in NVT ensemble. The size of cubic simulation box $a = 14.264$ Å corresponds to the experimental density of $\rho = 2.2$ g cm^{-3} of amorphous silica.

A scheme presenting the system preparation procedure is displayed in Figure 6.1. The initial crystalline structure of β-cristobalite was melted at 7000 K for a period of 2 ns and then, the system was slowly cooled down from 7000 K to 3000 K by steps of 100 K, each step lasted 200 ps. The resulting structure was used in a 2 ns MD run at 3000 K and the atomic configurations were saved each 50 ps of the last one nanosecond period (20 configurations). Odd configurations were further cooled down to $T = 1000$ K during 1 ns MD runs. In addition, 10 amorphous silica configurations at $T = 2000$ K were prepared by using a similar procedure starting from a silica melt at $T = 7000$ K. The resulting 30 configurations of the amorphous silica structures corresponding to the temperatures 1000, 2000, and 3000 K were used in single-point energy calculations with the CPMD program.

The CPMD calculations were performed with the revPBE XC functional [14]. This XC functional was chosen because it does not produce an artificial overbinding of van der Waals complexes as some other GGA XC functionals do, e.g. the PBE one [292–294]. The Si and O atoms were described with norm-conserving pseudopotentials generated according to Troullier-Martins

scheme [33]. The reference configuration for the both atoms in the pseudopotential generation corresponded to a neutral atom. The kinetic energy cutoff for the plane wave basis set was equal to 120 Ry. Such a large energy cutoff was necessary to ensure the convergence of the stress tensor. Given the large size of the supercell used in the calculations, the sampling of Brillouin zone was limited to the Γ-point. The CPMD calculations of 30 silica structures have provided the database of the energies, the atomic forces, and the stress tensors that were used in the force matching procedure as described below.

6.4 Effective interatomic potentials by cubic splines.

The first method, which was employed to obtain the short-range pair potentials in silica, is the force-matching method with the cubic splines. This approach allows one to avoid the use of a specific analytical form for the interaction potentials. Basically, any term in (3.1.1) can be approximated by spline functions. However, to perform the spline approximation, one should construct a grid of points at which the splines are defined. For the two-body E_2 term in (3.1.1) the grid is constructed only in one dimension, whereas the grid has to have the dimensionality of 3 and 6 (reduces to 5, if all 4 species lie in the same plane) for three-body or four-body potentials (the E_3 and E_4 terms), respectively, and the computation burden grows accordingly. In conformity with to the above discussion, the spline approximation in the present work was performed for pair potentials.

The method of force-matching by splines has first been proposed by Izvekov and co-workers [295] and has already been used for derivation of pair potentials in silica by Carré [296]. A significant advantage of the approach is that the Cartesian atomic forces can be represented as a linear form with respect to spline parameters, that implies the use of linear minimization techniques for seeking the solution.

Cubic splines approximation of a function.

A spline function $f(x)$ is defined as a sum of piecewise functions $S_i(x)$,

$$f(x) = \sum_{i}^{n} S_i(x), \tag{6.4.1}$$

where each function $S_i(x)$ is nonzero only in the interval $[x_i,\, x_{i+1}]$. This piecewise function $S_i(x)$ is often chosen to be in the form of a cubic function (spline) of x and is defined as

$$S_i(x) = \begin{cases} a_i(x-x_i)^3 + b_i(x-x_i)^2 + c_i(x-x_i) + d_i, & x \in [x_i, x_{i+1}] \\ 0, & x \in (-\infty, x_i) \cup (x_{i+1}, \infty). \end{cases} \tag{6.4.2}$$

As it follows from its definition, to obtain the spline approximation to a function $y(x)$ one has to define a grid of points $x_0,\, x_1, \ldots, x_n$ at which the values of the function $y_i \equiv y(x_i)$ are known and

to construct the cubic splines (6.4.2) such that the resulting spline function $f(x)$ is continuous and has continuous first and second derivatives on the whole interval $[x_0, x_n]$.

It can be shown that the cubic function with these properties can be represented in the form

$$S_i(x) = A_i(x)f_i + B_i(x)f_{i+1} + C_i(x)f_i'' + D_i(x)f_{i+1}'', \tag{6.4.3}$$

where

$$A_i(x) = \frac{x_{i+1} - x}{x_{i+1} - x_i}, \quad B_i(x) = 1 - A_i(x) = \frac{x - x_{i+1}}{x_{i+1} - x_i} \tag{6.4.4}$$

and

$$C_i(x) = \frac{1}{6}(A_i^3(x) - A)(x_{i+1} - x_i)^2, \quad D_i(x) = \frac{1}{6}(B_i^3(x) - B)(x_{i+1} - x_i)^2. \tag{6.4.5}$$

Now, one should apply the smoothness condition

$$\frac{dS_{i-1}}{dx}(x_i) = \frac{dS_i}{dx}(x_i) \tag{6.4.6}$$

that yields

$$\begin{aligned} 0 \;=\; & \frac{f_{i+1}}{x_{i+1} - x_i} - f_i\Big(\frac{1}{x_{i+1} - x_i} + \frac{1}{x_i - x_{i-1}}\Big) + \frac{f_{i-1}}{x_i - x_{i-1}} - \frac{1}{6}f_{i+1}''(x_{i+1} - x_i) - \\ & \frac{1}{3}f_i''(x_{i+1} - x_{i-1}) - \frac{1}{6}f_{i-1}''(x_i - x_{i-1}), \end{aligned} \tag{6.4.7}$$

where $f_i = f(x_i)$ and $f_i'' = f''(x_i)$ are introduced for convenience. In addition to eq. (6.4.7) one needs two more equations to define all second derivatives uniquely. Usually, it is done by applying the so-called natural conditions:

$$f_0'' = 0, \quad f_n'' = 0. \tag{6.4.8}$$

For the purpose of the present work one more condition has to be applied to the spline function $f(x)$

$$f_n = 0. \tag{6.4.9}$$

The condition (6.4.9) stems from the requirement of having zero force at the cut-off radius of short-range interactions, i.e. the use of shifted-force potential (*vide infra*).

Force-matching by cubic splines.

Assuming the pairwise interaction model, the α-component of force acting on atom i, F_i^α, can be written in the form

$$F_i^\alpha = \sum_{j \neq i} f^{ki\,kj}(r_{ij})\frac{\alpha_{ij}}{r_{ij}}, \tag{6.4.10}$$

where $f^{ki\,kj}(r_{ij})$ is a spline approximation of the pairwise force between atoms of kind ki and kj ($i \in ki$ and $j \in kj$). Provided that the function f is given by eqn. (6.4.1), it is clear that the force component F_i^α is a linear combination of the f_i and f_i'' values (cf. (6.4.3)).

The problem can then be recast in the matrix form

$$\mathbf{A} \cdot \mathbf{X}_s = \mathbf{F}, \tag{6.4.11}$$

where $\mathbf{A}$ is a matrix of coefficients (elements of the matrix depend only of the grid used in the spline construction and on the boundary conditions) that has the following form

$$\mathbf{A} = \left(\begin{array}{c} \mathbf{B} \\ \hline \mathbf{C} \end{array} \right). \tag{6.4.12}$$

The matrix $\mathbf{B}$ has the dimension $3N \times mn$ with N, n and m being the number of atoms, the number of points in the grid, and the number of types of pair interactions, respectively. Matrix $\mathbf{C}$ is the matrix corresponding to the boundary conditions (6.4.7)–(6.4.9). The vector $\mathbf{X_s}$ in (6.4.11) is the vector of unknowns and the vector $\mathbf{F} = (F_1^x, F_1^y, F_1^z, \ldots, F_N^x, F_N^y, F_N^z; 0, \ldots, 0)$ is constructed from the Cartesian components of *ab initio* atomic forces and is extended by the $3(n+1)$ zeros to account for the boundary conditions.

The number of unknowns in eqn. (6.4.11), being equal to nm, is generally less than the number of atomic forces and the equation (6.4.11) has the solution in the least-square sense. The present work used the singular value decomposition (SVD) method to obtain the vector $\mathbf{X_s}$ of the spline parameters.

To perform the spline approximation of force one should supply the following input data:

- The Cartesian coordinates of atoms and the reference values of the Cartesian forces on the atoms.
- A grid of points on which the splines will be constructed.

Special care should be taken while constructing the grid of points for the spline approximation. The simplest case is a regular grid, which however does not suit the force-matching with splines. The reason for this is that some of interatomic distances occur more often than others and consequently, the statistical weight of these more frequently sampled interactions will be more important that can bias the final results.

To circumvent the problem, the grid in the present work was constructed with the use of radial distribution function $g(r)$ for the specific pair interaction (i.e. O-O, Si-O, and Si-Si). For each of three interaction potentials the zero point x_0 was set equal to the minimum x value, where the $g(x)$ is non-zero. Then, it was assumed that each grid interval should include the same integral value of the radial distribution function. Mathematically, this means that

$$\int_{x_{i-1}}^{x_i} dx\, g(x) = \int_{x_i}^{x_{i+1}} dx\, g(x), \quad i = 1, .., n-1. \tag{6.4.13}$$

Such a choice of the grid leads to a situation, where the intervals including the maxima of RDF

are narrower than others, where the $g(r)$ has lower values. In particular, the presence of several intervals covering the first maximum in $g(r)$ can improve the description of a curvy potential function at the bonding distance.

Effective potentials by splines. The first question that arises while carrying out the force-matching by splines is to determine how many points are necessary in the spline grid for a proper description of a pair potential. For this purpose the O-O, Si-O, Si-Si radial distribution functions were constructed from the 30 configurations of amorphous silica used in the CPMD calculations and the choice of grid intervals was then performed with the use of the $g(r)$ functions as discussed above. The optimal number of points in the grid was then determined by computing the cost function (6.2.1) and the ΔF merit criterion (6.2.4). Figure 6.2 displays the obtained dependencies from the spline approximations of the total *ab initio* forces.

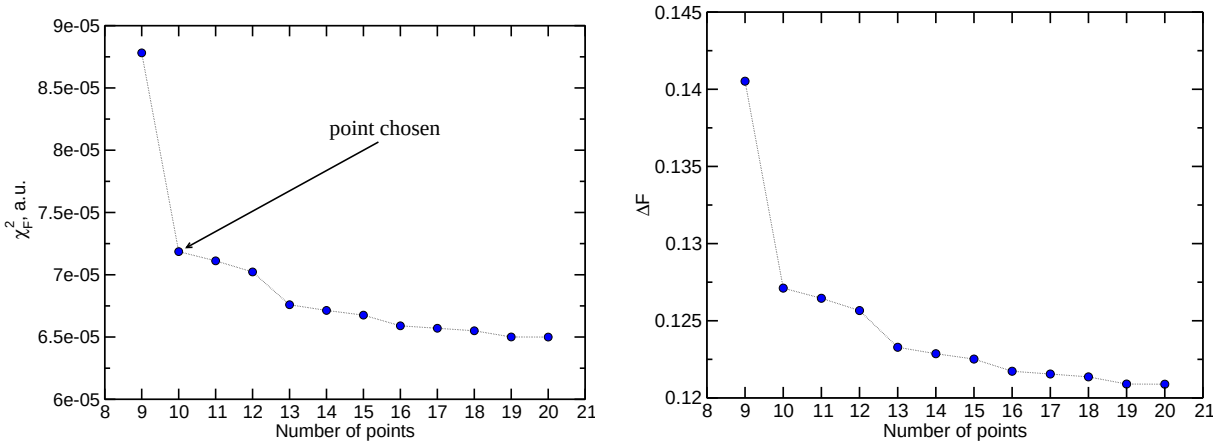

Figure 6.2: Dependence of the χ^2_F cost function (left) and of the ΔF merit criterion (right) on the number of points in the grid. The spline approximation was performed for the total *ab initio* force.

As it could be expected, the error in the approximation of the *ab initio* forces by the spline potential decreases as the number of points increases. Such a behaviour is explained by the increase of the number of spline parameters n and thus, by a more flexible representation of the forces by the model. Usually, the less the error the better potential is assumed to be. However, in the case of large number of grid points the potential behaves irregularly. This irregularity is seen as a nonmonotonical behaviour of the force (energy) (a wave-like form) in some intervals. The O-O and Si-Si pair potentials for each number of grid points were analyzed for the existence of the feature. The first derivatives of the force were calculated and their maximum positive values were compared for different grids (a monotonically decaying function should not have positive first derivative). The analysis has permitted to select a 10 points grid as the compromise between the precision of the approximation and the smoothness of the spline potential.

The total *ab initio* forces approximated with the splines and the effective potentials obtained by integration of the forces are shown in Figure 6.3. It is worthy of noticing that the Si-Si force

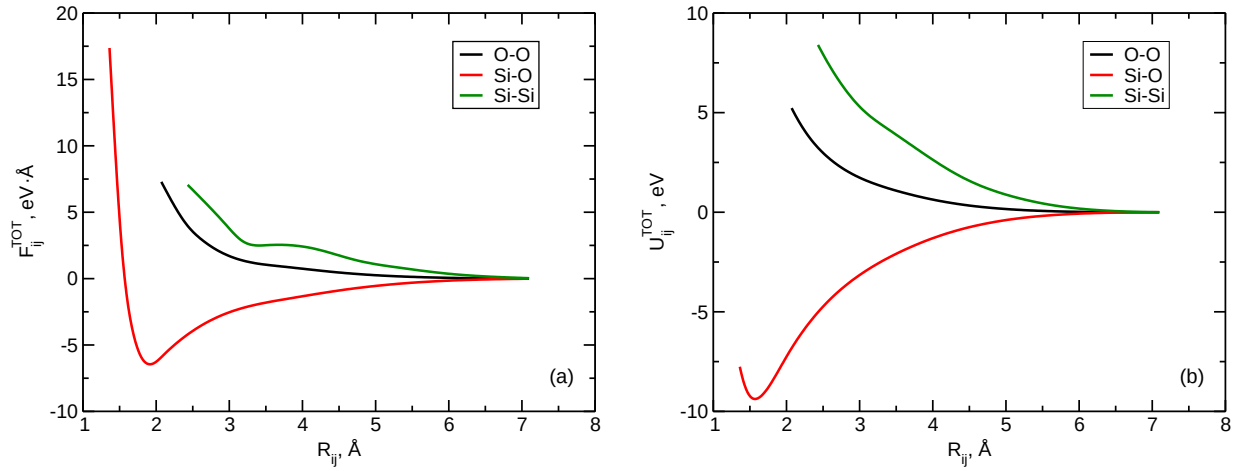

Figure 6.3: Pair forces (a) and effective pair potentials (b) obtained from the force-matching of the total *ab initio* forces with splines.

and potential curves still have the signature of the irregularities described above in the interval 3 to 4 Å. All attempts to remove this feature by varying the number of grid points or the grid selection scheme have failed that allows us to suppose that the feature is inherent to the Si-Si potential. It is interesting that a similar behaviour of the Si-Si spline potential has been found by Carré [296].

The merit criterion ΔF obtained in the fitting the total *ab initio* forces by splines is equal to $\Delta F = 0.1271$ and the potentials shown in Figure 6.3 can be considered as an "ideal" representation of the pair interaction potentials in silica. A force-matching fit that uses any analytical function for the representation of the pair potentials in the system is expected to yield higher values of the cost function χ_F^2 and of the merit criterion ΔF.

The main aim of the spline force-matching approach is to obtain the two-body short-range interaction potentials and to analyze their behaviour. As it was discussed previously, the short-range forces $\mathbf{F}_{\mathrm{SR}}$ were obtained as

$$\mathbf{F}_{\mathrm{SR}} = \mathbf{F}_{AI} - \mathbf{F}_{\mathrm{LR}}, \tag{6.4.14}$$

where $\mathbf{F}_{AI}$ correspond to total *ab initio* forces and $\mathbf{F}_{\mathrm{LR}}$ are the long-range electrostatic forces, that are to be calculated using the SQE model, described in Chapter 4. The LR forces were computed with the two SQE parametrizations, namely SQE/ESP and SQE/HI. The later parameterization modified for the solid state calculations was employed (Chapter 4). The number of grid points was chosen using the procedure outlined above and it was equal to 10 and 11 points for the SQE/ESP and SQE/HI parametrizations, respectively. Hereafter, the short-range potentials corresponding to the SQE/ESP and SQE/HI parameterizations of the long-range interactions are denoted as ESP and HI, respectively.

The resulting pair short-range forces and energies are shown in Figure 6.4 and Figure 6.5. Analysis of the potentials leads to an interesting observation. Let us suppose that the short-range

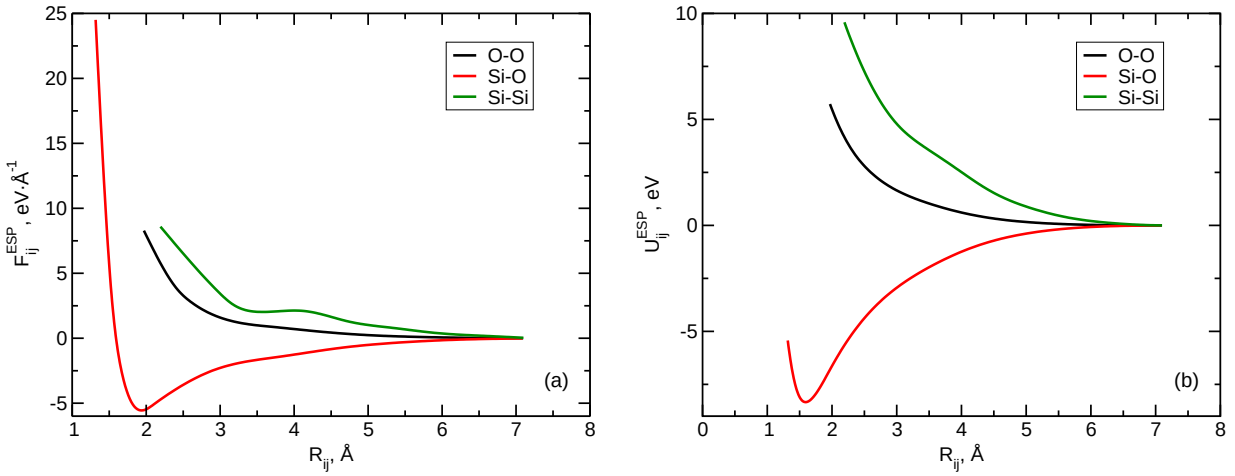

Figure 6.4: Pair forces (a) and effective pair potentials (b) obtained from the force-matching of forces with splines in the ESP ansatz.

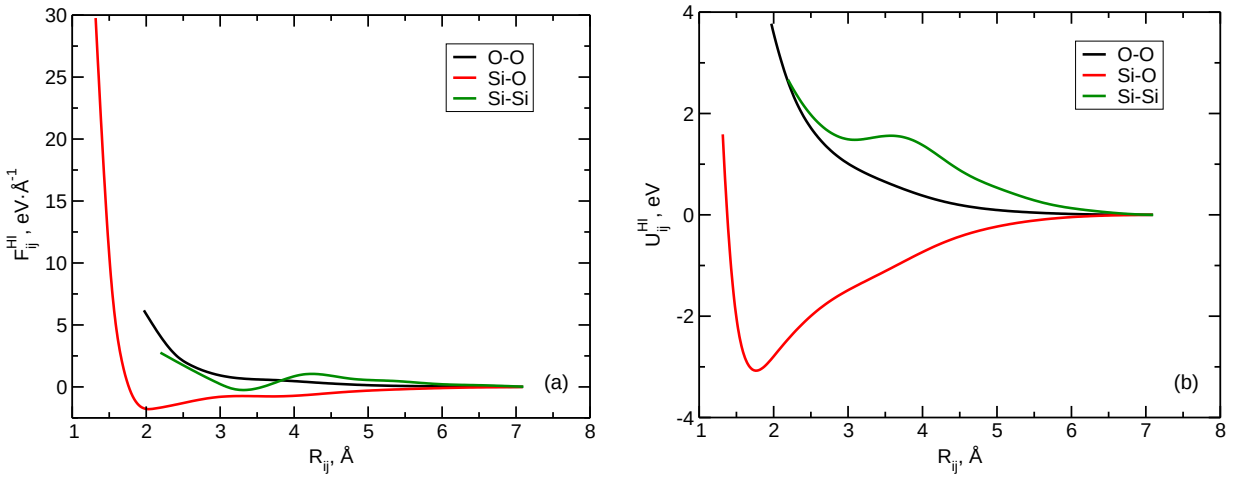

Figure 6.5: Pair forces (a) and effective pair potentials (b) obtained from the force-matching of forces with splines in the HI ansatz.

potentials in silica can be described with the RIM model (3.1.6), like in the BKS force field, but with the Coulomb term replaced by the SQE one. Given that the DFT-computed forces lack the dispersion energy, the short-range potentials would have, in the RIM case, the form of exponential repulsion. Figure 6.4 and Figure 6.5, however, show that, at least in the case of Si-O interactions, the effective potentials should include an attractive energy term.

The merit criterion ΔF for ESP- and HI-parametrized short-range effective potential is equal to $\Delta F_{\mathrm{ESP}} = 0.1498$ and $\Delta F_{\mathrm{HI}} = 0.1851$, respectively. Since in the case of the spline approximation, the merit criterion can be considered as a measure of the quality of the approximation with the pair potentials, one can conclude that the ESP short-range spline function better reproduces the reference forces than that based on the HI parameterization. Consequently, the subsequent force-matching procedure using analytical potential functions was based on the short-range forces obtained after subtracting the SQE/ESP long-range contributions from the total *ab initio* forces.

Additionally, one should mention that both the ESP and HI spline approximations feature the irregularity in the Si-Si pair force and energy. The comparison of the "irregularity values" (IV) and the merit criteria leads to: $\mathrm{IV}(AI) < \mathrm{IV}(\mathrm{ESP}) < \mathrm{IV}(\mathrm{HI})$ and $\Delta F_{AI} < \Delta F_{\mathrm{ESP}} < \Delta F_{\mathrm{HI}}$. These nonequalities may be interpreted in a way that ΔF and IV correlate with each other. Since the force-matching is performed assuming the pairwise interactions one may therefore conclude, that the specific behaviour of the potential function for the Si-Si interactions is directly related to the many-body interactions in the system. However, the question why these many-body interactions manifest themselves (at least, to the visible extent) only in the Si-Si potential remains unclear.

6.5 Force-matching approach.

The basic idea of the force-matching (FM) method, which was first proposed by Ercolessi and Adams [203], is to obtain realistic effective potentials by making use of a large amount of data available from the first-principles calculations. This approach is sometimes opposed to the method of fitting the potentials to a limited set of the experimentally determined characteristics. Within the FM approach one tries to mimic the *ab initio* forces by forces derived from the model potentials, as closely as possible. The optimization of the force field model is done by the minimization of a cost function in the parameters space whose dimensionality can become relatively large. The extension of the reference data by results of the calculations on numerous structures is aimed at the improvement of the transferability of the force field.

To perform the force-matching minimization one should select a functional form of effective potentials. In the previous section we have presented the FM approach with splines. The purpose of this section is to introduce the force-matching procedure with analytical functions which are not piecewise. Furthermore, it is highly desirable that the choice of the functions can be justified on the theoretical grounds.

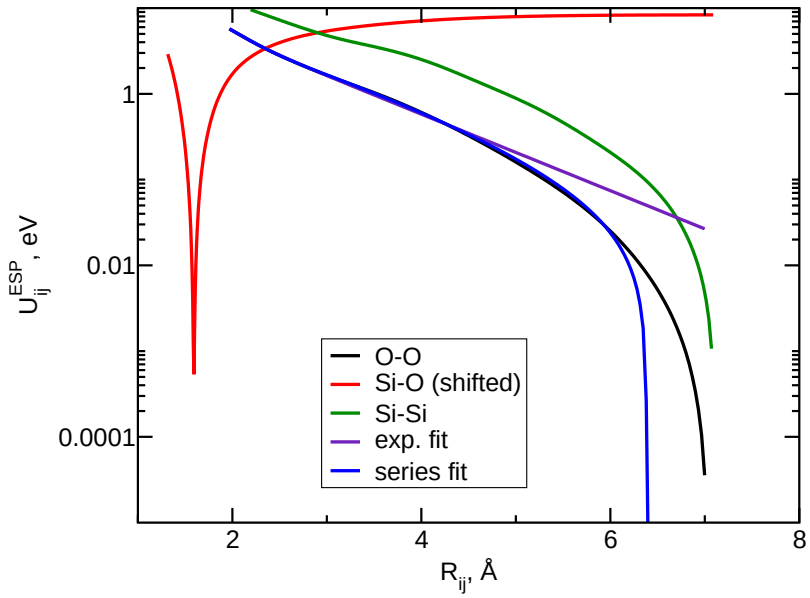

Figure 6.6: ESP short-range energy curves (Figure 6.4) in the logarithmic scale. Curve labeled "exponential fit" shows a fit of the O-O potential curve with function $F(r) = A_0e^{-br} + A_1e^{-2br} + A_2e^{-3br} + A_3e^{-4br}$, and curve labeled "power series fit" corresponds to a fit of the O-O potential with function $F_s(r) = Ae^{-br} - C_4r^{-4} - C_6r^{-6} - C_8r^{-8} - C_{10}r^{-10}$. The potential curve for Si-O was shifted into the positive energies region.

6.5.1 Ansatz for short-range interaction potential.

Asymptotic behaviour of pair potentials.

As it was discussed above, the total energy of the system is described as the sum of long-range and short-range energy contributions

$$E_{\mathrm{TOT}} = E_{\mathrm{LR}} + E_{\mathrm{SR}} \tag{6.5.1}$$

and the short-range energy, in its turn, can be written as the sum of different terms that account for the energy contributions due to the overlap of electron densities at short distances E_{rep}, exchange interactions E_{exc}, induction E_{ind}, and dispersion interactions E_{disp}

$$E_{\mathrm{SR}} = E_{\mathrm{rep}} + E_{\mathrm{ex}} + E_{\mathrm{ind}} + E_{\mathrm{disp}}. \tag{6.5.2}$$

The first term in (6.5.2) accounts for the repulsion interactions, whereas the remaining terms describe the energy of attraction. In order to understand which functional form can better be used for each of the terms, it is instructive to look at the asymptotic behaviour of pair potentials.

To do so we shall use the "ideal" ESP pairwise potentials (see Figure 6.4) approximated with the splines. Thus, representing the O-O energy *vs.* distance in the logarithmic scale (see Figure 6.6) one immediately sees that the energy follows an exponential law at short distances, while a strong deviation from the exponential behaviour is observed at the distances longer than 5

Å. The fit of the O-O spline potential by a power series better approximates the "ideal" two-body potential than the exponential fit. Nevertheless, Figure 6.6 shows that the power series also fail to mimic the long-range tail of the effective potential. The analysis of Figure 6.6 allows us to suggest functional forms for the different energy contributions in (6.5.2).

Analytical form of the short-range potential energy functions.

Repulsion energy. The E_{rep} energy term corresponds to the repulsion interactions due to the overlap of the electron densities at short distances. In most force fields this energy is commonly approximated with either a power function $1/r^n$ $(n > 9)$, as in Lennard-Jones or Stockmayer potentials, or an exponential function, as in Buckingham potential. In agreement with the asymptotical behaviour of the potential at short distances (Figure 6.6), we have chosen the exponential form for the repulsion potential, according to the pioneering work of Bohr [50]. Hence, the E_{rep} energy term is written in the form

$$E_{\text{rep}} = Ae^{-br}. \tag{6.5.3}$$

Attraction energy. The attraction energy consists of several contributions. In force field models these energy contributions are often combined into a single term represented with either a power function $1/r^m$ $(4 \leq m \leq n)$, as in the Lennard-Jones potential, or by an exponential function, as in the Morse potential.

The induction and dispersion energy terms account for the polarization interactions. However, the DFT calculations used to obtain the reference values of the "observables" do not take the dispersion interactions into account. Consequently, the polarization energy reduces to the term due to the induction energy. Within the second-order perturbation theory the induction interaction energy of two species A and B can be represented in the following form

$$E^{(2)}_{\text{ind}} = -\sum_{m\neq 0} \frac{|\langle \psi^{\text{A}}_0 \psi^{\text{B}}_m |V| \psi^{\text{A}}_0 \psi^{\text{B}}_0 \rangle|^2}{E^{\text{B}}_m - E^{\text{B}}_0} - \sum_{n\neq 0} \frac{|\langle \psi^{\text{A}}_n \psi^{\text{B}}_0 |V| \psi^{\text{A}}_0 \psi^{\text{B}}_0 \rangle|^2}{E^{\text{A}}_n - E^{\text{A}}_0}, \tag{6.5.4}$$

where V is the electrostatic potential operator of the interaction between the A and B, ψ^{A}_n and ψ^{B}_m are the wavefunctions of the species A and B characterized by the sets of quantum numbers n and m, respectively, and E^{A}_n and E^{B}_m are the eigenvalues corresponding to the quantum numbers n and m, respectively.

Within the approximation of the spherically symmetric electron density distribution and assuming the potential to be pairwise, one should include the induction contribution that describes the interaction between permanent charges with the multipole moments of the electron density induced by these charges. We truncate the multipole series after the first term, that corresponds to the induced dipoles. Therefore, the induction energy term has the form

$$E^{s}_{\text{ind}} = -\frac{C_4}{r^4}, \tag{6.5.5}$$

where $C_4 \geq 0$, because within the second-order perturbation theory (see (6.5.4)) this coefficient is proportional to the square of the corresponding matrix element of the electrostatic potential operator.

The value of the C_4 coefficient can be obtained using the following considerations. Let us assume that the permanent charges are the point charges and the induced dipoles are the point dipoles. Then, the $C_4^{(ij)}$ coefficient for the interaction of atoms i and j can be expressed in the following form

$$C_4^{(ij)} = Q_i^2 \alpha_j + Q_j^2 \alpha_i, \tag{6.5.6}$$

where Q_i and α_i are the charge and the dipole polarizability of the atom i and Q_j and α_j are the counterpart quantities of the atom j.

Taking terms (6.5.3) and (6.5.5) together one obtains the first model of the short-range interaction potential for the use in the force-matching procedure as

$$E_{\mathrm{SR}} = \sum_{j>i} \left[A_{ij} e^{-b_{ij} r_{ij}} - \frac{C_4^{(ij)}}{r_{ij}^4} \right] + E_{\mathrm{disp}}. \tag{6.5.7}$$

For the sake of convenience we will refer to this potential as *exp-4* potential.

The second model employs the form of an exponential function for the attraction potential. Consequently, the short-range interatomic potential function chosen for the use in the force-matching is Morse potential whose analytical form is given by

$$E_{\mathrm{Morse}} = D_0 (e^{-2b(r-r_0)} - 2e^{-b(r-r_0)}), \tag{6.5.8}$$

where r_0 is the equilibrium distance, D_0 is the well depth, and the b parameter controls the steepness of the potential well. Hence, the second model of the short-range interaction potential is given by

$$E_{\mathrm{SR}} = \sum_{j>i} \left[D_0^{(ij)} (e^{-2b_{ij}(r_{ij}-r_0^{(ij)})} - 2e^{-b_{ij}(r_{ij}-r_0^{(ij)})}) \right] + E_{\mathrm{disp}}. \tag{6.5.9}$$

Hereafter this model is called *Morse-SR* potential.

Although Figure 6.6 indicates that the function (6.5.8) has incorrect asymptotical behaviour at long distances, the Morse potential was successfully used in a number of force field models for solids. Thus, as it was already mentioned, the potential (6.5.8) combined with the charge-equilibration scheme well describes structural transformations of silica polymorphs [297]. Tangney and Scandolo [248] have constructed a very successful polarizable force-field in which a part of the short-range energy was represented with the Morse potential. Recently, Beck *et al.* [298] extended the Tangney and Scandolo model for MgO. These works justify the use of the function (6.5.8) to describe the short-range energy in silica, despite the deficiency mentioned above.

Note that the SR energy expressions (6.5.7) and (6.5.9) contain the dispersion energy term

given by

$$E_{\text{disp}} = -\sum_{j>i} \frac{C_6^{ij}}{r_{ij}^6}, \tag{6.5.10}$$

where the dispersion interactions were considered only between the oxygen atoms and the value of the coefficient C_6^{OO} was set to $C_6^{\text{OO}} = 56.91$ a.u., according to Table 5.3 in Chapter 5.

6.5.2 Basic expressions for derivatives.

Having the pairwise short-range potential $E(\{r_{ij}\})$ one has to obtain the expressions for the forces and the stress tensor components that are to be used in the force-matching procedure. The component $\beta = \{x, y, z\}$ of the force $\mathbf{F}$ acting on atom i is given by

$$F_i^\beta = \sum_{j\neq i} F_{ij}^\beta = -\sum_{j\neq i} \frac{\partial E}{\partial r_{ij}} \frac{\beta_{ij}}{r_{ij}}, \tag{6.5.11}$$

where $\beta_{ij} = \beta_i - \beta_j$, and the expression for the stress tensor component $S_{\alpha\beta}$ reads

$$S_{\alpha\beta} = \sum_{j\neq i} \alpha_{ij} F_{ij}^\beta. \tag{6.5.12}$$

The minimization procedure for finding the optimal set of parameters requires the calculation of the derivatives of the force with respect to the parameters (see Chapter 2, Levenberg-Marquardt method). For *exp-4* potential these derivatives are

$$\frac{\partial F_{ij}^\beta}{\partial A_{ij}} = b_{ij} e^{-b_{ij} r_{ij}} \frac{\beta_{ij}}{r_{ij}}, \tag{6.5.13}$$

$$\frac{\partial F_{ij}^\beta}{\partial b_{ij}} = A_{ij}(1 - b_{ij}) e^{-b_{ij} r_{ij}} \frac{\beta_{ij}}{r_{ij}}. \tag{6.5.14}$$

For the *Morse-SR* potential the equations are

$$\frac{\partial F_{ij}^\beta}{\partial D_0^{(ij)}} = \frac{F_{ij}^\beta}{D_0^{(ij)}} \frac{\beta_{ij}}{r_{ij}}, \tag{6.5.15}$$

$$\frac{\partial F_{ij}^\beta}{\partial b_{ij}} = F_{ij}^\beta \left(1 - b_{ij}(r - r_0)\right) - 2D_0^{(ij)} b_{ij}(r - r_0) e^{-2b_{ij}(r-r_0)} \frac{\beta_{ij}}{r_{ij}}, \tag{6.5.16}$$

$$\frac{\partial F_{ij}^\beta}{\partial r_{0ij}} = b_{ij} F_{ij}^\beta + 2D_0^{(ij)} b_{ij}^2 e^{-2b_{ij}(r-r_0)} \frac{\beta_{ij}}{r_{ij}}. \tag{6.5.17}$$

Using eqn. (6.5.12) together with eqs. (6.5.13) and (6.5.14) one can write the expressions for the derivatives of the stress tensor components with respect to the model parameters:

$$\frac{\partial S_{\alpha\beta}}{\partial A_{ij}} = \sum_{j\neq i}\alpha_{ij}\frac{\partial F_{ij}^{\beta}}{\partial A_{ij}} = \sum_{j\neq i}\alpha_{ij}b_{ij}e^{-b_{ij}r_{ij}}\frac{\beta_{ij}}{r_{ij}}, \tag{6.5.18}$$

$$\frac{\partial S_{\alpha\beta}}{\partial b_{ij}} = \sum_{j\neq i}\alpha_{ij}\frac{\partial F_{ij}^{\beta}}{\partial b_{ij}} = \sum_{j\neq i}\alpha_{ij}A_{ij}(1-b_{ij})e^{-b_{ij}r_{ij}}\frac{\beta_{ij}}{r_{ij}}. \tag{6.5.19}$$

Similarly, one obtains for the *Morse-SR* potential:

$$\frac{\partial S_{\alpha\beta}}{\partial D_0^{(ij)}} = \sum_{j\neq i}\alpha_{ij}\frac{\partial F_{ij}^{\beta}}{\partial D_0^{(ij)}} = \sum_{j\neq i}\alpha_{ij}\frac{F_{ij}^{\beta}}{D_0^{(ij)}}\frac{\beta_{ij}}{r_{ij}}, \tag{6.5.20}$$

$$\begin{aligned}\frac{\partial S_{\alpha\beta}}{\partial b_{ij}} &= \sum_{j\neq i}\alpha_{ij}\frac{\partial F_{ij}^{\beta}}{\partial b_{ij}} \\ &= \sum_{j\neq i}\alpha_{ij}\left[F_{ij}^{\beta}\left(1-b_{ij}(r-r_0)\right) - 2D_0^{(ij)}b_{ij}(r-r_0)e^{-2b_{ij}(r-r_0)}\frac{\beta_{ij}}{r_{ij}}\right],\end{aligned} \tag{6.5.21}$$

$$\frac{\partial S_{\alpha\beta}}{\partial r_{0ij}} = \sum_{j\neq i}\alpha_{ij}\frac{\partial F_{ij}^{\beta}}{\partial r_{0ij}} = \sum_{j\neq i}\alpha_{ij}\left[b_{ij}F_{ij}^{\beta} + 2D_0^{(ij)}b_{ij}^2e^{-2b_{ij}(r-r_0)}\frac{\beta_{ij}}{r_{ij}}\right]. \tag{6.5.22}$$

Shifted-force potential. In calculations where the size of the modelled system is limited, the interactions between atoms beyond a certain distance, cut-off radius R_c are neglected. However, despite the fact that the interactions are short-ranged and the cut-off radius is generally large enough, such a simple truncation of potential introduces a discontinuity in the potential that leads to a bad conservation of the energy in MD simulations. To remedy the problem, the potential is shifted as

$$V_s(r) = \begin{cases} V(r) - V(R_c), & r \leq R_c \\ 0, & r > R_c \end{cases}. \tag{6.5.23}$$

As the additional term in (6.5.23) is a constant for any pair of atoms, the modification does not affect the forces and therefore, the equations of motion. However, with the use of the shifted potential (6.5.23) the force has a discontinuity at R_c. To avoid this, one should use a shifted-force potential [300]

$$V_{sf}(r) = \begin{cases} V(r) - V(R_c) - \left(\frac{dV(r)}{dr}\right)_{r=R_c}(r-R_c), & r \leq R_c \\ 0, & r > R_c \end{cases}. \tag{6.5.24}$$

The force obtained by (6.5.24) goes smoothly to zero as the distance approaches R_c. This removes the problems with the energy conservation and the instabilities in the integration of the equations of motion in molecular dynamics simulations.

In the force-matching procedure and in the MD simulations performed in the present work

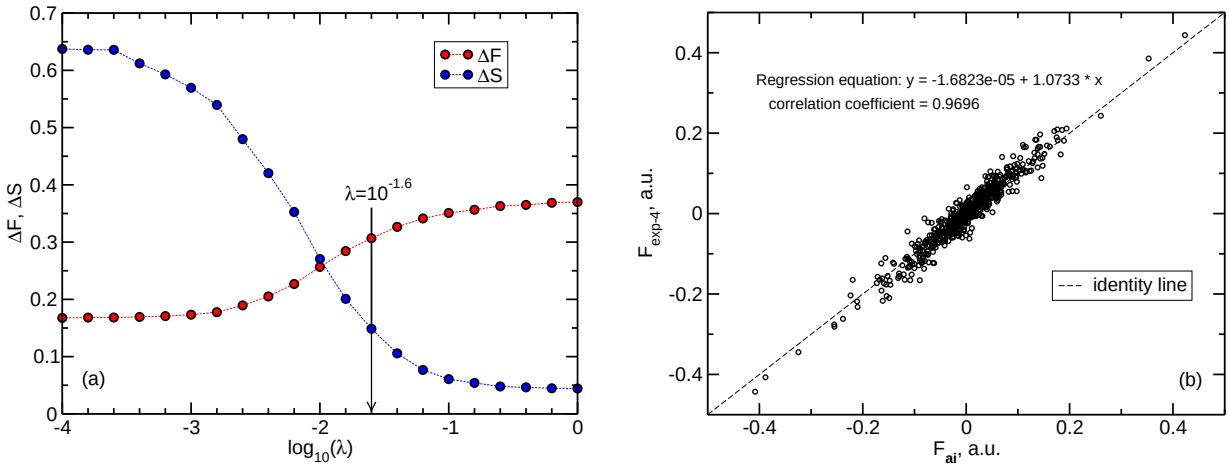

Figure 6.7: (a): dependence of the merit functions ΔF (6.2.4) and ΔS (6.2.5) on the coefficient λ for the *exp-4* potential. (b): correlation between the *ab initio* forces and the forces computed with the model.

the cutoff radius was taken to be equal to $R_c = \max(L_1, L_2, L_3)/2$, where L_i are the lengths of simulation box dimension sides. In the case of the force-matching performed on the configurations of amorphous silica structure with the simulation box sides $L = 14.263$ Å, the cut-off radius is equal to $R_c = 7.13$ Å.

6.5.3 Parameters of the short-range potential.

exp–4 potential.

Before starting the minimization we have fixed the value of the $C_4^{(ij)}$ coefficient according to eq. (6.5.6). The dipole polarizabilities of atoms were set to $\alpha_{\mathrm{O}} = 1.527$ Å^3 (see Table 5.3 in Chapter 5) and $\alpha_{\mathrm{Si}} = 0.033$ Å^3 [299]; values of atomic charges for each atomic kind were fixed to the mean values obtained by averaging the charges for each atomic species over all 30 configurations used in the force-matching procedure. These mean charges are equal to $Q_{\mathrm{O}} = -0.67$ a.u. and $Q_{\mathrm{Si}} = 1.34$ a.u.

To ensure a good exploration of the parameters space, each minimization was performed for different initial values of the parameters. The final results were found to be very sensitive to the initial guess, which means that there are several local minima in the meaningful region of parameters space ($A_{ij} \geq 0$ and $b_{ij} \geq 0$). To ensure that the optimized parameters lie in the correct region, boundary conditions were implemented into minimization procedure.

The total χ^2_{TOT} was constructed using eq. (6.2.3) and the value of the λ parameter was varied within the interval $[10^{-4}, 10^{0}]$. For each value of λ the model parameters A_{ij} and b_{ij} were optimized, whereas the C_4^{ij} coefficients were kept fixed at the values determined as discussed above. The dependence of the merit functions ΔF (6.2.4) and ΔS (6.2.5) on the value of the coefficient λ is shown in Figure 6.7 (a).

According to Figure 6.7 (a) we have chosen the parameters at $\lambda = 10^{-1.6}$ as a compromise

Table 6.1: Parameters of the *exp-4* model potential, eq. (6.5.7).

	A_{ij}, eV	b_{ij}, Å^{-1}	$C_4^{(ij)}$, eV Å^4
O-O	864.2509	2.4690	19.83
Si-O	18656.0874	5.4664	39.86
Si-Si	8287.3355	3.0645	1.71

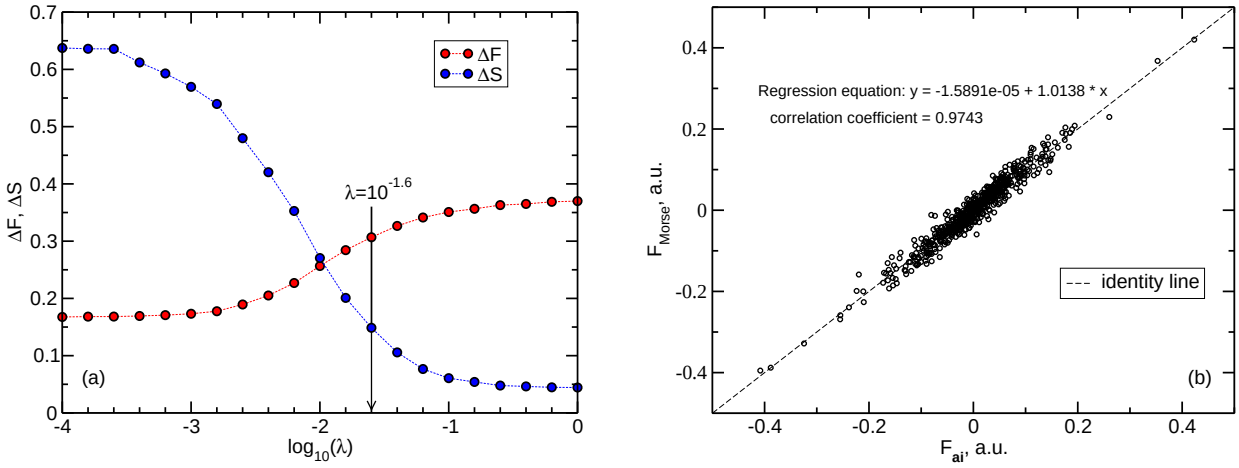

Figure 6.8: (a): dependence of the merit functions ΔF (6.2.4) and ΔS (6.2.5) on the coefficient λ for the *Morse-SR* potential. (b): correlation between the *ab initio* forces and the forces computed with the model.

between the quality in the description of the forces and the stress. Values of the parameters of the *exp-4* potential are presented in Table 6.1. Figure 6.7 (b) compares the total *ab initio* force with the forces computed using the *exp-4* model for one randomly chosen configuration and the figure illustrates a good agreement between the reference and model values.

Morse potential.

The same type of minimization procedure was performed for the *Morse-SR* potential. The boundary conditions $D_0^{(ij)}$, b_{ij}, $r_0^{(ij)} \geq 0$ were implemented to ensure the physically meaningful values of the parameters. The dependence of the merit functions ΔF (6.2.4) and ΔS (6.2.5) on the parameter λ value is shown in Figure 6.8 (a).

In agreement with the figure the parameters corresponding to the value of $\lambda = 10^{-1.6}$ were chosen for the use in the MD simulations. Figure 6.8 (a) presents the correlation between the *ab initio* and *Morse-SR* model forces obtained for the same reference configuration as for the *exp-4* model (Figure 6.7 (b)). Table 6.2 contains the values of the parameters obtained in the force-matching procedure for *Morse-SR* potential.

Results obtained with the *exp-4* and *Morse-SR* potential models are reported in the next chapter.

Table 6.2: Parameters of the *Morse-SR* model potential, eq. (6.5.9).

	$D_0^{(ij)}$, eV	b_{ij}, Å^{-1}	r_0^{ij}, Å
O-O	0.0457	1.2299	3.8218
Si-O	1.1691	2.6588	1.7236
Si-Si	0.0451	1.7491	3.6853

Chapter 7

Polarizable force field models for silica.

7.1 Introduction.

Development of the effective potential models for silica has a long history. Certainly, silicon dioxide plays the same role in the atomistic modeling of inorganic solids as water does in the modeling of molecular systems. The analogy is amplified by the fact that despite the long-lasting efforts, no force field capable of reproducing the complete palette of properties in a wide range of temperatures and pressures has been suggested for both the compounds yet.

First silica force fields made use of rigid ion model thus reflecting the semi-ionic character of the chemical bond. It has been realized that the inclusion of the polarization can significantly improve the agreement of the predicted characteristics with the experimental reference data. The phenomenon has been taken into account with the use of the shell model, later on other approaches have appeared some of which were discussed in Chapter 3. The end of 1980s has seen the appearance of silica force fields based on the *ab initio* calculations and several models have been put forward during the subsequent years [52,301–306]. Much of effort in that period has been put at the development of force fields for the microporous low-density silica modifications such as zeolites [302–307]. These models had a success in modeling of specific system characteristics, whereas they often suffered of limited transferability. A number of polarizable silica models have been developed since the late 1990s and revised versions of the RIM model appear regularly [288].

An effective potential model for silica based on the CPE approach has been proposed by Demiralp *et al.* [297]. The authors used the QEq scheme by Rappé and Goddard [168] to compute geometry-dependent charges in the system, whereas the short-range part of the potential energy has been represented with a Morse potential function (MS). The parameterization of the MS-QEq model was performed by using the experimental data (the cohesive energies, elastic moduli, density, etc.) of the α-quartz and stishovite structures. The MS-QEq model yields the structures of crystalline polymorphs in agreement with the experimental data, while it slightly (by 5 %) overestimates the density of silica glass. From the analysis of their results, the authors [297]

concluded that the three-body energy term is not necessary to describe the silica structures with a wide range of the Si-O-Si angles.

Another polarizable model for silica has been developed by Tangney and Scandolo [248] (TS) who used results of DFT calculations on liquid SiO_2 and the force-matching procedure to derive the parameters of their force field. The authors described the pair-wise interactions in the system as the sum of Coulomb and Morse potentials. The polarizable part of the TS force field has been represented by the induced point dipole model extended by a short-range contribution of the induced dipoles due to the interaction of oxygens polarizability with charges. Despite the fact that the TS model was designed using the data obtained for liquid silica, the authors of Ref. [248] found that their force field also describes the crystalline silica structures, such α-quartz, α-cristobalite, and coesite, thus revealing a good transferability. Recently, the TS force field was reformulated in a way avoiding the use of time-expensive Ewald summation [308]. The new model was re-parametrized and it was shown to produce the experimental characteristics of silica glass and α-quartz with a high degree of accuracy.

A thorough comparison of the MS-QEq and the original TS models with the BKS potential has been performed by Herzbach and co-workers [309]. The analysis of the vibrational density of states for the α-quartz structure showed the best performance of the TS model followed by the MS-QEq one. The TS potential also reproduced the structural pecularities of the crystalline silica polymorphs such as the c/a-ratio at $\alpha \rightarrow \beta$-quartz transition or the metastability of the β-cristobalite phase. However, the TS potential performed worse than the BKS one for the equation of state of the six-fold coordinated stishovite structure. Herzbach *et al.* [309] pointed to the fact that despite the notable improvement of the silica force field with the TS model, there remains a large room for further advances in this area.

An interesting and promising approach to the development of polarizable force field has been suggested by Tabacchi and co-workers [208]. The approach is directly related to the CPE formulation suggested by York and Yang [162]. The authors [208] based their model on the second-order expansion of the energy functional (2.1.32) with respect to the change of the external potential δv and the response density $\delta\rho$. The energy functional contains both the classical energy terms related to the electrostatic interactions between the charge densities and those with the nuclei, as well as a nonclassical short-range energy due to the $F[\rho]$ functional including the kinetic and XC contributions. Following the suggestion of York and Yang [162], the response density is expanded in a set of basis functions and the final expression for the energy functional needs to be minimized with respect to the expansion coefficients. Tabacchi *et al.* applied their approach to the development of the force field for lithium iodide. The parameterization of the model has been performed with the force-matching procedure making use the results of CPMD calculations. The resulting polarizable force field mimicked the structural properties of the LiI system with the accuracy comparable to that of CPMD calculations. Despite the success of the approach the authors of Ref. [208] have mentioned several issues related to the transferability of the parameters between different systems, such as the choice of the reference state for subsystems and the related computation of the hardness kernel.

The analysis of the literature data shows that last decade has seen a significant progress in the development of polarizable force fields for condensed phase systems. Several approaches have emerged, while no a universal model has been put forward so far. New studies in this direction are therefore of a great interest. The present Chapter reports first results of the application of the polarizable force field models discussed in Chapter 6 to silica structures. Results of these calculations are systematically compared with those obtained by one of the most successful non-polarizable model, namely the BKS potential [52]. Furthermore, whenever possible, the outcome of the calculations is also compared with the data obtained by using other polarizable force field models, and with the experimental data.

7.2 Computations.

The most of the results described in the present chapter were obtained with the use of the molecular dynamics simulation technique. The SQE model was included into the MD method in a straightforward way: the SQE equation (4.2.2) was solved for each configuration along the molecular dynamics trajectory thus yielding the geometry-dependent charges on atoms. Such a scheme, although being less computationally effective than the extended Lagrangian approach in the spirit of the Car-Parrinello method, warrants the fulfillment of the adiabatic approximation, i.e. the atomic charges are exactly those that minimize the potential energy. Furthermore, such an implementation allows using a large integration time-step equal to that of fixed-charge MD simulations. It is important to note that the use of the CPE approach with the geometry-dependent atomic charges does not modify the general expression for the long-range atomic forces (see Appendix E); because of the same reason no modification of the stress tensor computations is necessary.

Two types of MD simulations were carried out. The first type used the NPT statistical ensemble and these simulations were aimed at finding the equilibrium structural parameters of the system. This type of calculations employed the Rahmam-Parrinello algorithm coupled with a chain of Nosé-Hoover thermostats to keep the pressure and temperature constant, respectively. The equations of motion were integrated with the four-order predictor-corrector scheme and the fictitious mass associated with the simulation box parameters was equal to 20 times the mass of the Si atoms. The frequency parameter in the Nosé-Hoover method was taken to be equal to 1200 cm^{-1}, that approximately corresponds to the Debye frequency of silica. The second type of MD simulations was performed in the NVE statistical ensemble and they were used for computing the equilibrium characteristics of the system. In this type of simulations, the classical equations of motion were integrated with the velocity form of Verlet algorithm.

For all studied potential models the long-range electrostatic interactions were treated with the Ewald sum method. The cut-off radius for the direct space part of the Ewald sum and for the short-range interactions was taken to be equal the half of the minimum side of the simulation box, but not more than 12 Å. The shifted-force modification of the short-range potentials were used to ensure the smoothness of the potential and the force at the cut-off distance.

The integration time-step of 1 fs was used in both the NPT and NVE simulations. Typical

NPT simulation consisted of 35 000 time-steps from which the first 5 000 steps were used to equilibrate the system in the NVT ensemble. This equilibration period was followed by 30 000 steps in the NPT ensemble and the information about the simulation box parameters was collected for the last 20 000 steps. A typical NVE simulation consisted of 55 000 time-steps. The first 10 000 were used as the equilibration period, whereas the atomic coordinates, velocities, and charges were saved for a subsequent analysis each 10-th time-step during the last 40 000 steps.

Table 7.1: Number of atoms N, shape, and size of simulation cell for studied systems.

System	N	Simulation cell shape	Simulation cell parameters (Å)
α-quartz	576	trigonal	19.6536 × 19.6536 × 21.6208
α-cristobalite	576	orthorhombic	20.1248 × 20.1248 × 21.0534
β-cristobalite	648	cubic	21.4980 × 21.4980 × 21.4980

The parameters of the simulation box for the studies systems (initial configurations) are given in Table 7.1. The systems were characterized by computing the pair radial distribution functions, the bond angle distribution functions for the Si-O-Si and O-Si-O angles, as well as by the calculation of the complete set of the vibrational spectra, i.e. the density of vibrational states, the infrared and Raman spectra. Besides the Raman spectra corresponding to the powder samples, the polarized Raman spectra were computed as well. The computation of the polarized Raman spectra provides a very rigorous test of the symmetry properties of the atomic displacements in the vibrational modes. Multiple time-origins were used to reduce the noise in the calculated spectra.

The radial distribution function were computed with the use of (2.2.37). The time-correlation function formalism was employed for computing the vibrational spectra, as discussed in Chapter 2. The density of the vibrational states was obtained via the Fourier transform of the auto-correlation function of the mass-weighted Cartesian atomic velocities, equation (2.2.38). In the frame of the time-correlation functions formalism the expression for the infrared absorption coefficient $\alpha(\omega)$ reads [44, 47]

$$\alpha(\omega) = \frac{4\pi^2\omega}{3\hbar cn}\big(1 - \exp(-\hbar\omega/k_BT)\big)\int dt\, \exp(-\imath\omega t)\langle \mathbf{D}(t)\cdot\mathbf{D}(0)\rangle, \tag{7.2.1}$$

where c is the speed of light in the vacuum, n is the refraction index of the medium, and other quantities in (7.2.1) have their usual meaning. The $\mathbf{D}(t)$ quantity in (7.2.1) stands for the time-dependent dipole of the system which was calculated as

$$\mathbf{D}(t) = \sum_{i=1}^{N} q_i(t)\,\mathbf{r}_i(t) \tag{7.2.2}$$

with $q_i(t)$ and $\mathbf{r}_i(t)$ being the charge and the coordinate vector of the atom i, respectively.

Within the same approach the Raman spectrum is obtained via the auto-correlation function of the polarizability tensor [44, 48]. Thus, the intensity of the Raman scattering in the experimental

geometry $c(ab)c$ [†]

$$I_{c(ab)c} = \lambda_s^{-4} \int dt\, e^{\imath\omega t} \langle A_{ab}(0) \cdot A_{ab}(t) \rangle, \tag{7.2.3}$$

where $2\pi\lambda_s$ is the wavelength of the scattered radiation and A_{ab} stands for the ab component of the polarizability tensor (a, b, $c = x$, y, z). If the Cartesian axes coincide with the axes of crystal, the spectra of different tensor components reflect the Raman activity of the vibrational modes of a specific symmetry.

The computation of the time-dependent polarizability tensor was carried out using the bond-polarizability model [310, 311] that writes the polarizability tensor of the system $\mathbf{A}$ in Cartesian coordinated as

$$\mathbf{A}(t) = \sum_{i \in \text{bonds}} \mathbf{R}_i(t) \Big[\mathbf{a}_i^0 + \mathbf{a}_i' \big(r_i(t) - r_i^0 \big) \Big] \mathbf{R}_i^{-1}(t) \tag{7.2.4}$$

where the sum runs over bonds i, $\mathbf{R}_i$ is the matrix that rotates the principal axes of the bond i into the Cartesian ones, and the expression in the square brackets stands for the polarizability tensor of the bond i in the bond frame. The elements of the diagonal tensors $\mathbf{a}_i^0$ and $\mathbf{a}_i'$ are the longitudinal and transversal bond polarizability components and their derivatives with respect to the bond length, respectively. These elements are often referenced to as electro-optical parameters. Values of the parameters used in the calculations were equal to those obtained by Umari *et al.* [312] in the periodic DFT calculations of crystalline silica polymorphs.

7.3 Results and Discussion.

7.3.1 *exp-4* potential.

α-quartz.

Table 7.2 compares parameters of the α-quartz structure obtained in the NPT calculation with the experimental data and with the results of calculations using the BKS and TS force field models. The analysis of the table shows that the *exp-4* model contracts the a lattice constant and expands the c one that results in too large c/a ratio. Nevertheless, the structure remains stable along the MD trajectory and the standard deviation of the α, β and γ angles from their reference values does not exceed 0.6 deg.

Figure 7.1 (left panel) presents the Si-O radial distribution function computed in NVE MD run and compares it with $g_{\text{SiO}}(r)$ function obtained with the BKS potential. The first peak in $g(r)$ calculated with the *exp-4* potential is found at 1.72 Å and it is shifted by 0.11 Å to longer distances as compared to the experimental Si-O bond length. The longer Si-O bond length alone is not surprising since the potential was parameterized on the data obtained with GGA-family XC functional, which is known to overestimate bond lengths. The warning sign, however, is that the 7 % increase in the Si-O bond length yields a non-uniform change of the lattice parameters

[†]The experiment is carried out in the backscattering geometry with the light propagating along the axis c, while the polarization of the incident radiation is parallel to the axis a and the scattered light is analyzed with the polarization parallel to the axis b.

Table 7.2: Lattice parameters of the α-quartz structure computed with the *exp-4* potential.

Parameter	Present work	Exp. [313]	BKS [52]	TS [248]
a, Å	4.828 (0.004)	4.9134	4.941	4.925
c, Å	5.655 (0.007)	5.4052	5.449	5.386
c/a	1.171	1.100	1.103	1.094

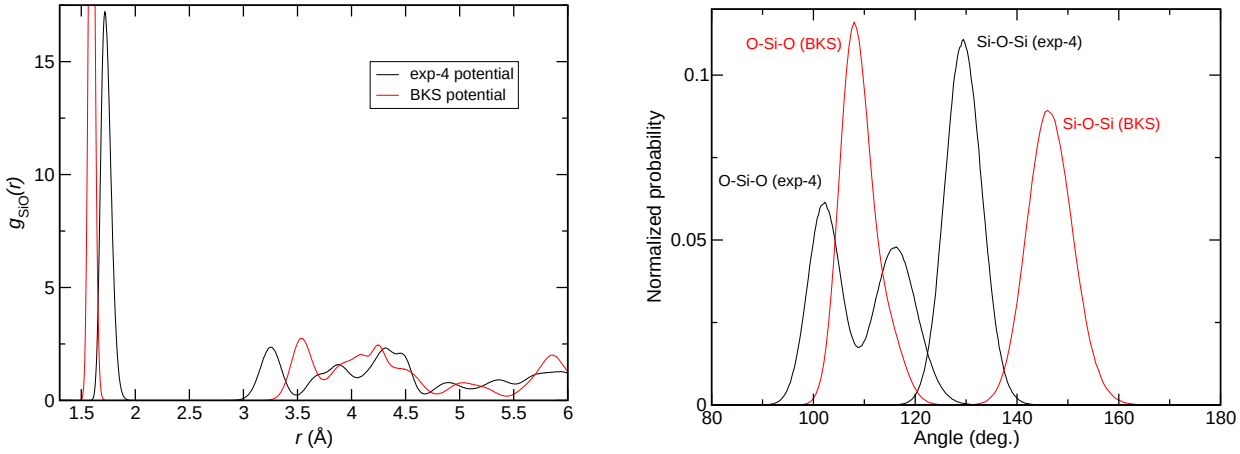

Figure 7.1: Si-O radial distribution function (left), and O-Si-O and Si-O-Si angle distribution functions (right) in α-quartz computed with the *exp-4* (black line) and BKS (red line) models.

such that the a and c constants become smaller and larger, respectively, than the corresponding experimental values. Such a result suggests that there should be some distortion of the bond angles. Figure 7.1 (right panel) shows the O-Si-O and Si-O-Si bond-angle distributions in the α-quartz computed with the *exp-4* and BKS potentials. Indeed, the O-Si-O angles distribution obtained with the former model has a bimodal character pointing to a distortion of the SiO_4 tetrahedra, whereas the calculation with the BKS potential results in a unimodal O-Si-O angle distribution centered at the tetrahedral angle. Furthermore, the angle of Si-O-Si bridges in the *exp-4* model has a mean value of ca. 129 deg. that is significantly smaller than the experimental value of 144 deg. and the angle of 148 deg. in the BKS model. The outcome of the test calculations reveals that the *exp-4* model fails to reproduce the correct structure of α-quartz

α-cristobalite. The NPT MD simulation of the α-cristobalite structure with the *exp-4* potential has shown instability of this crystalline phase. A spontaneous structural transition of α-cristobalite into a new structure was observed, where the shape of the simulation box remained orthorhombic, while the side lengths changed from $20.125 \times 20.125 \times 21.053$ Å to $16.4 \times 16.4 \times 25.3$ Å and the volume of the system decreased by ca. 20 % during the transition. Snapshots of the initial and final configurations are shown in Figure 7.2. The final structure strongly resembles the β-cristobalite phase, but contains a large number of too long Si-O contacts with the lengths more than 2 Å. The density of the new phase is ca. 2.8 g cm^3 that is larger than the density

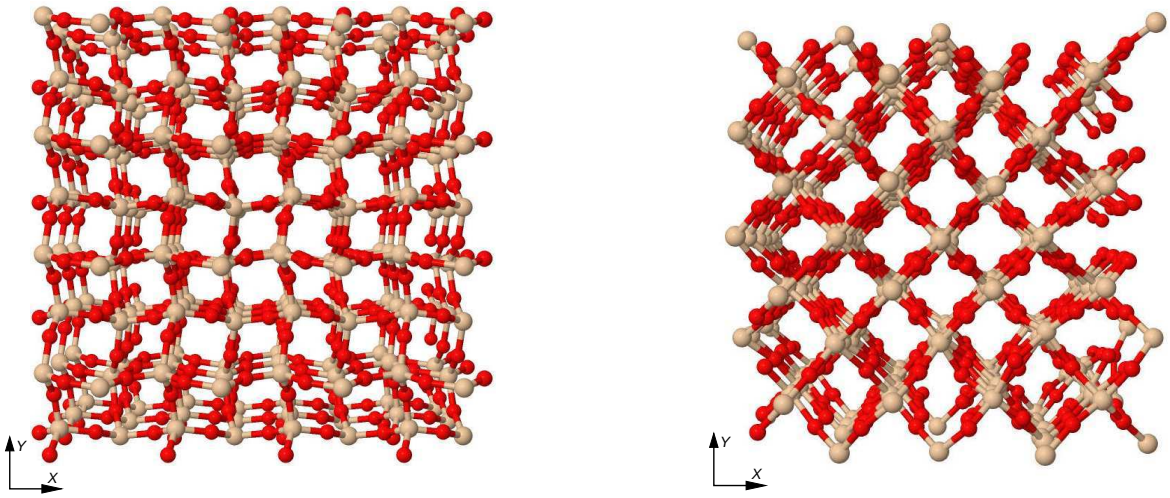

Figure 7.2: Snapshots of the initial (left) and final (right) configurations of the α-cristobalite structure from the NPT MD run with the *exp-4* model; prespective view is shown.

of cristobalite and quartz. The outcome of the α-cristobalite simulations is obviously an artifact that further highlights the failure of the *exp-4* model as a candidate for the polarizable force field for silica.

7.3.2 *Morse-SR* potential.

α-quartz.

Table 7.3 lists parameters of the α-quartz structure obtained in the NPT calculation using the *Morse-SR* potential and the reference data. The table shows that the computed values of the a and c lattice constant as well as the c/a ratio are in a good agreement with the experimental data and with the results obtained with other force field models. The mean value of the Si-O-Si angle is slightly overestimated by the model in comparison with the reference data. In contrast to the results obtained with the *exp-4* potential the O-Si-O angles in the SiO_4 tetrahedra has a unimodal distribution centered at 109.2 deg.

Table 7.3: Lattice parameters of the α-quartz structure obtained with the *Morse-SR* potential.

Parameter	Present work	Exp. [313]	BKS [52]	TS [248]
a, Å	4.963 (0.013)	4.9134	4.941	4.925
c, Å	5.439 (0.019)	5.4052	5.449	5.386
c/a	1.096	1.100	1.103	1.094
Si-O-Si, deg.	152.6	143.7	148.1	144.5

Figure 7.3 presents the Si-O and O-O radial distribution functions. The *Morse-SR* model yields the mean value of the Si-O bond length equal to 1.575 Å that is close to the experimental

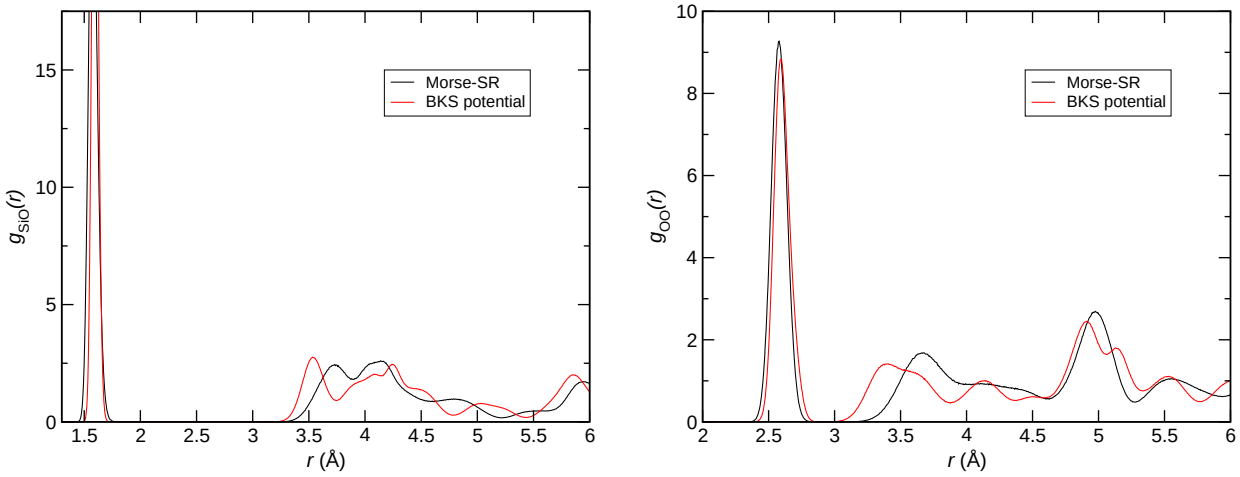

Figure 7.3: Si-O (left) and O-O (right) radial distribution functions for α-quartz in the *Morse-SR* model.

Si-O bond length of 1.60 Å and to the mean Si-O bond length of 1.597 Å in the BKS model. The first peak in the O-O radial distribution function computed with the *Morse-SR* and BKS models is at the distance of 2.58 and 2.59 Å, respectively. The difference between the RDFs computed with the two models for larger distances is obviously due to the difference in the Si-O-Si angles. The larger value of the angle obtained with the *Morse-SR* potential (Table 7.3) leads to the shift of the second and further peaks in $g(r)$ toward larger distances, in comparison to the reference BKS functions. The results presented above show that the *Morse-SR* model well reproduces the structural characteristics of α-quartz.

Figure 7.4 presents the density of vibrational states of the structure computed with the two models. The striking difference between the two spectra is the downward shift of the *Morse-SR* computed spectrum in comparison with the BKS one. The high-frequency edge of the density of vibrational states in silica is determined by the antisymmetric stretching vibrations of the Si-O bonds and too low energy of these vibrational modes points to a small value of the effective Si-O force constant in the *Morse-SR* model. Making use of the data reported in Table 6.2, estimated value of the force constant is equal to 2.65 mdyne/Å that is indeed too small as compared to a typical value of the Si-O force constant 4-5 mdyne/Å in valence force fields [305–307]. It is noteworthy that the MS-QEq force field by Demiralp *et al.* [297] is characterized by the Si-O force constant of 4.75 mdyne/Å and conseqently, the MS-QEq model correctly reflects the density of vibrational states of α-quartz [309]. Such a modest performance of *Morse-SR* model for the description of the vibrational dynamics is surprising since the potential yields a correct value of the Si-O bond length in the system and could be expected to mimic the interatomic force constants as well.

The point group of the α-quartz structure is D_3 and the Raman tensor has the following form [314]

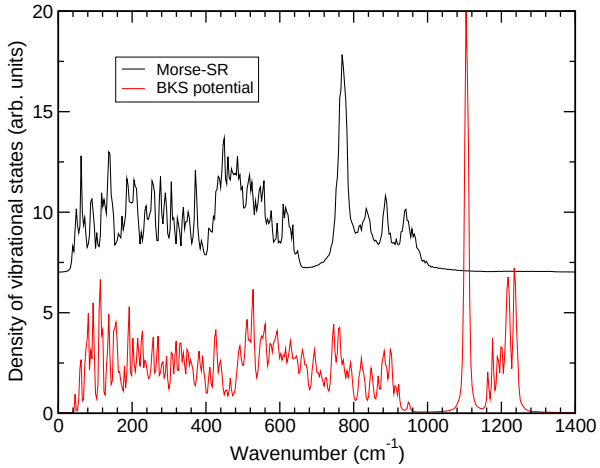

Figure 7.4: Density of vibrational states for α-quartz obtained with the *Morse-SR* model (black) and BKS potetial (red).

$$\begin{pmatrix} A_{xx} & & \\ A_{yx} & A_{yy} & \\ A_{zx} & A_{zy} & A_{zz} \end{pmatrix} = \overset{A_1}{\begin{pmatrix} a & & \\ & a & \\ & & b \end{pmatrix}} + \overset{E}{\begin{pmatrix} c & \\ -c & -c \\ -d & d \end{pmatrix}}.$$

The vibrational modes of the A_1 symmetry can be observed in the Raman spectrum measured in the $x(zz)x$ experimental geometry, whereas the E-modes change the off-diagonal elements of the polarizability tensor and contribute to the variation of the diagonal A_{xx} and A_{yy} components. To observe the modes of the E symmetry, the experiment should be conducted in, e.g. $x(yz)x$ geometry. Figure 7.5 compares the polarized Raman spectra computed using the *Morse-SR* and BKS potentials with the experimental Raman spectra. The theoretical spectra corresponding to the $x(zz)x$ and $x(yz)x$ experimental geometries were obtained by eqn. (7.2.3) for the A_{zz} and A_{yz} components of the polarizability tensor. The analysis of Figure 7.5 shows that the *Morse-SR* potential yields the correct symmetry of the vibrational modes, if one takes the downward shift of the Si-O stretching vibrations into account.

α-cristobalite.

The calculated lattice parameters of the α-cristobalite structure are reported in Table 7.4 which shows that the *Morse-SR* model reproduces the structure with the quality close to that of other potential models. The only notable difference is a small value of the c parameter obtained in the present calculations. The calculated Si-O bond length in the structure is equal to 1.580 Å that is smaller than the mean experimental bond length of 1.602 Å in the structure as determined by Pluth *et al.* [316]. The underestimation of the Si-O bond length seems to be inherent to the *Morse-SR* potential model.

Figure 7.6 shows the computed density of vibrational states (left panel) and the computed and measured powder Raman spectra (right panel) of α-cristobalite. Similarly to the spectra shown in

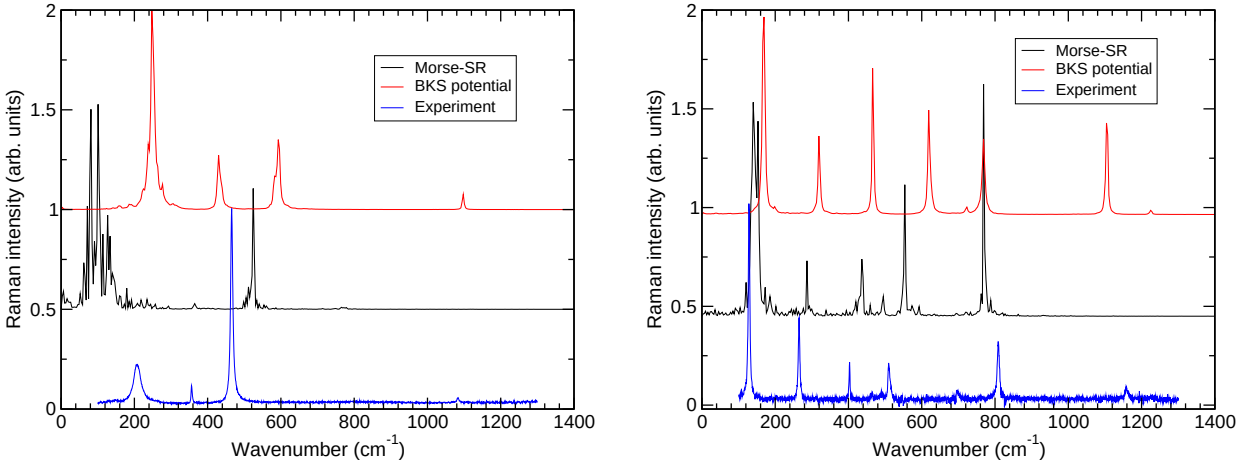

Figure 7.5: Polarized Raman spectra of α-quartz computed for the $x(zz)x$ (left) and $x(yz)x$ (right) geometries. The experimental spectra were taken from Ref. [315].

Table 7.4: Lattice parameters of the α-cristobalite structure obtained with the *Morse-SR* potential.

Parameter	Present work	Exp. [316]	BKS [52]	TS [248]
a, Å	4.947 (0.075)	4.9709	4.920	4.936
c, Å	6.427 (0.180)	6.9278	6.602	6.847
Si-O-Si, deg.	143.3	146.7	148.5	144.0

Figure 7.4 and Figure 7.5, the high-frequency edge in the *Morse-SR* calculated spectra in Figure 7.6 is shifted to lower wavenumbers in comparison with the experimental and BKS-computed spectra. Despite this deficiency, the general behaviour of the simulated spectrum agrees with the reference.

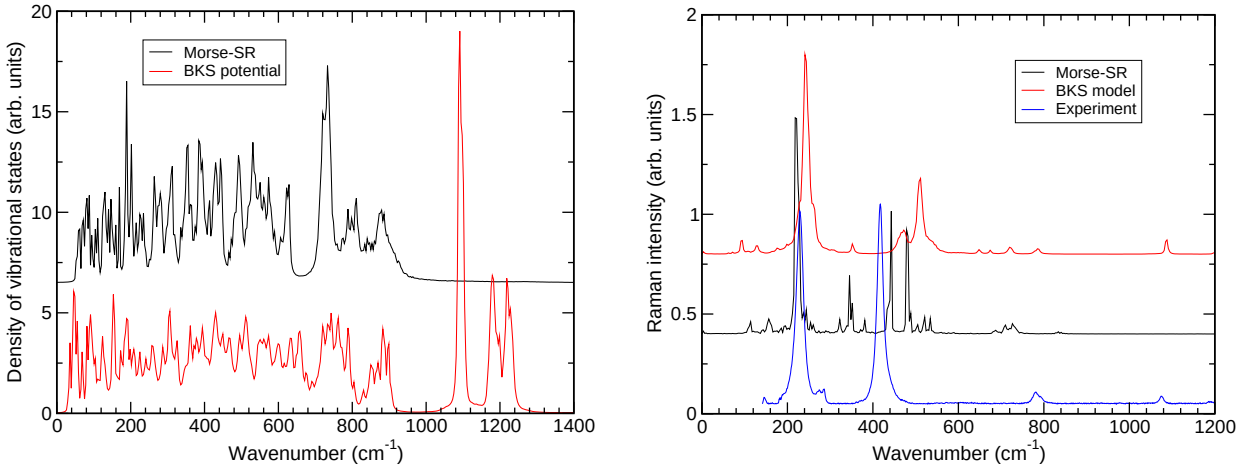

Figure 7.6: Density of vibrational states (left) and the powder Raman spectra (right) of α-cristobalite. The experimental spectrum was taken from Ref. [317].

Additional test of the proper description of the vibrational dynamics of the system was done by

computing the Raman activities the vibrational modes of different symmetries. The Raman active modes of the α-cristobalite structure (D_4 point group) are distributed over irreducible symmetry representations as follows

$$\Gamma_R = 4A_1 + 5B_1 + 4B_2 + 8E.$$

and the Raman tensor has the following form [314]

$$\begin{pmatrix} A_{xx} & & \\ A_{yx} & A_{yy} & \\ A_{zx} & A_{zy} & A_{zz} \end{pmatrix} = \underset{A_1}{\begin{pmatrix} a & & \\ & a & \\ & & b \end{pmatrix}} + \underset{B_1}{\begin{pmatrix} c & & \\ & -c & \\ & & \end{pmatrix}} + \underset{B_2}{\begin{pmatrix} & & \\ d & & \\ & & \end{pmatrix}} + \underset{E}{\begin{pmatrix} & & \\ & & \\ e & e & \end{pmatrix}}.$$

Consequently, computation of the spectra of the polarizability tensor elements (and of their linear combinations) permits to assign the Raman active vibrations to the modes of specific symmetry.

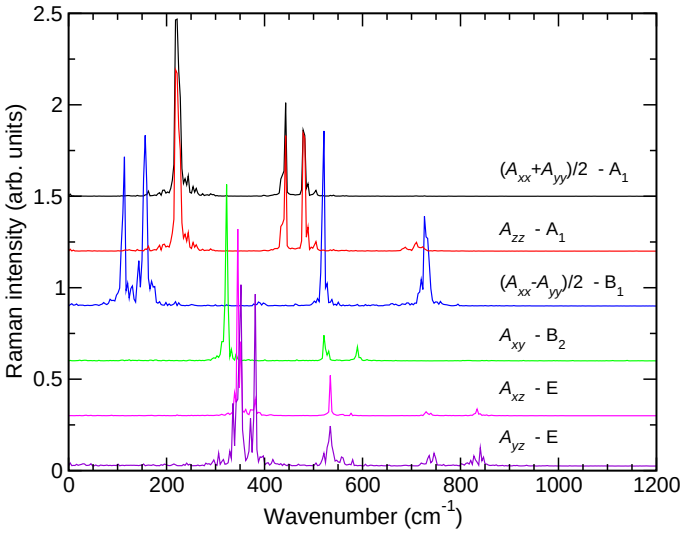

Figure 7.7: Spectra of the polarizability tensor components and of their combinations revealing the Raman active vibrational modes of specific symmetry in the structure of the α-cristobalite.

Figure 7.7 displays the spectra of the polarizability tensor elements corresponding to the modes of each symmetry species. Analysis of the spectra reveals a good agreement with the theoretical predictions. Thus, the spectra of the $(A_{xx} + A_{yy})/2$ combination and of the A_{zz} element have four peaks that corresponds to the expected number of the modes of the A_1 symmetry. It is, however, difficult to count how many modes are exactly present in each spectrum as the Raman activity of the modes is different. It should also be mentioned that the spectra presented in Figure 7.7 were computed from a single trajectory, while averaging over multiple runs is necessary to obtain reliable intensities. These results indicate that the *Morse-SR* potential correctly models the symmetry behaviour of the lattice vibrations.

β-cristobalite.

β-cristobalite is a high-temperature polymorph of silica that transforms into the α-phase upon cooling below ca. 550 K. The structure of the β-phase is stable above ca. 1740 K, but can exist in a metastable state below this temperature. The β-cristobalite lattice has $Fd\bar{3}m$ cubic symmetry and is characterized by a short, ca. 1.55 Å Si-O bonds and the Si-O-Si angles of 180 deg. [291]. The first test of the *Morse-SR* potential consisted in verifying the behaviour of β-cristobalite at the ambient conditions. The NPT MD simulation at $T = 300$ K and at the target pressure of 1 bar has shown a decrease of the volume of the simulation box by ca. 12 % with the corresponding increase of density from 2.17 to 2.44 g cm^3. The density of the final structure approaches the density of the α-cristobalite phase (2.54 g cm^3) obtained in the *Morse-SR* . The resulting structure is characterized by the Si-O-Si angle distribution with the maximum at ca. 146 deg. that is close to the mean Si-O-Si angle in α-cristobalite (Table 7.4).

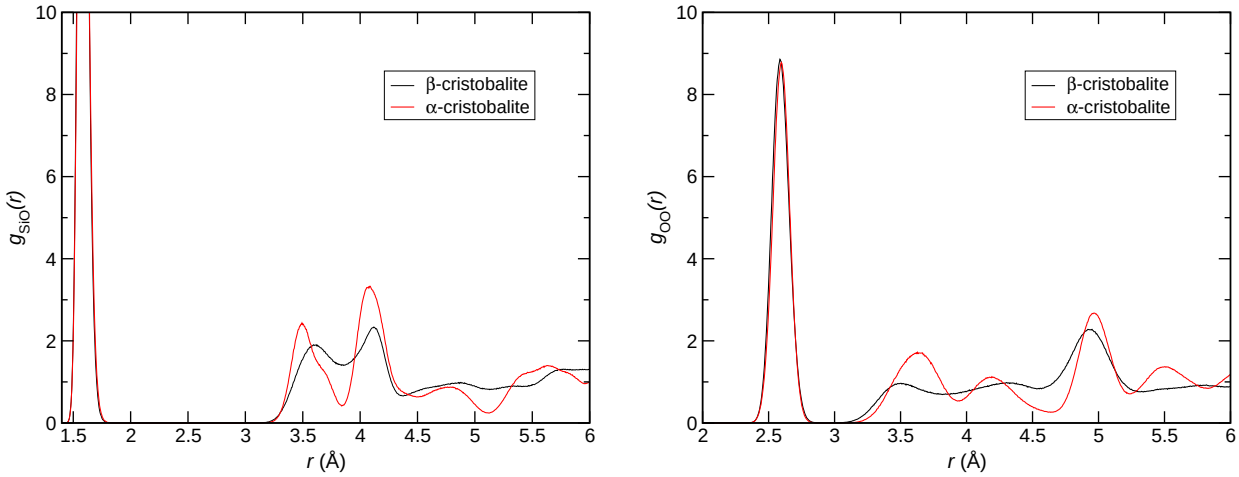

Figure 7.8: Si-O (left) and O-O (right) radial distribution functions for the β-cristobalite system at 300 K in comparison with the α-phase.

Figure 7.8 compares the Si-O and O-O radial distribution functions for the β- and α-cristobalite structures. One sees that the first peaks perfectly coincide and the peaks in the RDFs of β-cristobalite at longer distances well correlate with those in $g(r)$ of α-phase, but have larger widths. This finding points to a strong resemblance of the two structures and to a possible β to α phase transition occurred in the system. The larger width of the peaks in the RDFs of the β-cristobalite system can then be explained by a residual structural disorder. However, the visual comparison of the snapshots of the two systems did not allow us to confirm the transition. It should be noted that modelling phase transitions in a MD simulation can be hampered if the number of atoms in the system, which undergoes the transition, is not compatible with the number of atoms in the final structure or if the parameters of the simulation boxes for the two systems are incommensurate. Results of the simulation of the β-cristobalite system with the *Morse-SR* potential show that the structure is metastable at the ambient conditions.

7.4 Conclusions.

Two polarizable potential models presented in Chapter 6 were tested in the molecular dynamics simulations of crystalline silica polymorphs. Results of these calculations show that the *exp-4* force field fails to reproduce the correct structure of α-quartz. Furthermore, being used for modelling α-cristobalite, the force field produces a spontaneous transition of the structure to a new phase with a high density. These unsatisfactory results exclude the *exp-4* force field from the list of candidates of the polarizable force field for silica.

The *Morse-SR* model performs much better. The structural characteristics of α-quartz and α-cristobalite obtained with the model are in a good agreement with the experimental data and the results of calculations using other effective potentials, such as the BKS ot Tangney-Scandolo ones. The simulation of β-cristobalite at the ambient conditions reveals the metastability of the structure that transforms into a structure with structural characteristics close to those of the α-phase. The major deficiency of the *Morse-SR* force field is its bad performance in the description of the vibrational dynamics of the silica structures. Because of too small value of the Si-O force constant, the high-frequency edge of the silica phonon spectrum is shifted by ca. 300 cm^{-1} to lower wavenumbers. Despite this fact, the symmetry behaviour of the vibrational modes are in a good agreement with the theoretical predictions. A possible reason for the partial success of the *Morse-SR* force field is that the parameters optimization procedure in the force-matching method has ended in a local minimum. Additional studies are necessary for the improvement of the model.

Chapter 8

Conclusions and Perspectives.

The present work was aimed at the development of polarizable force field model for atomistic modelling of oxide materials. From the analysis of the literature and making use of the experience existing the laboratory, it was decided to base the development on a chemical potential equalization (CPE) approach that allows computing the environment- and geometry-dependent charge distribution in the system. These variable charges were expected to be able to subsequently serve a linkage between instantaneous charge distribution and the parameters of effective potential energy functions.

An extensive parametrization of two CPE models, the electronegativity equalization model (EEM) and split-charge equilibration (SQE) model, was performed for silicate materials on the basis of quantum-chemical calculations of oxide clusters containing aluminum, silicon, and zirconium cations. The calibration of the parameters in these models was done using the iterative Hirshfeld (HI) charges and the electrostatic potential (ESP) as the reference quantities. The transferability of parameters was assessed by a comparison with HI charges and ESP grid data for a validation set of isolated molecules and a number of crystalline structures.

The outcome of the CPE parametrization procedure allows us to conclude that the EEM model is capable of mimicking static characteristics only, while it fails to reproduce the response of the electronic distribution and ESP to an external electric field. On the other hand, the SQE model performs well for both static and response properties and also provides correct results for properties that were not explicitly included in the parameterization procedure, i.e. the dipole moment and the dipole polarizability.

Both the EEM and SQE calibrations can be used for a fast and reliable calculation of HI charges in isolated molecules that were not included in the training set, but they reveal a limited transferability to periodic systems. It is however possible to propose a correction to the atomic hardness parameters of each element based on the differences between the EEM/HI and the DFT/HI charges for the periodic systems. This correction amounts to an increase of the atomic hardness in the solid state, which can be related to decrease of atomic polarizability due to the confinement by the crystal field of the solids. These corrections were found to be transferable between the EEM and SQE models. The corrected EEM/HI and SQE/HI schemes reproduce the

reference DFT HI charges in the periodic systems with the mean relative error of less than 2 %.

The ESP-based parametrizations were found to provide reliable results only in those regions of space, where the electronic density values are lower than a threshold value $\rho = 10^{-3}$ a.u. The density criterion was found to be consistent with the distance criterion used in the popular Merz-Kollman ESP charge fitting scheme for a quick selection of grid points. The SQE/ESP model shows a good transferability from molecular to periodic structures, if the regions for computing ESP were chosen according to the criterion above. The EEM/ESP scheme was found to perform the worst among all models studied and hence the use of EEM method for computing the ESP-related characteristics, in particular of electrostatic forces, should be avoided.

This work has laid down a solid background for the description of the long-range part of the polarizable force field model with the CPE approach. On the basis of these results the SQE model may be considered as the best candidate for such a development.

The next step toward the polarizable force field was done by investigating how the dispersion energy in oxides depends on the structure of a particular system. So far, the treatment of the dispersion energy in most of the force field models, with few exceptions cited in Chapter 5, was done in largely empirical way. The dispersion coefficients, which determine the magnitude of the dispersion interactions, were often chosen arbitrarily and their values were sometimes meaningless from the physical viewpoint. In the present work the issue was addressed by computing the interatomic C_6^{ij} dispersion coefficients in crystalline and amorphous structures of silicon and zirconium dioxides with the approach based on the use of maximally-localized Wannier functions (MLWFs) for partitioning electronic density. The localization of MLWFs close to the nuclei in the oxide systems systems makes it possible to assign the functions to the atoms in an unambiguous way, thus allowing the calculation of dispersion coefficients in an atom-wise manner. It is found that the MLWFs can be condensed to effective orbitals centered on atoms and the spread of these effective orbitals can then be used for computing the interatomic dispersion coefficients. The effective orbitals method was shown to yield the results in a very good agreement with those obtained by the original approach.

The dispersion coefficient C_6^{OO} for oxygen atoms was found to have the largest variation. Its value changes not only from one oxide to another, but also for the different modifications of the same compound. On the other hand, the dispersion coefficient for the cations is almost insensitive to structural changes. The dipole polarizability of ions in the systems was computed making use of the Slater-Kirkwood formula and the values of self-atom dispersion coefficient appearing in the effective orbitals model. The oxide ion polarizability in silica structures is in a good agreement with the values obtained in previous calculations and from experimental data [280]. The values of the self-atom dispersion coefficient S_6^{OO} and of the oxide ion polarizability were found to correlate with the nearest-neighbours distance and with the coordination number of oxygen atoms in the structures. This feature is attributed to the effect of confinement by electrostatic potential. Simple functional forms for the dependence of the S_6^{OO} coefficients on the nearest-neighbour distance and

coordination number are suggested. The values of the coefficient and $\alpha_{O^{2-}}$ quantity were related to the oxygen charges computed with the SQE/ESP method. In all studied systems the atoms having larger absolute values of charge were found to have a smaller S_6^{OO} value and to be less polarizable because of stronger confinement of electrons in a more deep potential wall. The results of the work provide a clue on the judicious choice of the dispersion coefficients in the oxide systems and give an insight into the factors affecting the coefficient values.

The final part of the work dealt with the development of effective potential model for the short-range interactions. Silicon dioxide was chosen as a reference system for the development. This part of the work made use of force-matching procedure that was based on the results of DFT calculations of numerous structures of amorphous silica. Assuming the two-body nature of the short-range interactions the "ideal" pairwise interatomic potentials in silica were obtained within the force-matching method using spline potentials. The analysis of the spline-approximated potentials has allowed devising the functional form of different contributions into the short-range potential energy. Furthermore, the SQE/ESP model for the long-range part of the force field was chosen with the use of merit criteria. Two short-range potential models were constructed in this way. Being combined with the long-range electrostatic energy and with the dispersion energy term, the complete model was tested in the molecular dynamics simulations of the structural and dynamical characteristics of crystalline silica polymorphs with four-fold coordinated silica atoms.

The first model, *exp-4* force field, was found to fail reflecting the correct structure of α-quartz. Furthermore, being used for modelling of α-cristobalite, the force field produced a spontaneous transition of the structure to a new phase with a high density, above the densities of quartz and cristobalite structures. These unsatisfactory results exclude the *exp-4* force field from the list of candidates of the polarizable force field for silica.

The second model, in which the short-range interaction were represented via a Morse potential, yileds more satisfactory results. Thus, the structural characteristics of α-quartz and α-cristobalite obtained with the model are in a good agreement with the experimental data and with the results of calculations using other force fields. The simulation of β-cristobalite at the ambient conditions reveals the metastability of the structure. The main drawback of the *Morse-SR* model is its unsatisfactory performance in the description of the vibrational dynamics of the silica structures. Because of too small value of the Si-O force constant, the high-frequency edge of the silica phonon spectrum is shifted to lower wavenumbers. Consequently, one can also expect that the model will yield too small values of the bulk moduli and elastic constants. Despite this fact, the symmetry behaviour of the vibrational modes are in a good agreement with the theoretical predictions.

The main results of the work presented above allow us to conclude that this study has laid a solid background for the development polarizable force field model for oxides. Future studies making use these results should certainly lead to a more reliable and transferable models and some of such works are already on the way.

The experience gained while performing the research presented in the dissertation allows us to envisage the following perspectives for the development of polarizable force field. The perspectives are related to both the approach, which was used to deal with the long-range interactions, and the models of the short-range interactions.

The first aspect concerning the model of the long-range interactions is the extending the SQE model to the *sp*-basis, in a line with the results of Refs. [162, 163]. Indeed, the SQE approach used in the present work makes use of the spherical distribution of charge density around the atoms. Adding *p*-like functions is expected to yield a better description of the charge distribution and its variation upon interaction with the surrounding. The use of distributed dipoles mimicked by the *p*-like functions may greatly improve the description of the electrostatic interactions and of the response of the charge distribution to a perturbation by external potential. For example, water molecule has zero out-of-plane polarizability component within the *s*-model, whereas the addition of the *p*-like dipole components allows to recover the correct values of the components of the polarizability tensor. In a certain sense, the extension is similar to adding the polarization functions in the quantum-chemical basis sets, that is known to greatly improve the quality of the results.

The second perspective pertains the fact that the present SQE model does not allow to take the bond breaking and bond formation phenomena into account. The need of such an improvement is vital for the investigation of such phenomena as structural phase transitions, which can involve bond breaking/formation mechanisms. Furthermore, the modeling of glass formation is not possible within the present approach as the charge transfer channels between the atoms remain the same regadless the atoms environment. A possible way of designing such a reactive SQE-based force field was presented in Ref. [206] and the way makes use of the bond-order concept to decide between which atoms the charge transfer is allowed for.

The envisaged improvements of model of the short-range interactions are also many-fold. An obvious improvement concerns the dispersion energy term. In the present work the dispersion energy rapidly grows as the interatomic distance decreases, whereas in the reality the interaction is screened by electron distributions at short distances. The screening is commonly introduced by a damping function, e.g. in the form suggested by Tang and Toennies [271]. In the frame of the present work it is worth studying whether the parameter of the damping function can be related to the characteristics of MLWFs. Furthermore, Chapter 5 outlined the way of making the dispersion coefficients environment-dependent. This model was not explored in the present work for the sake of simplicity and should certainly be tested in future studies.

It is also clear that the choice of the functional form for the short-range interaction potentials plays very important role. Indeed, despite the fact that both the *exp-4* and *Morse-SR* models equally well reproduce the *ab initio* forces and have very similar values of the merit criteria, the former force field fails to mimic the characteristics of crystalline silica polymorphs, whereas the later performs reasonably well. It is expected that the extension of the SQE approach to

the *sp* model will also yield the improvement of the short-range interaction potentials. Indeed, the charge-charge, charge-dipole, and dipole-dipole interactions explicitly treated in the *sp* SQE model are the first terms in the inverse power series, which, as it was shown in Chapter 6, well reproduces effective interaction potentials at the medium-range distances. The force-matching method using splines can be of significant help for choosing the functional form of the short-range interaction potentials and to judge whether three-body (and higher) terms need to be included into the model.

It is necessary to admit that the goal, which was claimed at the time when the work has been started, is not achieved. As it was mentioned above the geometry-dependent atomic charges could serve a linkage between the instantaneous charge distribution in the system and the environment-dependent parameters. However, the force fields developed and tested in the work do not contain the parameters of the short-range interaction potentials that depend on the environment of atoms. Basically, the dependence of the parameters of the existing potentials on atomic coordinates can be introduced using the atomic charges. Such a modification of the potential will, however, lead to some extra effort in the force and stress calculations (for detail see Appendix F). Introducing the charge-dependent parameters of short-range potentials can make the force field more flexible and transferable. However, a careful estimation of the computational expenses needs to be carried out prior to introducing the extension.

In summary, it is possible to say that the perspectives presented above show how much room does still exist in the development of polarizable force field models.

Appendix A

Electrostatic interactions for distributed charges and dipoles.

The orthonormalized Gaussian form factors $s_i(\mathbf{r})$ and $\mathbf{p}_i(\mathbf{r})$ in (3.2.41) are defined as

$$s_i(\mathbf{r}) = \left(\frac{\xi_i}{\pi}\right)^{3/2} \exp(-\xi_i|\mathbf{r}-\mathbf{r}_i|^2), \tag{A.0.1}$$

$$p_{\beta i}(\mathbf{r}) = 2\pi\left(\frac{\zeta_i}{\pi}\right)^{5/2} (r_\beta - r_{\beta i})\exp(-\zeta_i|\mathbf{r}-\mathbf{r}_i|^2), \tag{A.0.2}$$

with $\mathbf{r}_i$ and $r_{\beta i}$ being radius vector of an atom i and its $\beta-$th component, $\beta = x, y, z$. ξ and ζ are the parameters related to the spread of charge distribution. It is noteworthy that since $s_i(\mathbf{r})$ and $\mathbf{p}_i(\mathbf{r})$ are othonormalized Q_i and $\mathbf{d}_i$ are atomic charge and dipole moment respectively.

The ss, $J(\mathbf{r}_{ij}, \xi_i, \xi_j)$, sp, $F_\beta(\mathbf{r}_{ij}, \xi_i, \zeta_j)$, and pp, $P_{\beta\gamma}(\mathbf{r}_{ij}, \zeta_i, \zeta_j)$, interactions are obtained from the following expressions

$$J(\mathbf{r}_{ij}, \xi_i, \xi_j) = \int\int \frac{s_i(\mathbf{r})s_j(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}d\mathbf{r}' = \frac{\mathrm{erf}(\varepsilon_{ij}r_{ij})}{r_{ij}}, \tag{A.0.3}$$

where $\mathrm{erf}(x)$ is the error function and

$$\varepsilon_{ij} = \left(\frac{\xi_i\xi_j}{\xi_i+\xi_j}\right)^{1/2}. \tag{A.0.4}$$

$$F_\beta(\mathbf{r}_{ij}, \xi_i, \zeta_j) = \int\int \frac{s_i(\mathbf{r})p_{\beta j}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}d\mathbf{r}' = \frac{r_{\beta ij}}{r_{ij}^2}\left[\frac{\mathrm{erf}(z_{ij}r_{ij})}{r_{ij}} - \frac{2z_{ij}}{\sqrt{\pi}}\exp(-z_{ij}^2r_{ij}^2)\right], \tag{A.0.5}$$

where

$$z_{ij} = \left(\frac{\xi_i\zeta_j}{\xi_i+\zeta_j}\right)^{1/2}. \tag{A.0.6}$$

$$P_{\beta\gamma}(\mathbf{r}_{ij},\zeta_i,\zeta_j) = \int\int \frac{p_{\beta i}(\mathbf{r})p_{\gamma j}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}d\mathbf{r}' = \frac{\delta_{\beta\gamma}}{r_{ij}^2}\left[\frac{\mathrm{erf}(\kappa_{ij}r_{ij})}{r_{ij}} - \frac{2\kappa_{ij}}{\sqrt{\pi}}\exp(-\kappa_{ji}^2 r_{ij}^2)\right]$$
$$- 3\frac{r_{\beta ij}r_{\gamma ij}}{r_{ij}^4}\left[\frac{\mathrm{erf}(\kappa_{ij}r_{ij})}{r_{ij}} - \frac{2\kappa_{ij}}{\sqrt{\pi}}\exp(-\kappa_{ji}^2 r_{ij}^2)(1+\frac{2}{3}\kappa_{ij}^2 r_{ij}^2)\right], \quad \text{(A.0.7)}$$

where

$$\kappa_{ij} = \left(\frac{\zeta_i\zeta_j}{\zeta_i+\zeta_j}\right)^{1/2}. \quad \text{(A.0.8)}$$

One can obtain the atomic hardness for s and p distributions. For s distribution one should take $i=j$ in the eq. (A.0.3), which leads to the limit case $r_{ij}\to 0$. For p distribution one should take the same limit $r_{ij}\to 0$ and also apply two conditions: $i=j$ and $\beta=\gamma$.

$$\eta_s(\xi_i) = \left(\frac{2\xi_i}{\pi}\right)^{1/2}, \quad \text{(A.0.9)}$$

$$\eta_p(\zeta_i) = \left(\frac{2\zeta_i}{9\pi}\right)^{1/2}\zeta_i. \quad \text{(A.0.10)}$$

Appendix B

Hardness matrix elements of the CPE and SQE models for periodic systems.

For periodic systems the electrostatic interactions are generally treated with the use of the Ewald sum method [318]. A combination of the method with a screened Coulombic potential, see for example eq. (4.2.7), results in the following expression for the electrostatic interaction energy E_{el} [319]

$$\begin{aligned} E_{el} &= \frac{1}{2}\sum_{i=1}^{N}\sum_{j\neq i}^{N} q_i q_j \Big[\, J(r_{ij}) - \frac{\mathrm{erf}(\beta r_{ij})}{r_{ij}} \,\Big] \\ &+ \frac{2\pi}{V}\sum_{k\neq 0}^{k_{max}} \frac{e^{-k^2/4\beta^2}}{k^2} \sum_{i=1}^{N}\sum_{j=1}^{N} q_i q_j\, e^{\imath \mathbf{k}\cdot(\mathbf{r}_i-\mathbf{r}_j)} - \frac{\beta}{\sqrt{\pi}}\sum_{i=1}^{N} q_i^2, \end{aligned} \tag{B.0.1}$$

where $J(r)$ is the screened electrostatic potential, and β and k_{max} are the parameters controlling the convergence of the Ewald sums in the direct and reciprocal space, respectively. For periodic systems the equation (B.0.1) replaces the last term in (4.2.3) and (4.2.6).

Making use of (4.2.3), (4.2.6), and (B.0.1) the charge-dependent energy of periodic system in the EEM and SQE models can be rewritten as

$$E_{\mathrm{EEM}} = \sum_i \big[\chi_i^* q_i + \frac{1}{2}\tilde{\eta}_i q_i^2\big] + \frac{1}{2}\sum_i\sum_{j\neq i} q_i q_j\, J'(r_{ji}), \tag{B.0.2}$$

$$\begin{aligned} E_{SQE} &= \sum_i\sum_j \big[\xi_{ji}^* p_{ji} + \frac{1}{2}\tilde{\kappa}_{ji} p_{ji}^2\big] + \frac{1}{2}\sum_i \tilde{\eta}_i \sum_m\sum_{l\neq m} p_{mi} p_{li} \\ &+ \frac{1}{2}\sum_i\sum_{j\neq i} J'(r_{ji})\big(\sum_m p_{mi}\big)\big(\sum_l p_{lj}\big), \end{aligned} \tag{B.0.3}$$

where the quantities $\tilde{\eta}_i$, $\tilde{\kappa}_{ij}$ and $J'(r_{ji})$ stand for

$$\tilde{\eta}_i = \eta_i^* + \frac{4\pi}{V}\sum_{k\neq 0}^{k_{max}} \frac{e^{-k^2/4\beta^2}}{k^2} - \frac{2\beta}{\sqrt{\pi}}, \tag{B.0.4}$$

$$\tilde{\kappa}_{ij} = \kappa_{ij}^* + \tilde{\eta}_j, \tag{B.0.5}$$

$$J'(r_{ji}) = J(r_{ij}) - \frac{\mathrm{erf}\,(\beta r_{ij})}{r_{ij}} + \frac{4\pi}{V}\sum_{k\neq 0}^{k_{max}} \frac{e^{-k^2/4\beta^2}}{k^2} e^{\imath\mathbf{k}\cdot(\mathbf{r}_i-\mathbf{r}_j)}. \tag{B.0.6}$$

As before, a generic formula for hardness matrix element $H_{(ji)(lm)}$ in SQE model can be obtained by the differentiation of eqn. (B.0.3) with respect to the CTPs q_{ji} and q_{lm}

$$\begin{aligned} H_{(ji)(lm)} = {} & \delta_{jl}\delta_{im}\big(\tilde{\kappa}_{ji} + \tilde{\kappa}_{ij} - 2J'_{ji}\big) + \big[\delta_{im}(1-\delta_{jl})\tilde{\eta}_i - \delta_{il}\tilde{\eta}_i - \delta_{jm}\tilde{\eta}_j + \delta_{lj}(1-\delta_{im})\tilde{\eta}_j\big] + \\ & \big(1-\delta_{jl}\delta_{im}\big)\big[\big(1-\delta_{lj}\big)J'_{lj} - \big(1-\delta_{li}\big)J'_{li} - \big(1-\delta_{mj}\big)J'_{mj} + \big(1-\delta_{mi}\big)J'_{mi}\big], \end{aligned} \tag{B.0.7}$$

where $J'_{ji} \equiv J'(r_{ji})$ and δ_{ji} is the Kronecker symbol.

Appendix C

Static and response cost functions.

Cost functions.

External perturbation. In addition to the static quantities (A), such as HI charges and electrostatic potential, their response to an external perturbation V_{ext} was computed and the resulting response characteristics were then employed in constructing the corresponding response cost function. The external perturbation was given by

$$V_{\text{ext}} = \mathcal{E}_x x + \mathcal{E}_y y + \mathcal{E}_z z \tag{C.0.1}$$

where $\mathcal{E}_\alpha$ ($\alpha = x, y, z$) is an applied uniform electric field in x, y and z directions. The derivative of quantity A with respect to the perturbation $\mathcal{E}_\alpha$ was obtained by the finite differences as

$$A'_\alpha \approx \frac{A(\mathcal{E}_\alpha = +\epsilon) - A(\mathcal{E}_\alpha = -\epsilon)}{2\epsilon}. \tag{C.0.2}$$

A value of $\epsilon = 0.0019$ a.u. is used, which is equal to the default value in Gaussian03 code when a numerical differentiation is carried out with respect to external electric field.

Static HI cost function. The static HI cost function uses the iterative Hirshfeld charges and it has the following form

$$X_{\text{S,HI}} = \frac{\sum_{m=1}^{M}\sum_{n=1}^{N_m} w_{mn}\left[q_{mn}^{0} - q_{mn}^{\text{MODEL}}\right]^2}{\sum_{m=1}^{M}\sum_{n=1}^{N_m} w_{mn}}, \tag{C.0.3}$$

where the sums in the numerator run over M molecules and N_m atoms in a given molecule, respectively; the superscript "0" denotes the reference value and MODEL stands for value obtained with either EEM or SQE model. The weight w_{mn} associated with a certain atomic charge is inverse proportional to the prevalence of the corresponding chemical element in the training set

(Table 4.1). This choice compensates for the fact that some chemical elements are more abundant than others. The cost function (C.0.3) can be seen as a weighted average of the individual squared errors.

Response HI cost function. The response cost function related to the HI charges takes the following form:

$$X_{\mathrm{R,HI}} = \frac{\sum_{m=1}^{M} \sum_{\alpha=x,y,z} \sum_{n=1}^{N_m} w_{mn} \left[\left(q_{mn}^{0}\right)'_{\alpha} - \left(q_{mn}^{\mathrm{MODEL}}\right)'_{\alpha} \right]^2}{\sum_{m=1}^{M} \sum_{n=1}^{N_m} 3 w_{mn}} \tag{C.0.4}$$

The derivatives of the model HI charges towards the external field are computed analytically. The weights are identical to those of the static HI cost function (C.0.3).

Static ESP cost function. For every molecule m, a set G_m of grid points $\mathbf{r}_{m,g}$ ($g \in G_m$) with their associated weights $w_{m,g}$ was defined. At these grid points the DFT and model electrostatic potentials are computed and used to construct static ESP cost function as follows:

$$X_{\mathrm{S,ESP}} = \frac{1}{M} \sum_{m=1}^{M} \left(\frac{\sum_{g=1}^{G_m} w_{m,g} \left[V_m^0(\mathbf{r}_{m,g}) - V_m^{\mathrm{MODEL}}(\mathbf{r}_{m,g}) \right]^2}{\sum_{g=1}^{G_m} w_{m,g}} \right). \tag{C.0.5}$$

The grid consists of a uniform 3D grid whose edges are separated by at least 5 Åfrom the nearest atom. The spacing between the grid points is 0.2 Å. The part between brackets in (C.0.5) represents an ESP cost function for a molecule m. The evaluation of this molecular cost function is made more efficient by reducing it to a simple quadratic function of charges prior to the calibration procedure.

The weight function $w_m(\mathbf{r}_{m,g})$ is smooth, becomes one in a shell surrounding the molecule and decays to zero inside the molecule and at large distances, i.e. at the edges of the grid. This weight function is constructed as the product of two switching functions:

$$w(\mathbf{r}) = \mathrm{switch}\left[-\alpha(\log_{10}(\rho(\mathbf{r})) - \log_{10}(\rho_0))\right] \times \mathrm{switch}\left[-\beta(\mathrm{d}(\mathbf{r}) - d_0)\right] \tag{C.0.6}$$

where $\mathrm{switch}(x)$ and $\mathrm{d}(\mathbf{r})$ are defined as

$$\mathrm{switch}(x) = \begin{cases} 0 & \text{if } x < -1 \\ (1 + \sin(\pi x/2))/2 & \text{if } -1 < x < 1 \\ 1 & \text{if } x > 1 \end{cases} \tag{C.0.7}$$

$$\mathrm{d}(\mathbf{r}) = -a_0 \log\left[\sum_{j=1}^{N} \exp(-\|\mathbf{R}_j - \mathbf{r}\|/a_0)\right]. \tag{C.0.8}$$

The first factor in the weight function switches to zero inside the molecule. The second factor switches of the weight function at long distances. The function $d(\mathbf{r})$ is a smooth function that approximates the distance to the closest atom. The parameters are fixed to $\alpha = 1.0$, $\rho_0 = 3\text{e-}4$ a.u., $\beta = 0.5$ Å, $d_0 = 4.0$ Å, and a_0 is 1 Bohr.

The smoothness of the weight function guarantees that a small rotation of the molecules with respect to the grid only leads to small changes in the cost function. A second advantage of this approach is that it never includes grid points in the fitting procedure where the electron density is significantly non-zero.

Response ESP cost function. The response ESP cost function takes the following form:

$$X_{\mathrm{R,ESP}} = \frac{1}{3M} \sum_{m=1}^{M} \sum_{\alpha=x,y,z} \left(\frac{\sum_{g=1}^{G_m} w_{m,g} \left[\left(V_m^0\right)'_\alpha(\mathbf{r}_{m,g}) - \left(V_m^{\mathrm{MODEL}}\right)'_\alpha(\mathbf{r}_{m,g}) \right]^2}{\sum_{g=1}^{G_m} w_{m,g}} \right). \tag{C.0.9}$$

The derivatives of the model ESP towards the external field was computed numerically. The weights were identical to those in the static ESP cost function (C.0.5). Similar to the static ESP cost function, the part between brackets was first reduced to a quadratic function of charges for each molecule to lower the computational cost of the calibration procedure.

Appendix D

Details on structural models.

α-quartz.

The supercell was constructed from 1×2×2 unit cells of the α-quartz structure in orthorhombic setting (2×2×2 conventional trigonal unit cells). The lattice parameters and fractional coordinates of atoms were taken from [320].

α-cristobalite.

The lattice parameters and the fractional coordinates of atoms in the tetragonal unit cell of α-cristobalite were taken from [321]. The supercell was built of 2×2×2 unit cells (96 atoms).

Amorphous silica.

The structure of a-SiO_2 was obtained in a classical molecular dynamics simulation of a system constructed from 2×2×2 unit cells of β-cristobalite (192 atoms). The interatomic interactions were described with the BKS potential [52]. The system was first melted at $T = 7000$ K for the period of 1 ns and then cooled down to the temperature of 3000 K during 4 ns. The simulations were carried out in the NPT statistical ensemble allowing the isotropic variation of the cubic simulation cell. Since the NPT MD simulations using the BKS potential produce too dense a-SiO_2 structures, the cell parameters were scaled to yield the density corresponding to the experimental density of 2.2 g/cm^3. The resulting structure was relaxed in a NVT MD run at $T = 3000$ K during 2 ns period and the final configuration of the run was used as the "ini" structure.

Tetragonal ZrO_2.

The supercell used in the calculations was built of 3×3×2 unit cells of t-ZrO_2 structure with the lattice parameters and atomic coordinates taken from [322].

Cubic ZrO_2.

The supercell containing 96 atoms was built of 2×2×2 unit cells of c-ZrO_2 structure with the lattice parameters and atomic coordinates taken from [323].

Amorphous zirconia.

a-ZrO_2 model was obtained in classical MD simulations starting from the structure of amorphous silica. The Si atoms in the a-SiO_2 model were substituted with Zr atoms and the simulation cell parameters of the silica model were scaled to produce the density of 5.32 g/cm^3 [324, 325]. The interatomic interactions were described with a BKS-like potential [326]. The model structure was used in a NPT MD simulation in which the temperature was first raised from 300 K to 5000 K during the period of 4 ns and then lowered to 2800 K during 2 ns. The resulting structure was then relaxed in a NVT simulation at $T = 2800$ K during 500 ps and the final configuration was used in CPMD calculations. The model a-ZrO_2 structure obtained by the above procedure had the density of 4.8 g/cm^3 and in the text this structural model is denoted as low-density (LD) a-ZrO_2. A more dense a-ZrO_2 structure named the high-density (HD) model was obtained by scaling the cell parameters of the LD model to obtain the target density of 5.32 g/cm^3. Both the LD and HD structures were then used as a starting configuration in CPMD runs.

Appendix E

Forces and Hessian matrix in CPE models.

General expression for the electrostatic energy E_q in CPE models (EEM, SQE, etc.) is

$$E_q = \mathbf{x}^{\mathrm{T}}\mathbf{q} + \frac{1}{2}\mathbf{q}^{\mathrm{T}}\mathbf{H}\mathbf{q}. \tag{E.0.1}$$

The ζ component of the electrostatic force F_i^{ζ} acting on atom i is obtained as

$$F_i^{\zeta} = \frac{\partial E_q}{\partial \zeta_i} = \frac{\partial}{\partial \zeta_i}\left[\mathbf{x}^{\mathrm{T}}\mathbf{q} + \frac{1}{2}\mathbf{q}^{\mathrm{T}}\mathbf{H}\mathbf{q}\right] \tag{E.0.2}$$

and it reads

$$F_i^{\zeta} = \left[\mathbf{x}^{\mathrm{T}} + \mathbf{q}^{\mathrm{T}}\mathbf{H}\right]\frac{\partial \mathbf{q}}{\partial \zeta_i} + \frac{1}{2}\mathbf{q}^{\mathrm{T}}\frac{\partial \mathbf{H}}{\partial \zeta_i}\,\mathbf{q}. \tag{E.0.3}$$

One can note that $\mathbf{x}^{\mathrm{T}} + \mathbf{q}^{\mathrm{T}}\mathbf{H}$ in the square brackets is nothing as $\partial E_q/\partial \mathbf{q}$ and since our charges minimize the energy (E.0.1), the relation $\partial E_q/\partial \mathbf{q} = 0$ holds. Note that in the case of EEM model $\partial E_q/\partial \mathbf{q} = \tilde{\mathbf{x}}^{\mathrm{T}}$, where $\tilde{\mathbf{x}}$ is the global electronegativity of the molecule, and therefore, the first term in (E.0.3) is simply

$$\tilde{\mathrm{x}}\,\frac{\partial}{\partial \zeta_i}\sum_j q_j = 0 \tag{E.0.4}$$

because of the charge conservation condition. Hence,

$$F_i^{\zeta} = \frac{1}{2}\mathbf{q}^{\mathrm{T}}\,\frac{\partial \mathbf{H}}{\partial \zeta_i}\,\mathbf{q} \tag{E.0.5}$$

i.e. the atomic force in a variable-charge CPE model is identical to that of a fixed-charge model.

Let us look at the Hessian matrix elements

$$\frac{\partial^2 E_q}{\partial \alpha_i \partial \beta_j} = \frac{\partial}{\partial \beta_j}\left(\frac{1}{2}\,\mathbf{q}^{\mathrm{T}}\,\frac{\partial \mathbf{H}}{\partial \alpha_i}\,\mathbf{q}\right) = \mathbf{q}^{\mathrm{T}}\,\frac{\partial \mathbf{H}}{\partial \alpha_i}\,\frac{\partial \mathbf{q}}{\partial \beta_j} + \frac{1}{2}\,\mathbf{q}^{\mathrm{T}}\,\frac{\partial^2 \mathbf{H}}{\partial \alpha_i \partial \beta_j}\,\mathbf{q}. \tag{E.0.6}$$

In the framework of CPE models the atomic charges are obtained as

$$\mathbf{q} = -\mathbf{H}^{-1}\mathbf{x} \tag{E.0.7}$$

and their derivatives with respect to the β coordinate of atom j read

$$\frac{\partial \mathbf{q}}{\partial \beta_j} = \mathbf{H}^{-1}\frac{\partial \mathbf{H}}{\partial \beta_j}\mathbf{H}^{-1}\mathbf{x} = -\mathbf{H}^{-1}\frac{\partial \mathbf{H}}{\partial \beta_j}\mathbf{q}. \tag{E.0.8}$$

Inserting this expression to the first term in the right-hand part of (E.0.6) gives

$$\mathbf{q}^{\mathrm{T}}\frac{\partial \mathbf{H}}{\partial \alpha_i}\frac{\partial \mathbf{q}}{\partial \beta_j} = -\mathbf{q}^{\mathrm{T}}\frac{\partial \mathbf{H}}{\partial \alpha_i}\mathbf{H}^{-1}\frac{\partial \mathbf{H}}{\partial \beta_j}\mathbf{q}. \tag{E.0.9}$$

Finally, the expression for the Hessian matrix element is obtained in the following form

$$\frac{\partial^2 E_q}{\partial \alpha_i \partial \beta_j} = \mathbf{q}^{\mathrm{T}}\left[\frac{1}{2}\frac{\partial^2 \mathbf{H}}{\partial \alpha_i \partial \beta_j} - \frac{\partial \mathbf{H}}{\partial \alpha_i}\mathbf{H}^{-1}\frac{\partial \mathbf{H}}{\partial \beta_j}\right]\mathbf{q}. \tag{E.0.10}$$

Appendix F

Short-range forces and virial for the environment-dependent ansatz.

While the CPE model does not modify the expression for the long-range electrostatic forces, the short-range force $\mathbf{F}_i^{\mathrm{SR}}$ acting on atom i is given in the charge-dependent ansatz by

$$\mathbf{F}_i^{\mathrm{SR}} = -\boldsymbol{\nabla}_i E_{\mathrm{SR}}(\{\mathbf{r}_i\}) = \mathbf{F}_i^0 + \mathbf{F}_i^q, \tag{F.0.1}$$

where $\mathbf{F}_i^0$ is the conventional force, which can be obtained from (6.5.11) assuming the constant charges, and the $\mathbf{F}_i^q$ quantity is the force that results from the fact, that the charges q_i depend explicitly on the atomic positions $\{\mathbf{r}_i\}$. This additional force can be expressed as

$$\mathbf{F}_i^q = -\frac{\partial E_{\mathrm{SR}}(\{\mathbf{r}_i\}, \mathbf{q})}{\partial \mathbf{q}} \frac{\partial \mathbf{q}}{\partial \mathbf{r}_i}, \tag{F.0.2}$$

with $\mathbf{q}$ being the vector of charges. One sees that the force (F.0.2) has a many-body nature because each charge depends on the positions of all N atoms in the system.

The following analysis uses the fact that the charges are computed via the SQE formalism. In fact, the derivative of charge q_i with respect to the coordinate β of atom k can easily be calculated analytically as (cf. (4.2.2))

$$\frac{\partial q_i}{\partial \beta_k} = \sum_j \frac{\partial p_{ij}}{\partial \beta_k} = -\sum_j \frac{\partial}{\partial \beta_k} \left[\mathbf{H}^{-1}_{(ij)(lm)}\, \mathbf{x}_{(lm)} \right] = -\sum_j \frac{\partial \mathbf{H}^{-1}_{(ij)(lm)}}{\partial \beta_k}\, \mathbf{x}_{(lm)}, \tag{F.0.3}$$

where $\mathbf{H}^{-1}_{(ij)(lm)}$ is the element of the inverse of the hardness matrix, which corresponds to interaction between the p_{ij} and p_{lm} charge-transfer parameters, and $\mathbf{x}_{(lm)}$ is the element of the bond electronegativy vector corresponding to the parameter p_{lm}. The derivative of the inverse hardness matrix with respect to β_k reads

$$\frac{\partial\, \mathbf{H}^{-1}}{\partial \beta_k} = -\mathbf{H}^{-1} \frac{\partial \mathbf{H}}{\partial \beta_k} \mathbf{H}^{-1} \tag{F.0.4}$$

and therefore, the derivative of the vector of the charge-transfer parameters with respect to the coordinate β_k is given by

$$\frac{\partial \mathbf{q}}{\partial \beta_k} = \mathbf{H}^{-1} \frac{\partial \mathbf{H}}{\partial \beta_k} \mathbf{H}^{-1} \mathbf{x} = \mathbf{H}^{-1} \frac{\partial \mathbf{H}}{\partial \beta_k} \mathbf{p}. \tag{F.0.5}$$

And the derivatives of the Hardness matrix $\partial \mathbf{H}/\partial \beta_k$ for SQE model, while applied to periodic structures, can be obtained by using the eq. (B.0.7). Eq. (F.0.5) shows that the calculation of the derivatives $\partial \mathbf{q}/\partial \beta_k$ necessitates $3N$ computations of the $\partial \mathbf{H}/\partial \beta_k$ matrix, that can be a very time consuming operation.

An additional complication in the charge-dependent ansatz comes from the fact that the short-range contribution to stress tensor cannot now be computed by equation (6.5.12). The reason is that the $\mathbf{F}_i^q$ force (F.0.2) has become a many-body quantity. For a potential, which is not pairwise, the virial should be calculated, according to Louwerse and Baerens [327], with the following expression

$$W(\{\mathbf{r}\}) = \sum_{i=1}^{N} \mathbf{r}_i \cdot \mathbf{F}_i - 3V \left(\frac{\partial E}{\partial V} \right)_{\{\mathbf{r}\}}. \tag{F.0.6}$$

Bibliography

[1] Born, M.; Oppenheimer, R. *Ann. Physik* 84, 457 (1927).

[2] Fock, V. F. *Z. Phys.* 61, 126 (1930).

[3] Hohenberg, P.; Kohn, W. *Phys. Rev. B* 136, 864 (1964).

[4] Kohn, W.; Sham, L. J. *Phys. Rev.* 140, A1133 (1965).

[5] Dirac, P. A. M. *Proc. Cambridge Phil. Soc.* 26, 376 (1930).

[6] Gell-Mann, M.; Brueckner, K. A. *Phys. Rev.* 106, 364 (1957).

[7] Carr, W. J.; Maradudin, A. A. *Phys. Rev.* 133, A371 (1964).

[8] Carr, W. J. *Phys. Rev.* 122, 1437 (1961).

[9] Ceperley, D. M.; Alder, B. J. *Phys. Rev. Lett.* 45, 566 (1980).

[10] Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* 58, 1200 (1980).

[11] Perdew, J. P.; Zunger, A. *Phys. Rev. B* 23, 5048 (1981).

[12] Perdew, J. P.; Wang, Y. *Phys. Rev. B* 45, 13244 (1992).

[13] Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* 77, 3865 (1996).

[14] Zhang, Y.; Yang. W. *Phys. Rev. Lett.* 80, 890 (1998).

[15] Becke, A. D. *J. Chem. Phys.* 98, 1372 (1992).

[16] Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *J. Phys. Chem.* 98, 11623 (1994).

[17] Becke, A. D. *Phys. Rev. B* 38, 3098 (1988).

[18] Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* 37, 785 (1993).

[19] Pople, J. A.; Head-Gordon, M.; Fox, D. J.; Raghavachari, K.; Curtiss, L. A.; *J. Chem. Phys.* 90, 5622 (1989).

[20] Bouckaert, L. P.; Smoluchowski, R.; Wigner, E. *Phys. Rev.* 50, 58 (1936).

[21] Baldereschi, A. *Phys. Rev. B* 7, 5212 (1973).

[22] Chadi, D. J.; Cohen, M. L. *Phys. Rev. B* 8, 5747 (1973).

[23] Monkhorst, H. J.; Pack, J. D. *Phys. Rev. B* 13, 5188 (1976).

[24] Chadi, D. J. *Phys. Rev. B* 16, 1746 (1977).

[25] Evarestov, R. A.; Smirnov, V. P. *Phys. Stat. Sol. b* 119, 9 (1983).

[26] Froyen, S. *Phys. Rev. B* 39, 3168 (1989).

[27] Slater, J. C. *Phys. Rev.* 36, 57 (1930).

[28] Boys, S. F. *Proc. R. Soc. Lond. A* 200, 542 (1950).

[29] Obara, S.; Saika, A. *J. Chem. Phys.* 84, 3963 (1986).

[30] Pulay, P. *Mol. Phys.* 17, 197 (1969).

[31] Hamann, D. R.; Schlüter, M.; Chiang, C. *Phys. Rev. Lett.* 43, 1494 (1979).

[32] Bachelet, G. B.; Hamann, D. R.; Schlüter. *Phys. Rev. B* 26, 4199 (1982).

[33] Troullier, N.; Martins, J. S. *Phys. Rev. B* 43, 1993 (1991).

[34] Rappé, A. M.; Rabe, K. M.; Kaxiras, E.; Joannopoulos, J. D. *Phys. Rev. B* 41, 1993 (1990).

[35] Vanderbilt, D. *Phys. Rev. B* 41, 7892 (1990).

[36] Alder, B. J.; Wainwright, T. E. *J. Chem. Phys.* 27, 1208 (1957).

[37] Alder, B. J.; Wainwright, T. E. *J. Chem. Phys.* 31, 459 (1959).

[38] Rahman, A. *Phys. Rev.* 136, A405 (1964).

[39] Nosé, S. *Mol. Phys.* 52, 255 (1984).

[40] Nosé, S. *J. Chem. Phys.* 81, 511 (1984).

[41] Hoover, W. G. *Phys. Rev. A* 31, 1695 (1985).

[42] Car, R.; Parinello, M. *Phys. Rev. Lett.* 55, 2471 (1985).

[43] Laasonen, K.; Pasquarello, A.; Car, R.; Lee, C.; Vanderbilt, D. *Phys. Rev. B* 47, 10142 (1993).

[44] McQuarrie, D.A. *Statistical Mechanics*; Harper & Row, 1976.

[45] Gordon, R. G. *Adv. Mag. Resonance* 3, 1 (1968).

[46] Gordon, R. G. *J. Chem. Phys.* 42, 3658 (1965)

[47] Berens, P. H.; Wilson, K. R. *J. Chem. Phys.* 74, 4872 (1981).

[48] Berens, P. H.; White, S. R.; Wilson, K. R. *J. Chem. Phys.* 75, 515 (1981).

[49] Press, W. H.; Flannery, B. P.; Teukolsky, S. A.; Vetterling, W. T. Conjugate Gradient Methods in Multidimensions. In *Numerical Recipes in C: The Art of Scientific Computing*; Cambridge University Press: New York, **1992**; p 420.

[50] Born, M.; Mayer, J. E. *Z. Physik* 75, 1 (1932).

[51] Earnshaw, S. *Trans. Camb. Phil. Soc.* 7, 97 (1842).

[52] van Beest, B. W. H.; Kramer, G. J.; van Santen, R. A. *Phys. Rev. Lett.* 64, 1955 (1990).

[53] Halgren, T. A.; Damm, W. *Curr. Opin. Struct. Biol.* 11, 236 (2001).

[54] Lopes, P. E. M.; Lamoureux, G.; Mackerell, A. D. *J. Comp. Chem.* 30, 1821 (2009).

[55] Rick, S. W.; Stuart, S. J. *Rev. Comp. Chem.* 18, 89 (2002).

[56] Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* 91, 6269 (1987).

[57] Applequist, J.; Carl, J. R.; Fung, K.-K. *J. Am. Chem. Soc.* 94, 2952 (1972).

[58] Thole, B. T. *Chem. Phys.* 59, 341 (1981).

[59] Stillinger, F. H. *J. Chem. Phys.* 71, 1647 (1979).

[60] Ahlström, P.; Wallqvist, A.; Engström, S., Jönsson, B. *Mol. Phys.* 68, 563 (1989).

[61] Brodholt, J.; Sampoli, M.; Vallauri, R. *Mol. Phys.* 86, 149 (1995).

[62] Saboungi, M.-L.; Rahman, A.; Halley, J. W.; Blander, M. *J. Chem. Phys.* 88, 5818 (1988).

[63] Sprik, M.; Klein, M. L. *J. Chem. Phys.* 89, 7556 (1988).

[64] van Belle, D.; Lippens, M. F. G.; Wodak, S. J. *Mol. Phys.* 77, 239 (1992).

[65] Halley, J. W.; Rustad, J. R.; Rahman, A. *J. Chem. Phys.* 98, 4110 (1993).

[66] Ren, P.; Ponder, J. W. *J. Comp. Chem.* 23, 1497 (2002).

[67] Grossfield, A.; Ren, P.; Ponder, J. W. *J. Am. Chem. Soc.* 125, 15671 (2003).

[68] Ren, P.; Ponder, J. W. *J. Phys. Chem. B* 107, 5933 (2003).

[69] Ren, P.; Ponder, J. W. *J. Phys. Chem. B* 108, 13427 (2004).

[70] Grossfield, A. *J. Chem. Phys.* 122, 024506 (2005).

[71] Jiao, D.; King, C.; Grossfield, A.; Darden, T. A.; Ren, P. *J. Phys. Chem. B* 110, 18553 (2006).

[72] Applequist, J. *J. Math. Phys.* 24, 736 (1982).

[73] Olson, M. L.; Sundberg, K. R. *J. Chem. Phys.* 69, 5400 (1978).

[74] Applequist, J. *J. Phys. Chem.* 97, 6016 (1993).

[75] Jensen, L.; Åstrand, P.-O.; Mikkelsen, K. V. *Int. J. Quantum Chem.* 84, 513 (2001).

[76] Stern, H. A.; Rittner, F.; Berne, B. J.; Friesner, R. A. *J. Chem. Phys.* 115, 2237 (2001).

[77] Mayer, A. *Phys. Rev. B* 75, 045407 (2007).

[78] Mayer, A.; Åstrand P.-O. *J. Phys. Chem. A* 112, 1277 (2008).

[79] Smalø, H. S.; Åstrand P.-O.; Jensen, L. *J. Chem. Phys.* 131, 044101 (2009).

[80] Sundberg, K. R. *J. Chem. Phys.* 66, 114 (1977).

[81] Sundberg, K. R. *J. Chem. Phys.* 66, 1475 (1977); 67, 4314 (E) (1977).

[82] Jensen, L.; Sylvester-Hvid, K. O.; Mikkelsen, K. V.; Åstrand, P.-O. *J. Phys. Chem. A* 107, 2270 (2003).

[83] Jensen, L.; Åstrand, P.-O.; Mikkelsen, K. V. *Nano Lett.* 3, 661 (2003).

[84] Applequist, J. *J. Am. Chem. Soc.* 95, 8255 (1973).

[85] Applequist, J. *J. Am. Chem. Soc.* 95, 8258 (1973).

[86] Applequist, J. *J. Chem. Phys.* 58, 4251 (1973).

[87] Sundberg, K. R. *J. Chem. Phys.* 68, 5271 (1978).

[88] Applequist, J.; Quicksall, C. O. *J. Chem. Phys.* 66, 3455 (1977).

[89] Bocian, D. F.; Schick, G. A.; Birge, R. R. *J. Chem. Phys.* 74, 3660 (1981).

[90] Bocian, D. F.; Schick, G. A.; Birge, R. R. *J. Chem. Phys.* 75, 2626 (1981).

[91] Bocian, D. F.; Schick, G. A.; Birge, R. R. *J. Chem. Phys.* 75, 3215 (1981).

[92] Bocian, D. F.; Schick, G. A.; Hurd, J. K.; Birge, R. R. *J. Chem. Phys.* 76, 4828 (1982).

[93] Bocian, D. F.; Schick, G. A.; Hurd, J. K.; Birge, R. R. *J. Chem. Phys.* 76, 6454 (1982).

[94] Applequist, J.; Sundberg, K. R.; Olson, M. L.; Weiss, L. C. *J. Chem. Phys.* 70, 1240 (1979); 71, 2330 (1979).

[95] Jensen, L.; Schmidt O. H.; Mikkelsen, K. V.; Åstrand, P.-O. *J. Phys. Chem.* 104, 10462 (2000).

[96] Jensen, L.; Åstrand, P.-O.; Mikkelsen, K. V. *J. Phys. Chem. A* 108, 8795 (2004).

[97] Jensen, L.; Åstrand, P.-O.; Osted, A.; Kongsted, J.; Mikkelsen, K. V. *J. Chem. Phys.* 116, 4001 (2002).

[98] Jensen, L.; Åstrand, P.-O.; Mikkelsen, K. V. *J. Phys. Chem. B* 108, 8226 (2004).

[99] Kongsted, J.; Osted, A.; Jensen, L.; Åstrand, P.-O.; Mikkelsen, K. V. *J. Phys. Chem. B* 105, 10243 (2001).

[100] Hansen, T.; Jensen, L.; Åstrand, P.-O.; Mikkelsen, K. V. *J. Chem. Theor. Comput.* 1, 626 (2005).

[101] Guillaume, M.; Champagne, B. *Struct. Chem.* 15, 385 (2004).

[102] Guillaume, M.; Champagne, B. *Phys. Chem. Chem. Phys.* 7, 3284 (2005).

[103] Lybrand, T. P.; Kollman, P. A. *J. Chem. Phys.* 83, 2923 (1985).

[104] Cieplak, P.; Kollman, P. A.; Lybrand, T. P. *J. Chem. Phys.* 92, 6755 (1990).

[105] Stillinger, F. H.; David, C. W. *J. Chem. Phys.* 69, 1473 (1978).

[106] Kuwajima, S.; Warshel, A. *J. Phys. Chem.* 94, 460 (1990).

[107] Niesar, U.; Corongiu, G.; Clementi, E.; Kneller, G. R.; Bhattacharya, D. K. *J. Phys. Chem.* 94, 7949 (1990).

[108] Caldwell, J.; Dang, L. X.; Kollman, P. A. *J. Am. Chem. Soc.* 112, 9144 (1990).

[109] Kozack, R. E.; Jordan, P. C. *J. Chem. Phys.* 96, 3120 (1992).

[110] Dang, L. X. *J. Chem. Phys.* 97, 2659 (1992).

[111] Wallqvist, A.; Berne, B. J. *J. Phys. Chem.* 97, 13841 (1993).

[112] Bernardo, D. N.; Ding, Y.; Krogh-Jespersen, K.; Levy, R. M. *J. Phys. Chem.* 98, 4180 (1994).

[113] Caldwell, J. W.; Kollman, P. A. *J. Phys. Chem.* 99, 6208 (1995).

[114] Chialvo, A. A.; Cummings, P. T. *J. Chem. Phys.* 105, 8274 (1996).

[115] Dang, L. X.; Chang, T.-M. *J. Chem. Phys.* 106, 8149 (1997).

[116] Burnham, C. J.; Li, J.; Xantheas, S. S.; Leslie, M. *J. Chem. Phys.* 110, 4566 (1999).

[117] Hasegawa, T.; Tanimura, Y. *J. Phys. Chem. B* 115, 5545 (2011).

[118] Xie, W.; Pu. J.; MacKerell, A. D.; Gao, J. *J. Chem. Theor. Comp.* 3, 1878 (2007).

[119] MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L., Jr.; Evanseck, J. D.; Field, S.; Fischer, M. J.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E., III; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* 102, 3586 (1998).

[120] MacKerell, A. D., Jr. *J. Comp. Chem.* 25, 1584 (2004).

[121] Dick, B. G.; Overhauser, A. W. *Phys. Rev.* 112, 90 (1958).

[122] Mitchell, P. J.; Fincham, D. *J. Phys.: Condens. Matter* 5, 1031 (1993).

[123] Cao, J.; Berne, B. J. *J. Chem. Phys.* 99, 2213 (1993).

[124] van Maaren, P. J.; van der Spoel, D. *J. Phys. Chem.* 105,2618 (2001).

[125] Shröder, U. *Solid State Commun.* 4, 347 (1966).

[126] Melvin, J. S.; Pirie, J. D.; Smith, T. *Phys. Rev.* 175, 1082 (1968).

[127] Sangster, M. J. L.; Peckham, G.; Saunderson, D. H. *J. Phys. C: Solid State Phys.* 3, 1026 (1970).

[128] Basu, A. N.; Sengupta, S. *Phys. Stat. Solidi* 29, 367 (1968).

[129] Basu, A. N.; Sengupta, S. *J. Phys. C: Solid State Phys.* 5, 1158 (1972).

[130] Basu, A. N.; Sengupta, S. *Phys. Rev. B* 8, 2982, (1973).

[131] Stoneham, A. M.; Harding, J. H. *Ann. Rev. Phys. Chem.* 37, 53 (1986).

[132] Harding, J. H. *Rep. Progr. Phys.* 53, 1403 (1990).

[133] Lewis, G. V.; Catlow, C. R. A. *J. Phys. C: Salid State Phys.* 18, 1149 (1985).

[134] Tomlinson, S. M.; Catlow, C. R. A.; Harding, J. H. *J. Phys. Chem. Solids* 51, 477 (1990).

[135] Ferneyhough, R.; Fincham, D.; Price, G. D.; Gillan, M. J. *Model. Simul. Mater. Sci. Eng.* 2, 1101 (1994).

[136] Lindan, P. J. D.; Gillan, M. J. *Phylos. Mag. B* 69, 535 (1994).

[137] Lindan, P. J. D. *Mol. Simul.* 14, 303 (1995).

[138] Rambaut, C.; Jobic, H.; Jaffrezic, H.; Kohanoff, J.; Fayeulle, S. *J. Phys.: Condens. Matter* 10, 4221 (1998).

[139] Islam, M. S. *J. Mater. Chem.* 10, 1027 (2000).

[140] Baudin, M.; Hermansson, K. *Surf. Sci.* 474, 107 (2001).

[141] Ricci, D.; Pacchioni, G. *J. Chem. Phys.* 117, 2844 (2002).

[142] Zaoui, A.; Sekkal, W. *Phys. Rev. B* 66, 174106 (2002).

[143] Liu, Z. J.; Sun, X. W.; Chen, Q. F.; Cai, L. C.; Tan, X. M.; Yang, X. D. *Physics Letters A* 353, 221 (2006).

[144] Ramo, D. M.; Sushko, P. V.; Gavartin, J. L.; Shluger A.L. *Phys. Rev. B* 78, 235432 (2008).

[145] Rosen, J.; Warschkow, O. *Phys. Rev. B* 80, 115215 (2009).

[146] Dai, S.; Dunn, M. L.; Park, H. S. *Nanotechnology* 21, 445707 (2010).

[147] Sun, X. W.; Song, T.; Chu, Y. D.; Liu, Z. J.; Zhang, Z. R.; Chen, Q. F. *Solid State Commun.* 150, 1785 (2010).

[148] Lee, W.-J., Chang, J.-G.; Ju, S.-P.; Weng, M.-H.; Lee, C.-H. *Nanoscale Research Letters* 6, 352 (2011).

[149] Devynck, F.; Ianuzzi, M.; Krack, M. *Phys. Rev. B* 85, 184103 (2012).

[150] Behera, R. K.; Deo, C. S. *J. Phys.: Condens. Matter* 24, 215405 (2012).

[151] Hill, J.-R.; Minihan, A. R.; Wimmer, E.; Adams, C. J. *Phys. Chem. Chem. Phys.* 2, 4255 (2000).

[152] Sauer, J.; Sierka, M. *J. Comput. Chem.* 21, 1470 (2000).

[153] Catlow, C. R. A.; Diller, K. M.; Norgett, M. J. *J. Phys. C: Solid State Phys.* 10, 1395 (1977).

[154] Dabringhaus, H. *Surf. Sci.* 600, 941 (2006).

[155] Bichoutskaia, E.; Pyper, N. C. *J. Chem. Phys.* 129, 154701 (2008).

[156] Ivanov, O. V.; Maksimov, E. G. *Phys. Rev. Lett.* 69, 268 (1992).

[157] Ivanov, O. V.; Maksimov, E. G. *Zh. Eksp. Teor. Fiz.* 108, 1841 (1995) [*JETP 81, 1008 (1995)*].

[158] Gordon, R. G.; Kim, Y. S. *J. Chem. Phys.* 56, 3122 (1972).

[159] Ivanov, O. V.; Maksimov, E. G. *Solid State Commun.* 81, 69 (1992).

[160] Boyer, L. L.; Mehl, M. J.; Feldman, J. L.; Hardy, J. R.; Flocken, J. W.; Fong, C. Y. *Phys. Rev. Lett.* 54, 1940 (1985).

[161] Hedin, L.; Lundquist, B. I. *J. Phys. C* 4,2064 (1971).

[162] York, D. M.; Yang, W. *J. Chem. Phys.* 104, 159 (1996).

[163] Chelli, R.; Procacci, P. *J. Chem. Phys.* 117, 9175 (2002).

[164] Sanderson, R. T. *Science* 114, 670 (1951).

[165] Mortier, W.; Ghosh, S.; Shankar, S. *J. Am. Chem. Soc.* 108, 4315 (1986).

[166] Parr, R. G.; Yang, W. *Density-functional theory of atoms and molecules.*, Oxford university press, New-York (1989).

[167] Nistor, R. A.; Polihronov, J. G.; Müser, M. H.; Mosey, N. J. *J. Chem. Phys.* 125, 094108 (2006).

[168] Rappé, A.; Goddard III, W. *J. Phys. Chem.* 95, 3358 (1991).

[169] Chelli, R.; Procacci, P.; Righini, R.; Califano, S. *J. Chem. Phys.* 111, 8569 (1999).

[170] Baekelandt, B. G.; Mortier, W. J.; Lievens, J. L.; Schoonheydt, R. A. *J. Am. Chem. Soc.* 113, 6730 (1991).

[171] Rick, S. W.; Stuart, S. J.; Berne, B. J. *J. Chem. Phys.* 101, 6141 (1994).

[172] Toufar, H.; Baekelandt, B. G.; Janssens, G. O. A.; Mortier, W. J.; Schoonheydt, R. A. *J. Phys. Chem.* 99, 13876 (1995).

[173] Bret, C.; Field, M. J.; Hemmingsen, L. *Mol. Phys.* 98, 751 (2000).

[174] Medeiros, M. *Theor. Chem. Acc.* 113, 178 (2005).

[175] Chelli, R.; Schettino, V.; Procacci, P. *J. Chem. Phys.* 122, 234107 (2005).

[176] Chelli, R.; Barducci, A.; Bellucci, L.; Schettino, V.; Procacci, P. *J. Chem. Phys.* 123, 194109 (2005).

[177] Piquemal, J.-P.; Chelli, R.; Procacci, P.; Gresh, N. *J. Phys. Chem. A* 111, 8170 (2007).

[178] Stuart, S. J.; Berne, B. J. *J. Phys. Chem.* 100, 11934 (1996).

[179] Stuart, S. J.; Berne, B. J. *J. Phys. Chem. A* 103, 10300 (1999).

[180] Chelli, R.; Ciabatti, S.; Cardini, G.; Righini, R.; Procacci, P. *J. Chem. Phys.* 111, 4218 (1999).

[181] Llanta, E.; Ando, K.; Rey, R. *J. Phys. Chem. B* 105, 7783 (2001).

[182] Patel, S.; Brooks, C. L. *J. Comput. Chem.* 25, 1 (2004).

[183] Patel, S.; Mackerell, A. D.; Brooks, C. L. *J Comput. Chem.* 25, 1504 (2004).

[184] Chelli, R.; Pagliai, M.; Procacci, P.; Cardini, G.; Schettino, V. *J. Chem. Phys.* 122, 074504 (2005).

[185] Tanaka, M.; Siehl, H.-U. *Chem. Phys. Lett.* 457, 263 (2008).

[186] Rick, S. W.; Berne, B. J. *J. Am. Chem. Soc.* 118, 672 (1996).

[187] Chen, B.; Xing, J.; Siepmann, J. I. *J. Phys. Chem. B* 104, 2391 (2000).

[188] Stern, H. A.; Kaminski, G. A.; Banks, J. L.; Zhou, R.; Berne, B. J.; Friesner, R. A. *J. Phys. Chem. B* 103, 4730 (1999).

[189] Banks, J. L.; Kaminski, G. A.; Zhou, R.; Mainz, D. T.; Berne, B. T. *J. Chem. Phys.* 110, 741 (1999).

[190] Kitao, O.; Ogawa, T. *Mol. Phys.* 101, 3 (2003).

[191] Bultinck, P.; Langenaeker, W.; Carbó-Dorca, R.; Tollenaere, J. P. *J. Chem. Inf. Comput. Sci.* 43, 422 (2003).

[192] Ogawa, T.; Kurita, N.; Sekino, H.; Kitao, O.; Tanaka, S. *Chem. Phys. Lett.* 397, 382 (2004).

[193] Bultinck, P.; Vanholme, R.; Popelier, P. L. A.; De Proft, F.; Geerlings, P. *J. Phys. Chem. A* 108, 10359 (2004).

[194] Mathieu, D. *J. Chem. Phys.* 127, 224103 (2007).

[195] Verstraelen, T.; van Speybroeck, V.; Waroquier, M. *J. Chem. Phys.* 131, 044127 (2009).

[196] Mikulski, P. T.; Knippenberg, M. T.; Harrison, J. A. *J. Chem. Phys.* 131, 241105 (2009).

[197] Smirnov, K. S.; van de Graaf, B. *J. Chem. Soc., Faraday Trans.* 92, 2475 (1996).

[198] Smirnov, K. S.; Thibault-Starzyk, F. *J. Phys. Chem. B* 103, 8595 (1999).

[199] Brodersen, S.; Wilke, S.; Leusen, F. J. J.; Engel, G. E. *Cryst. Growth Des.* 5, 925 (2005).

[200] Van Genechten, K. A.; Mortier, W. J.; Geerlings, P. *J. Chem. Phys.* 86, 5063 (1987).

[201] Janssens, G. O. A.; Baekelandt, B. G.; Toufar, H.; Mortier, W. J.; Schoonheydt, R. A. *J. Phys. Chem.* 99, 3251 (1995).

[202] Sefcik, J.; Demiralp, E.; Cagin, T.; Goddard, W. A. *J. Comput. Chem.* 23, 1507 (2002).

[203] Ercolessi, F.; Adams, J. B. *Europhys. Lett.* 26, 583 (1994).

[204] Silvestrelli, P. L. *Phys. Rev. Lett.* 100, 053002 (2008).

[205] Yang, Z. Z.; Wang, C. S. *J. Phys. Chem. A* 101, 6315 (1997).

[206] Chen, J.; Martínez, T. J. *Chem. Phys. Lett* 438, 315 (2007).

[207] Giese, T. J.; York, D. M. *J. Chem. Phys.* 120, 9903 (2004).

[208] Tabacchi, G.; Mundy, C. J.; Hutter, J.; Parrinello, M. *J. Chem. Phys.* 117, 1416 (2002).

[209] Masia, M.; Probst, M.; Rey, R. *J. Chem. Phys.* 121, 7362 (2004).

[210] Njo, S. L.; Fan, J. F.; van de Graaf, B. *J. Mol. Catal. A* 134, 79 (1998).

[211] Warren, G. L.; Davis, J. E.; Patel, S. *J. Chem. Phys.* 128, 144110 (2008).

[212] Verstraelen, T.; Pauwels, E.; De Proft, F.; Van Speybroeck, V.; Geerlings, P.; Waroquier, M. *J. Chem. Theory Comput.* 8, 661 (2012).

[213] Nistor, R. A.; Müser, M. H. *Phys. Rev. B* 79, 104303 (2009).

[214] Weber, V.; Niklasson, A. M. N.; Challacombe, M. *Phys. Rev. Lett.* 92, 193002 (2004).

[215] Krishtal, A.; Senet, P.; Yang, M.; Van Alsenoy, C. *J. Chem. Phys.* 125, 034312 (2006).

[216] Kirtman, B.; Toto, J. L.; Robins, K. A.; Hasan, M. *J. Chem. Phys.* 102, 5350 (1995).

[217] van Faassen, M.; de Boeij, P. L.; van Leeuwen, R.; Berger, J. A.; Snijders, J. G. *Phys. Rev. Lett.* 88, 186401 (2002).

[218] Parr, R. G.; Ayers, P. W.; Nalewajski, R. F. *J. Phys. Chem. A* 109, 3957 (2005).

[219] Bultinck, P.; Langenaeker, W.; Lahorte, P.; De Proft, F.; Geerlings, P.; Van Alsenoy, C.; Tollenaere, J. P. *J. Phys. Chem. A* 106, 7895 (2002) .

[220] Jiroušková, Z.; Svobodová Vařeková, R.; Vaněk, J.; Koča, J. *J. Comput. Chem.* 30, 1174 (2009).

[221] Geerlings, P.; De Proft, F.; Langenaeker, W. *J. Chem. Rev.* 103, 1793 (2003).

[222] Bultinck, P.; Van Alsenoy, C.; Ayers, P. W.; Dorca, R. C. *J. Chem. Phys.* 126, 144111 (2007).

[223] Bultinck, P.; Ayers, P. W.; Fias, S.; Tiels, K.; Van Alsenoy, C. *Chem. Phys. Lett.* 444, 205 (2007).

[224] Van Damme, S.; Bultinck, P.; Fias, S. *J. Chem. Theory Comput.* 5, 334 (2009).

[225] Catak, S.; D'hooghe, M.; Verstraelen, T.; Hemelsoet, K.; Van Nieuwenhove, A.; Ha, H.-J.; Waroquier, M.; De Kimpe, N.; Van Speybroeck, V. *J. Org. Chem.* 75, 4530 (2010).

[226] Baerlocher, C.; McCusker, L.; Olson, D. *Atlas of Zeolite Framework Types*, 6th ed.; Elsevier: Amsterdam, The Netherlands, p 1 (2007). http://www.iza-structure.org/databases/

[227] Becke, A. D. *J. Chem. Phys.* 98, 5648 (1993).

[228] Raghavachari, K.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* 72, 650 (1980).

[229] McLean, A. D.; Chandler, G. S. *J. Chem. Phys.* 72, 5639 (1980).

[230] Hay, P. J.; Wadt, W. R. *J. Chem. Phys.* 82, 270 (1985).

[231] Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; and Pople, J. A. *Gaussian 03, Revision D.01*, Gaussian, Inc., Wallingford CT, 2004..

[232] Dovesi, R.; Saunders, V. R.; Roetti, C.; Orlando, R.; Zicovich-Wilson, C. M.; Pascale, F.; Civalleri, B.; Doll, K.; Harrison, N. M.; Bush, I. J.; D'Arco, P.; Llunell, M. *CRYSTAL06 User's Manual*. University of Torino: Torino, 2006.

[233] Gatti, C.; Saunders, V.R.; Roetti, C.; *J. Chem. Phys.* 101, 10686 (1994).

[234] Catti, M.; Valerio, G.; Dovesi R.; M. Causà, M. *Phys. Rev. B* 49, 14179 (1994).

[235] Pascale, F.; Zicovich-Wilson, C. M.; Orlando, R.; Roetti, C.; Ugliengo, P.; Dovesi, R. *J. Phys. Chem. B* 109, 6146 (2005).

[236] http://www.tcm.phy.cam.ac.uk/˜mdt26/crystal.html

[237] http://molmod.ugent.be/code/wiki/HiPart

[238] Becke, A. D. *J. Chem. Phys.* 88, 2547 (1988).

[239] Verstraelen, T.; Bultinck, P.; Van Speybroeck, V.; Ayers, P. W.; Van Neck, D.; Waroquier, M. *J. Chem. Theory Comput.* 7, 1750 (2011).

[240] Press, W. H.; Flannery, B. P.; Teukolsky, S. A.; Vetterling, W. T. Conjugate Gradient Methods in Multidimensions. In *Numerical Recipes in C: The Art of Scientific Computing*; Cambridge University Press: New York, 1992; p 420.

[241] Politzer, P. *J. Chem. Phys.* 86, 1072 (1987).

[242] Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* 5, 129 (1987).

[243] Mantina, M.; Chamberlin, A. C.; Valero, R.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. A* 113, 5806 (2009).

[244] Kokalj, A. *Comp. Mater. Sci.* 28, 155 (2003). Code available from http://www.xcrysden.org/.

[245] Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. *J. Phys. Chem.* 97, 10269 (1993).

[246] Campanã, C.; Mussard, B.; Woo, T. K. *J. Chem. Theory Comput.* 5, 2866 (2009).

[247] Pyper, N. C. In *Advances in Solid-State Chemistry* vol 2 ed Catlow C. R. A. (London: JAI Press) p.223 (1991).

[248] Tangney, P.; Scandolo, S. *J. Chem. Phys.* 117, 8898 (2002).

[249] Beck, P.; Brommer, P.; Roth, J.; Trebin, H.-R. *J. Chem. Phys.* 135, 234512 (2011).

[250] Meister, J.; Schwarz, W. H. E. *J. Phys. Chem.* 98, 8245 (1994).

[251] Grimme, S. *J. Comput. Chem.* 27, 1787 (2006).

[252] Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* 132, 154104 (2010).

[253] Johnson, E. R.; Becke, A. D. *J. Chem. Phys.* 123, 024101 (2005).

[254] Becke, A. D.; Johnson, E. R. *J. Chem. Phys.* 127, 154108 (2007).

[255] Silvestrelli, P. L. *J. Phys. Chem. A* 113, 5224 (2009).

[256] Tkatchenko, A.; Scheffler, M. *Phys. Rev. Lett.* 102, 053002 (2009).

[257] Krishtal, A.; Geldof, D.; Vanommeslaegje, K.; Van Alsenoy, C.; Geerlings, P. *J. Chem. Theory Comput.* 8, 125 (2012).

[258] Grimme, S. *WIREs Comput. Mol. Sci.* 1, 211 (2012).

[259] Johnson, E. R.; Mackie, I. D.; DiLabio, G. A. *J. Phys. Org. Chem.* 22, 1127 (2009).

[260] Klimeš, J.; Michaelides, A. *J. Chem. Phys.* 137, 120901 (2012).

[261] Andersson, Y.; Langreth, D. C.; Lundqvist, B. I. *Phys. Rev. Lett.* 76, 102 (1996).

[262] Silvestrelli, P. L. *Chem. Phys. Lett.* 475, 285 (2009).

[263] Costanzo, F.; Venuti, E.; Della Valle, R. G.; Brillante, A.; Silvestrelli, P. L. *J. Phys. Chem. C* 114, 20068 (2010).

[264] Andrinopoulos, L.; Hine, H. D. M.; Mostofi, A. A. *J. Chem. Phys.* 135, 154105 (2011).

[265] Rotenberg, B.; Salanne, M.; Simon, C.; Vuilleumier, R. *Phys. Rev. Lett.* 104, 138301 (2010).

[266] Tazi, S.; Molina, J. J.; Rotenberg, B.; Turq, P.; Vuilleumier, R.; Salanne, M. *J. Chem. Phys.* 136, 114507 (2012).

[267] Catlow, C. R. A. *Computational Materials Science* (*NATO Science Series* vol 187) eds C. R. A. Catlow and E. A. Kotomin (Amsterdam: IOS Press) p.1 (2003).

[268] Conesa, J. C. *J. Chem. Phys. C* 114, 22718 (2010).

[269] Boys, S. F. *Rev. Mod. Phys.* 32, 296 (1960).

[270] Marzari, N., Mostofi, A. A.; Yates, J. R.; Souza, I.; Vanderbilt, D. *Rev. Mod. Phys.* 84, 1419 (2012).

[271] Tang, K. T.; Toennies, J. P. *J. Chem. Phys.* 80, 3726 (1984).

[272] Thygesen, K. S.; Hansen, L. B.; Jacobsen, K. W. *Phys. Rev. Lett.* 94, 026405 (2005).

[273] Thygesen, K. S.; Hansen, L. B.; Jacobsen, K. W. *Phys. Rev. B* 72, 125119 (2005).

[274] Souza, I.; Marzari, N.; Vanderbilt, D. *Phys. Rev. B* 65, 035109 (2001).

[275] Tang, K. T. *Phys. Rev.* 177, 108 (1969).

[276] Kramer, H. L.; Herschbach, D. R. *J. Chem. Phys.* 53, 2792 (1970).

[277] Cambi, R.; Cappelletti, D.; Liuti, G.; Pirani, F. *J. Chem. Phys.* 95, 1852 (1991).

[278] CPMD, http://www.cpmd.org/, Copyright IBM Corp. 1990-2008, Copyright MPI für Festkörperforschung Stuttgart 1997-2001.

[279] Wilson, M.; Madden, P. A.; Pyper, N. C.; Harding, J. H. *J. Chem. Phys.* 104, 8068 (1996).

[280] Dimitrov, V.; Sakka, S. *J. Appl. Phys.* 79, 1736 (1996).

[281] Pyykkö, P.; Atsumi, M. *Chem. Eur. J.* 15, 186 (2009).

[282] Heaton, R. J.; Madden, P. A.; Clark, S. J.; Jahn, S.; Turq, P.; Guillot, B. *J. Chem. Phys.* 125, 144104 (2006).

[283] Salanne, M.; Vuilleumier, R.; Madden, P. A.; Simon, C.; Turq, P.; Guillot, B. *J. Phys.: Condens. Matter* 20, 494207 (2008).

[284] Fowler, P. W.; Harding, J. H.; Pyper, N. C. *J. Phys.: Condens. Matter* 6, 10593 (1994).

[285] Verstraelen, T.; Sukhomlinov, S. V.; van Speybroeck, V.; Waroquier, M.; Smirnov, K. S. *J. Phys. Chem. C* 116, 490 (2012).

[286] Knippenberg, M. T.; Mikulski, P. T.; Ryan, K. E.; Stuart, S. J.; Gao, G.; Harrison, J. A. *J. Chem. Phys.* 136, 164701 (2012).

[287] Vuilleumier, R.; Sator, N.; Guillot, B. *J. Non-Cryst. Solids* 357, 2555 (2011).

[288] Carré, A.; Horbach, J.; Ispas, S.; Kob, W. *Europhys. Lett.* 82, 17001 (2008).

[289] Cochran, W. *CRC Crit. Rev. Solid State Mater. Sci.* 2, 1 (1971).

[290] Ree, F.; Holt, A. *Phys. Rev. B* 8, 826 (1973).

[291] Peacor, D. R. *Z. Krist.* 138, 274 (1973).

[292] Zhang, Y.; Pan, W.; Yang, W. *J. Chem. Phys.* 107, 7921 (1997).

[293] Wu, X.; Vargas, M. C.; Nayak, S.; Lotrich, V.; Scoles, G. *J. Chem. Phys.* 115, 8748 (2001).

[294] Dion, M.; Rydberg, H.; Schröder, E.; Langreth, D. C.; Lundqvist, B. I. *Phys. Rev. Lett.* 92, 246401 (2004).

[295] Izvekov, S.; Parrinello, M; Burnham, C. J.; Voth, G. A. *J. Chem. Phys.* 120, 10896 (2004).

[296] Carré, A. PhD dissertation (2007).

[297] Demiralp, E.; Çağin, T.; Goddard, W. A. III *Phys. Rev. Lett.* 82, 1708 (1999).

[298] Beck, P.; Brommer, P.; Roth, J.; Trebin, H.-R. *J. Chem. Phys.* 135, 234512 (2011).

[299] Dimitrov, V.; Sakka, S. *J. Appl. Phys.* 79, 1736 (1996).

[300] Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids* Oxford Science Publication (1987).

[301] Tsuneyuki, S.; Tsukada, M.; Aoki, H.; Matsui, Y. *Phys. Rev. Lett.*, 61, 869 (1988).

[302] Hill, J.-R.; Sauer, J. *J. Phys. Chem.*, 98 1238 (1994).

[303] Hill, J.-R.; Sauer, J. *J. Phys. Chem.*, 99 9536 (1995).

[304] Schröder, K.-P.; Sauer, J. *J. Phys. Chem.*, 100, 11043 (1996).

[305] Ermoshin, V. A.; Smirnov, K. S.; Bougeard, D. *Chem. Phys.*, 202, 53 (1996).

[306] Ermoshin, V. A.; Smirnov, K. S.; Bougeard, D. *Chem. Phys.*, 209, 41 (1996)

[307] Demontis, P.; Suffritti, G. B.; Quartieri, S.; Fois, E. S.; Gamba, A. *J. Phys. Chem.*, 92, 867 (1988).

[308] Kermode, J. R.; Cereda, S.; Tangney, P.; De Vita, A. *J. Chem. Phys.*, 133, 094102 (2010).

[309] Herzbach, D.; Binder, K.; Müser M. H. *J. Chem. Phys.*, 123, 124711 (2005).

[310] Wolkenshtein, M. W. *Compt. R. Acad. Sci. URSS*, 30, 791 (1941).

[311] D. Bougeard, D.; Smirnov, K. S. *J. Raman Spectrosc.*, 40, 1704 (2009)

[312] Umari, P.; Pasquarello, A.; Dal Corso A. *Phys. Rev. B*, 63, 094305 (2001)

[313] Antao S. M.; Hassan I.; Wang J.; Lee P. L; Toby B. H. T. Can. Mineral. 46, 1501 (2008).

[314] Loudon, R. *Adv. Phys.* 13, 423 (1964).

[315] Smirnov, K. S.; Bougeard, D.; Tandon. P. *J. Phys. Chem. A*, 110, 4516 (2006).

[316] Pluth J. J.; Smith J. V.; Faber, J. *J. Appl. Phys.*, 57, 1045 (1985).

[317] Database of Raman spectroscopy: http://www.rruff.info.

[318] Ewald, P. P. *Ann. Phys.* 64, 253 (1921).

[319] Smirnov, K. S.; Bougeard, D. *Chem. Phys.* 292, 53 (2003).

[320] Le Page, Y.; Donnay, G. *Acta Crystallogr. B* 32, 2456 (1976).

[321] Pluth, J. J.; Smith, J. V.; Faber, J. Jr. *J. Appl. Phys.* 57, 1045 (1985).

[322] Bouvier, P.; Djurado, E.; Ritter, C.; Dianoux, A. J.; Lucazeau, G. *Int. J. Inorg. Mater.* 3, 647 (2001).

[323] Howard, C. J.; Hill, R. J.; Reichert, B. E. *Acta Crystallogr.* B44, 116 (1988).

[324] Vanderbilt, D.; Zhao, X.; Ceresoli, D. *Thin Solid Films* 486, 125 (2005).

[325] Ceresoli, D.; Vanderbilt, D. *Phys. Rev. B* 74, 125108 (2006).

[326] Yu, J.; Devanathan, R.; Weber, W. J. *J. Mater. Chem.* 19, 3923 (2009).

[327] Louwerse, M. J.; Baerends, E. J. *Chem. Phys. Lett.* 421, 138 (2006).

List of Figures

2.1 Periodic boundary conditions in two-dimensions. The system of particles is enclosed into a cell (shown as shaded rectangle) with the cell parameters $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$. The principal cell is then replicated in the space (plane of the figure) to form an infinite 2D system. The particles in the principal cell interact not only with each other, but also with particles in the images of the cell. 21

3.1 Schematic representation of atom i in the shell model. The shell particle with negative charge Y_i is attached to the core having positive charge Z_i by a harmonic spring with spring constant K_i. The core-shell separation distance is d_i and net atomic charge Q_i is equal to $Q_i = Z_i + Y_i$. 52

4.1 Dependence of EEM/HI total X_λ, static X_S, and response X_R cost functions on the weight λ, 4.2.10. Horizontal blue and red lines indicate "worst-case" values of static and response cost functions, respectively (see text for discussion). 72

4.2 Dependence of SQE/ESP total X_λ, static X_S, and response X_R cost functions on weight λ, 4.2.10. Horizontal blue and red lines indicate "worst-case" values of static and response cost functions, respectively (see text for discussion; vertical dashed lines indicate the selected point. 74

4.3 Correlation between the reference DFT HI charges and those predicted by the EEM and SQE models for molecules in the validation set. For the sake of clarity the EEM and SQE data were equally shifted down and up, respectively, along the y-axis. . . 76

4.4 Relative error in ESP values and correlation coefficient between reference and model ESP values *vs* electronic density for molecules in validation set; results obtained with SQE/ESP parameters of 4.3. Vertical dashed line corresponds to the threshold density value. 81

4.5 Plot of electronic density value in a given point as a function of distance from closest oxygen atom. Dashed lines indicate coordinates for $\rho = 10^{-3}$ a.u. 82

4.6 Comparison of the dipole moment components for the molecules in the validation set, obtained with the reference DFT method and with the SQE/ESP model. . . . 83

4.7 Comparison of all principal components α_i of the polarizability tensors for the molecules in the validation set, obtained with the reference DFT method and with the SQE/ESP model. 84

4.8 Left: image of α-quartz structure showing the region, where electronic density value is below 10^{-3} a.u. (in yellow color); silicon and oxygen atoms are shown as cyan and red balls, respectively. Right: image of sodalite unit cell with the isodensity surface corresponding to $\rho = 10^{-3}$ a.u. The images were generated with XCrySDen program [244]. 84

4.9 Relative error in ESP values and correlation coefficient between reference and model ESP values *vs* electronic density in α-quartz structure; results obtained with SQE/ESP parameters of 4.3. Vertical dashed line corresponds to the threshold density value $\rho = 10^{-3}$ a.u. 85

4.10 Relative error in ESP values and correlation coefficient between reference and model ESP values *vs* electronic density in sodalite structure; results obtained with SQE/ESP parameters of 4.3. Vertical dashed line corresponds to the threshold density value $\rho = 10^{-3}$ a.u. 85

4.11 ESP (upper panel) and ESP gradient (lower panel) obtained in periodic DFT and SQE/ESP computations along the $< 100 >$ direction in sodalite structure (lines connecting the centers of opposite 4R rings). The electron density in the region between the dashed lines is below the threshold value $\rho = 10^{-3}$ a.u. 86

5.1 Position of Wannier function centers in crystalline silicon (a) and in amorphous silica (b) (fragments of structures are shown). The silicon and oxygen atoms, and Wannier function centers are the wheat, red, and small green balls, respectively. . 93

5.2 The correlation of the C_6^{ij} coefficients computed by the effective orbital method (blue circles) and by the combination rules (5.2.15) and (5.2.16) (green squares and magenta diamonds, respectively) with the reference C_6^{ij} values obtained by the original Sivestrelli approach (x axis) for the oxygen-oxygen (a) and zirconium-oxygen (b) interactions in the LD structure of a-ZrO_2. For the sake of clarity, in figure (a) the C_6^{OO} values computed by the combination rules (5.2.15) and (5.2.16) were equally shifted up and down, respectively, along the y axis. Red dashed line is the identity line. 97

5.3 Distribution of the C_6^{OO} dispersion coefficient values in silica and zirconia "cpmd" structures. 99

5.4 Self-atom dispersion coefficients S_6^{OO} in silica (a) and zirconia (b) structures as a function of mean nearest-neighbour distance R_{MO} (M=Si, Zr). In figure (b) CN stands for the coordination number of oxygen atoms. 101

5.5 Correlation of self-atom dispersion coefficients S_6^{OO} with the oxygen atom charge Q_{O} computed with SQE/ESP method [285]. 102

5.6 Correlation of S_6^{OO} coefficients computed by (5.4.1) (y-axis) with reference S_6^{OO} values (x-axis) in a-SiO_2 structure. 104
5.7 Correlation of interatomic C_6^{OO} coefficients computed by (5.4.2) and (5.2.16) (y-axis) with reference C_6^{OO} values (x-axis) in "cpmd" structures of tetragonal (grey circles) and cubic (open squares) ZrO_2. 105
5.8 Correlation of dispersion energies (a) and x-component of Cartesian force (b) computed with fixed (x-axis) and structure-dependent (y-axis) dispersion coefficients for oxygen atoms in amorphous silica structure. The superscript of the axis labels denotes the method used to compute the corresponding quantity (see text for details). 106

6.1 Scheme showing the preparation of amorphous silica configurations starting from the crystalline structure of β-cristobalite. 113
6.2 Dependence of the χ_F^2 cost function (left) and of the ΔF merit criterion (right) on the number of points in the grid. The spline approximation was performed for the total *ab initio* force. 117
6.3 Pair forces (a) and effective pair potentials (b) obtained from the force-matching of the total *ab initio* forces with splines. 118
6.4 Pair forces (a) and effective pair potentials (b) obtained from the force-matching of forces with splines in the ESP ansatz. 119
6.5 Pair forces (a) and effective pair potentials (b) obtained from the force-matching of forces with splines in the HI ansatz. 119
6.6 ESP short-range energy curves (Figure 6.4) in the logarithmic scale. Curve labeled "exponential fit" shows a fit of the O-O potential curve with function $F(r) = A_0e^{-br} + A_1e^{-2br} + A_2e^{-3br} + A_3e^{-4br}$, and curve labeled "power series fit" corresponds to a fit of the O-O potential with function $F_s(r) = Ae^{-br} - C_4r^{-4} - C_6r^{-6} - C_8r^{-8} - C_{10}r^{-10}$. The potential curve for Si-O was shifted into the positive energies region. 121
6.7 (a): dependence of the merit functions ΔF (6.2.4) and ΔS (6.2.5) on the coefficient λ for the *exp-4* potential. (b): correlation between the *ab initio* forces and the forces computed with the model. 126
6.8 (a): dependence of the merit functions ΔF (6.2.4) and ΔS (6.2.5) on the coefficient λ for the *Morse-SR* potential. (b): correlation between the *ab initio* forces and the forces computed with the model. 127

7.1 Si-O radial distribution function (left), and O-Si-O and Si-O-Si angle distribution functions (right) in α-quartz computed with the *exp-4* (black line) and BKS (red line) models. 134
7.2 Snapshots of the initial (left) and final (right) configurations of the α-cristobalite structure from the NPT MD run with the *exp-4* model; prespective view is shown. 135

7.3 Si-O (left) and O-O (right) radial distribution functions for α-quartz in the *Morse-SR* model. 136
7.4 Density of vibrational states for α-quartz obtained with the *Morse-SR* model (black) and BKS potetial (red). 137
7.5 Polarized Raman spectra of α-quartz computed for the $x(zz)x$ (left) and $x(yz)x$ (right) geometries. The experimental spectra were taken from Ref. [315]. 138
7.6 Density of vibrational states (left) and the powder Raman spectra (right) of α-cristobalite. The experimental spectrum was taken from Ref. [317]. 138
7.7 Spectra of the polarizability tensor components and of their combinations revealing the Raman active vibrational modes of specific symmetry in the structure of the α-cristobalite. 139
7.8 Si-O (left) and O-O (right) radial distribution functions for the β-cristobalite system at 300 K in comparison with the α-phase. 140

List of Tables

4.1 Number of atoms of each element in training and validation sets for molecular systems. 69
4.2 Structural characteristics of periodic systems. 69
4.3 EEM and SQE parameters obtained in the NLLSQ fits. 75
4.4 Atomic charges in periodic systems obtained with different schemes. 77
4.5 Elements of **G** matrix (cf. 4.3.7) computed for the oxygen and silicon atoms in crystalline silica structures using the parameters of EEM/HI scheme. 79
4.6 Hirshfeld-I atomic charges computed with the molecular EEM/HI and SQE/HI models, and with the models corrected for solid state calculations (EEM_S/HI and SQE_S/HI); the right-hand column reports the reference Hirshfeld-I charges obtained in periodic DFT computations (DFT/HI). 80

5.1 Number of atoms N, shape, and size of simulation cell for studied systems. 96
5.2 C_6^{ij} dispersion coefficients (in a.u.) in MO_2 oxide structures (M=Si, Zr), values in parentheses are the standard deviations. 98
5.3 Self-atom dispersion coefficients S_6^{ii} (in a.u.), atomic polarizabilities α_i (in $Å^3$), and mean values of bond lengths R_{M-O} (in Å) in MO_2 oxide structures (M=Si, Zr). Values in parentheses are the standard deviations. 100

6.1 Parameters of the *exp-4* model potential, eq. (6.5.7). 127
6.2 Parameters of the *Morse-SR* model potential, eq. (6.5.9). 128

7.1 Number of atoms N, shape, and size of simulation cell for studied systems. 132
7.2 Lattice parameters of the α-quartz structure computed with the *exp-4* potential. . 134
7.3 Lattice parameters of the α-quartz structure obtained with the *Morse-SR* potential. 135
7.4 Lattice parameters of the α-cristobalite structure obtained with the *Morse-SR* potential. 138

Développement de modèle de potentiels effectifs d'interactions interatomiques pour la modélisation d'oxydes

Le modèle de potentiels effectifs d'interactions interatomiques (champ de forces) pour la modélisation d'oxydes a été développé avec l'utilisation de calculs *ab initio* basés sur la théorie de la fonctionnelle de la densité. Le champ de forces décrit l'énergie potentielle totale du système par la somme de l'énergie électrostatique, celle de dispersion, et l'énergie d'interactions à courte portée.

L'énergie électrostatique a été décrite par le modèle d'équilibration des tranferts de charge (SQE) basé sur l'approche d'égalisation du potentiel chimique (CPE). Le calcul de coefficients de dispersion, qui déterminent les interactions de dispersion, a été réalisé avec l'utilisation de fonctions Wannier maximalement localisées (MLWF). Dans les oxydes la position des centres des MLWF près d'atomes permet de calculer les coefficients de dispersion pour chaque atome. Les calculs de ces coefficients ont montré que leur valeur dépend du nombre d'atomes et du rayon de la première sphère de coordination. Le développement de potentiels d'interactions à courte portée a été réalisé avec l'utilisation de la méthode "force-matching", ce qui a permis de choisir la forme analytique des potentiels. Les paramètres des composants du champ de force ont été obtenus sur la base de calculs de chimie quantique de systèmes isolés et périodiques de structures de silicates. Les paramètres du modèle SQE on été calibrés en utilisant le potentiel électrostatique comme la grandeur de référence.

Le champ de forces complet a été testé par simulation de polymorphes cristallins de la silice par la méthode de la dynamique moléculaire. Les résultats des calculs ont permis de choisir le meilleur modèle. Le champ de forces sélectionné reproduit bien les caractéristiques structurelles de α-quartz et α-cristobalite. Le calcul de spectres vibrationnels des systèmes montre que le champ de forces sous-estime les constantes de forces Si-O, ce qui conduit à un déplacement de spectres vibrationnels vers les basses fréquences par rapport aux spectres expérimentaux. Des voies visant l'amélioration de la performance du champ de forces sont proposées.

Mots-clés : champ de forces, calculs de la fonctionnelle de la densité, égalisation du potentiel chimique, interactions de dispersion, silicates, silice, structure, dynamique.

Development of effective interatomic potentials for computer simulation of oxides

The effective interatomic potential model (force field) for the atomistic modeling of oxide materials was developed with the extensive use of *ab initio* density functional calculations. The force field represents the total potential energy of system as a sum of the long-range electrostatic, dispersion, and short-range energy contributions.

The long-range energy electrostatic energy was described with the use of split-charge equilibration (SQE) model based on the chemical potential equalization (CPE) approach. The electrostatic potential was used as the reference quantity for the calibration of parameters of the SQE model. The computation of dispersion coefficients, which determine the magnitude of the dispersion interactions, was carried out with the use of maximally localized Wannier functions (MLWF). The position of MLWF centers close to the nuclei in oxides permits the computation of the dispersion coefficients in an atom-wise manner. The values of the dispersion coefficients were found to be affected by the coordination number and the radius of the first coordination sphere of atom. The short-range (SR) interaction potentials were designed with the use of force-matching method, which has allowed a judicious choice of the functional form of the SR potentials. The parametrization of the force field components was performed on the basis of extensive quantum-chemical calculations of isolated and periodic silicate systems.

The complete force field was tested in the molecular dynamics simulations of crystalline silica polymorphs. Results of the calculations allowed to choose the best model. The selected force field well reproduces structural characteristics of the α-quartz and α-cristobalite polymorphs. The calculation of the vibrational spectra of the systems has shown that the model underestimates the Si-O force constants that leads to a downward shift of the vibrational spectra in comparison with the experimental data. A number of ways aimed at improving the force field's performance are suggested.

Keywords : force field, density functional calculations, chemical potential equalization, dispersion interactions, silicates, silica, structure, dynamics.