Dissertation zur Erlangung des Grades eines
„Doktor rerum naturalium (Dr. rer. nat.)“ der Fachbereiche:
08 - Physik, Mathematik und Informatik,
09 - Chemie, Pharmazie und Geowissenschaften,
10 - Biologie,
04 - Universitätsmedizin
am Max-Planck-Graduate-Center mit der
Johannes Gutenberg–Universität Mainz

Prediction of New Crystal Structures under Extreme Conditions

von Daniel Schärf

geb. am 07.09.1984 in Dernbach

Datum der Einreichung: 29. 10. 2014
Referent:
Korreferent:
Datum der mündlichen Prüfung: 16. 12. 2014

Abstract

One of the most important challenges in chemistry and material science is the connection between the contents of a compound and its chemical and physical properties. In solids, these are greatly influenced by the crystal structure.

The prediction of hitherto unknown crystal structures with regard to external conditions like pressure and temperature is therefore one of the most important goals to achieve in theoretical chemistry. The stable structure of a compound is the global minimum of the potential energy surface, which is the high dimensional representation of the enthalpy of the investigated system with respect to its structural parameters. The fact that the complexity of the problem grows exponentially with the system size is the reason why it can only be solved via heuristic strategies.

Improvements to the artificial bee colony method, where the local exploration of the potential energy surface is done by a high number of independent walkers, are developed and implemented. This results in an improved communication scheme between these walkers. This directs the search towards the most promising areas of the potential energy surface.

The minima hopping method uses short molecular dynamics simulations at elevated temperatures to direct the structure search from one local minimum of the potential energy surface to the next. A modification, where the local information around each minimum is extracted and used in an optimization of the search direction, is developed and implemented. Our method uses this local information to increase the probability of finding new, lower local minima. This leads to an enhanced performance in the global optimization algorithm.

Hydrogen is a highly relevant system, due to the possibility of finding a metallic phase and even superconductor with a high critical temperature. An application of a structure prediction method on SiH_{12} finds stable crystal structures in this material. Additionally, it becomes metallic at relatively low pressures.

Zusammenfassung

Eine der wichtigsten Herausforderungen in der Chemie und der Materialwissenschaft ist die Verknüpfung der Zusammensetzung einer Verbindung mit ihren chemischen und physikalischen Eigenschaften. In Festkörpern sind diese erheblich durch die Kristallstruktur des Systems beeinflusst.

Die Vorhersage von bisher unbekannten Kristallstrukturen in Bezug auf äußere Bedingungen wie Druck und Temperatur ist daher eines der wichtigsten Ziele in der theoretischen Chemie. Die stabile Struktur einer Verbindung ist das globale Minimum ihrer Potentialfläche, welche die hochdimensionale Darstellung der Enthalpie des untersuchten Systems hinsichtlich seiner Strukturparameter ist. Die Tatsache, dass die Komplexität des Problems exponentiell mit der Systemgröße ansteigt ist der Grund, warum es nur durch heuristische Strategien gelöst werden kann.

Verbesserungen der Artificial Bee Colony Methode, bei der die lokale Erkundung der Potentialfläche durch eine hohe Anzahl von unabhängigen Walkern durchgeführt wird, wurden entwickelt und implementiert. Daraus resultiert ein verbessertes Kommunikationsschema zwischen diesen Walkern. Dadurch wird die Suche auf die vielversprechendsten Bereiche der Potentialfläche konzentriert wird.

Die Minima Hopping Methode verwendet kurze Molekulardynamik Simulationen bei erhöhten Temperaturen, um die Strukturaufklärung von einem lokalen Minimum der Potentialfläche zum nächsten zu leiten. Eine Modifikation, bei der die lokalen Informationen um jedes Minimum extrahiert und in einer Optimierung der Suchrichtung verwendet werden, ist entwickelt und implementiert worden. Unsere Verbesserung verwendet diese lokalen Informationen, um die Wahrscheinlichkeit zu erhöhen neue, energetisch niedrigere, lokale Minima zu finden. Dies führt zu einer verbesserten Performance des globalen Optimierungsalgorithmus.

Wasserstoff ist ein hoch relevantes System, da die Möglichkeit besteht, dass unter hohem Druck eine metallische Phase und Supraleitung mit einer hohen Sprungtemperatur gefunden werden kann. Eine Anwendung eines Strukturvorhersage-Algorithmus auf SiH_{12} findet stabile Kristallstrukturen in diesem Material. Zusätzlich finden wir eine Metallisierung bei verhältnismäßig niedrigen Drücken.

Contents

1 Introduction

The structure of a solid, as defined by the order of the atoms in the crystal structure, hugely influences its chemical and physical properties. Systems composed from the same types of atoms may exhibit vastly different characteristics, depending on their structure. This polymorphism is for example found in carbon, where many diverse allotropes are known. Diamond, graphite, fullerene and graphene all are only composed from elemental carbon, and still differ greatly in many respects.

In chemistry and material science, where the prediction of properties of unknown materials is one of the most fundamental goals, the stable crystal structures is therefore one of the most important features one can aim to predict. This structure is the energetical minimum of the system depending on external conditions like temperature and pressure. These conditions have therefore to be considered in the simulations, which poses additional challenges.

With the discovery of X-rays by Wilhelm Conrad Röntgen in 1895 and the development of X-ray diffraction by Max von Laue and William Lawrence Bragg it became possible to probe the three-dimensional structure of crystalline solids. This method allows us to extract valuable information about systems relatively easily. The structure determination via X-ray diffraction is a routinely applied technique and provides good results. Yet the theoretical prediction of the crystal structure of a system based on the chemical composition alone is still very challenging. This fact was already addressed and called a scandal by John Maddox in 1988 [59].

> "One of the continuing scandals in the physical sciences is that it remains in general impossible to predict the structure of even the simplest crystalline solids from a knowledge of their chemical composition."

This difficulty arises from the exponential nature of the problem. The number of possible solutions increases vastly with the increased number of parameters, which makes a direct solution only feasible for the smallest of systems. To overcome this issue, global optimization is employed. General examples of problems where global optimization is used are the travelling salesman problem, protein folding, the here mentioned structure prediction, just to name a few.

Many different approaches to these global optimization problems have been proposed over the years. Even though huge advances have been achieved in this area, these methods are still computationally expensive, and hence there is still room for improvement [97, 101].

The difficulty of global optimization arises from the fact that generally many local minima are present. The discovery of these minima via local minimization is comparatively simple. Unfortunately there is no straightforward way of determining if a local minimum is also the global one. Additionally, there is no information available where this global minimum is located in the search space. To predict the stable structure of a solid the whole potential energy space has to be explored to make sure the global minimum has been found. Local minima of the potential energy surface represent metastable structures, while the global minimum is the stable ground state of the investigated system. This energy surface is a function of the potential energy of the system with respect to structural parameters like atomic positions and the unit cell parameters. Consequently it is high dimensional and generally very complex. An example of such a surface is illustrated in Figure 1.1, here one can see that the surface is usually very uneven and many (local) minima are present.

The main motivation for this work are predictions for experimentally inaccessible structures, in particular under high pressure. Diamond anvil cells, as depicted in Figure 1.2, allow us to reach pressures up to 400 GPa and thus fundamentally change the properties of most materials. However, many questions cannot be answered solely by such experiments. For example the application of X-ray diffraction, which is the standard method for crystal structure determination, is highly non-trivial for hydrogen-rich compounds, due to the small diffraction coefficient of the light hydrogen atoms. Simulations allow us to investigate materials at pressures which are experimentally not accessible and help

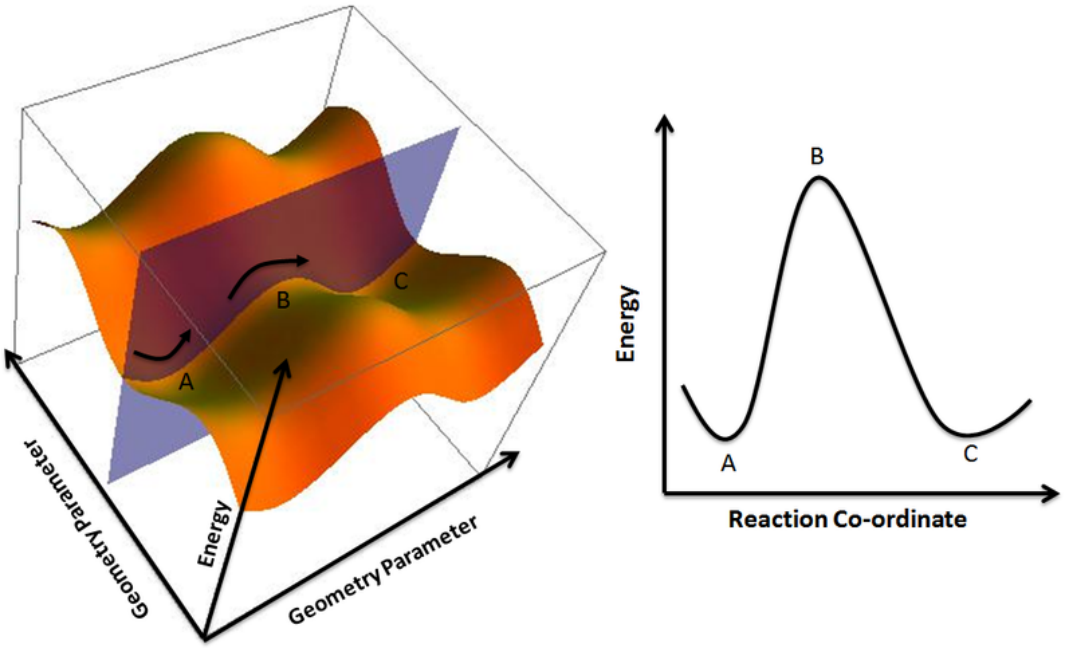

Figure 1.1: Schematic sketch of a potential energy surface [1]

solve the questions where experiments are not sufficient. Theoretical predictions can also provide excellent prospects for new experiments, which makes this work reside in the intersection between theoretical solid state chemistry and experimental high pressure physics.

1.1 Hydrogen Under Pressure

We are interested in hydrogen and hydrogen-rich systems. Hydrogen has been predicted by Wigner and Huntington in 1935 [99] to become an atomic solid that is additionally a metallic conductor under very high pressure. The existence of metallic hydrogen has implications for astrophysics, because gas giants, like Jupiter and Saturn, are mainly made up from light elements like hydrogen and helium, yet possess some of the strongest planetary magnetic fields in the solar system. Proving the existence of a metallic hydrogen core may provide a convincing explanation for this extraordinarily high magnetic field [6, 79, 80]. A sketch of the planetary model of Jupiter is shown in Figure 1.3. Additionally the possibility exists that metallic hydrogen is a high-T_c superconductor [8, 78]

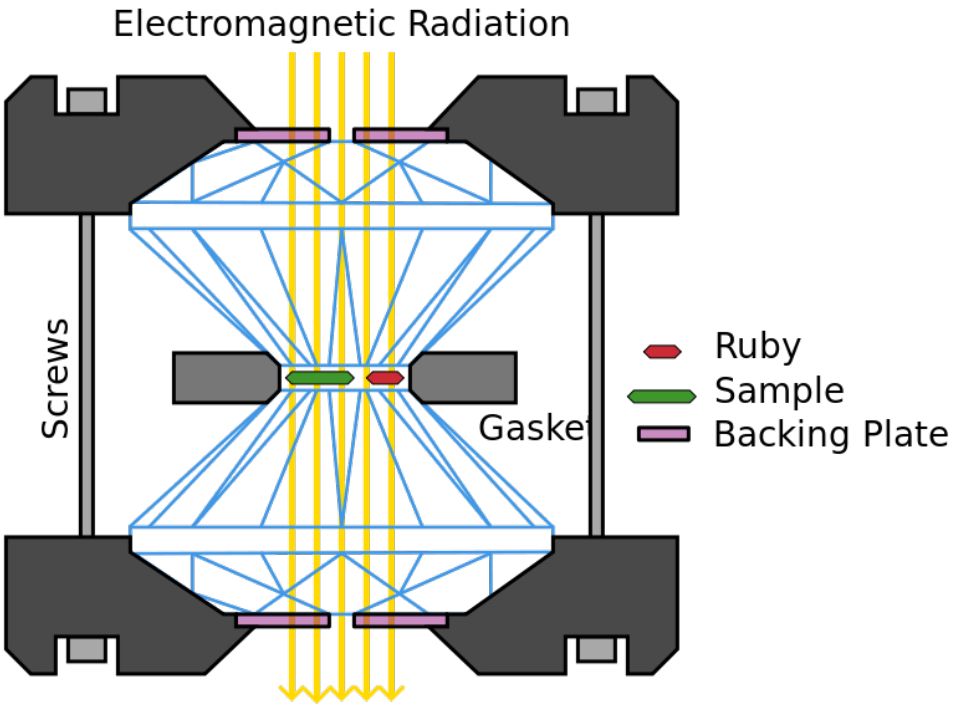

Figure 1.2: Cross section of a diamond anvil cell. These Cells are used to reach very high pressures in experiments. The sample is compressed between two opposing diamonds [2].

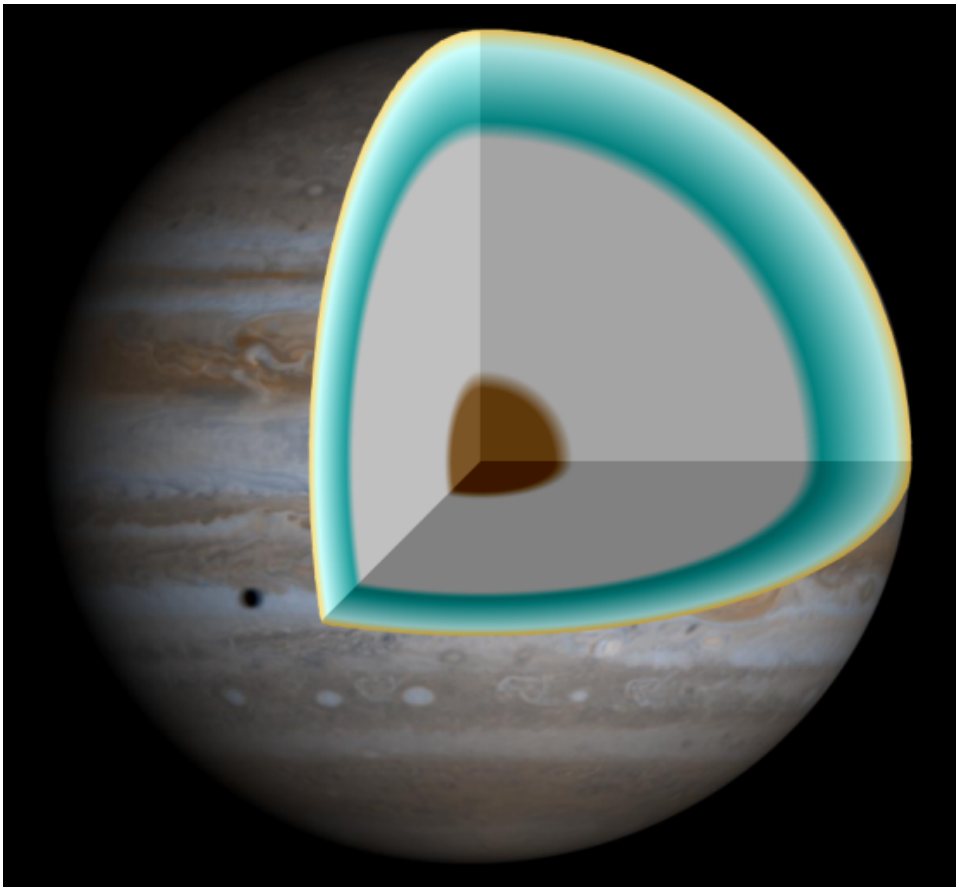

Figure 1.3: Model of the interior of Jupiter. A core of heavy elements is surrounded by a layer of metallic hydrogen. Further outwards liquid hydrogen is found. [3]

and may posses a metallic liquid ground state [14]. This makes the theoretical examination of the phase diagram of hydrogen under high pressures an important field of research.

The difficulty of reaching these high pressures make it hard to verify predictions and the previously mentioned problems of performing X-ray diffraction on hydrogen complicates the determination of the structure of experimentally attained high pressure phases. In 2011, Eremets and Troyan reported the discovery of metallic hydrogen at room temperature and a pressure of 300 GPa [30].

It is still not possible to reach the needed pressures of above 400 GPa in experiments to dissociate hydrogen, therefore other ways of attaining metallic states have been pro-

posed. In hydrogen rich compounds like silane a metallic state at a comparatively low pressure has previously been found [29]. This effect can be explained via "chemical pre-compression". The presence of the silicon atoms in the system vastly decreases the distance between the hydrogen atoms. To reach the same distance between two atoms in elemental hydrogen a way higher pressure is required. We are investigating hydrogen rich compounds with lower silicon contents in order to try to find similar effects.

This work is structured as follows: After an introduction to the applied computational methods in chapter 2 an overview over the current state of crystal structure prediction is given in chapter 3. In chapter 4 and 5 developed improvements to two different global optimization methods are presented. Finally, in chapter 6 the electronic structure of SiH_{12} is calculated under pressure to determine its metallization pressure.

2 Computational Methods

An important basis to carry out the global optimizations (GO) are methods of calculating or approximating the potential energy surface (PES). The accuracy at which this is done determines the efficiency and reliability of the performed GO. In this chapter some classical potentials to approximate the PES are presented. Additionally, density functional theory (DFT), one of the most used electronic structure methods to calculate the PES from first principles, is introduced [53]. An introduction to molecular dynamics is given at the end of this chapter due to the fact that it is frequently used to propagate the investigated system from one local minimum to the next [33].

2.1 Force Fields and Classical Potentials

Classical potentials are an easy and fast way to calculate the total energy of a system. These potentials are a mathematical description of the forces acting between the particles in the system, and provide a relatively simple functional form from which the potential energy of the system can be calculated. Potentials have to be generated for each special problem or system one wants to investigate. Different types of atoms interact via different forces with each other. Additionally, some potentials are optimized to represent a certain aspect as accurate as possible. One such aspect is the geometry and potentials which are designed to optimally represent the geometry should be preferred for GO.

2.1.1 Lennard-Jones Cluster

Systems of particles interacting via a Lennard-Jones (LJ) type potential, which is one of the easiest and most basic interaction potentials, are called Lennard-Jones cluster. Due to the fact that its particles are solely interacting via a classical potential the total energy and derivatives are easily available. Therefore, extensive databases of energy values and geometries of the global minimum are available for a wide range of cluster sizes.

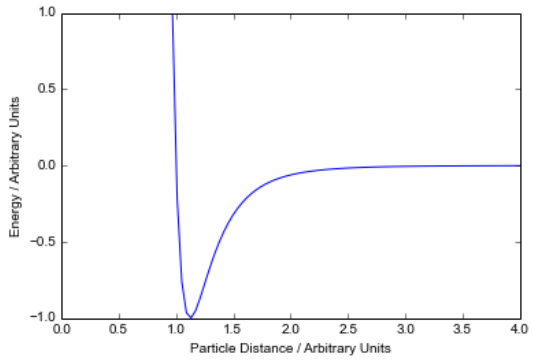

Figure 2.1: Lennard-Jones Potential, $\epsilon = \sigma = 1$

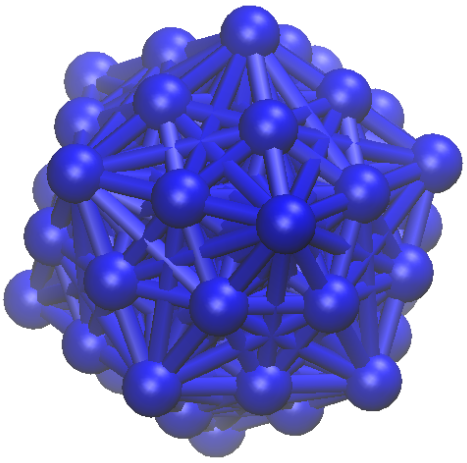

Figure 2.2: Lennard-Jones-cluster 55 particles

The Lennard-Jones interaction potential between two particles is of the form

$$V_{ij}(r_{ij}) = 4\epsilon \left[\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^{6} \right]. \tag{2.1}$$

Here, r_{ij} denotes the distance between the two particles, denoted with i and j. ϵ and σ are parameters that can be freely chosen and define the energy ϵ of a pair of particles in the equilibrium distance $2^{1/6}\sigma$. However, in order to get results comparable with literature, values of $\epsilon = 1$ and $\sigma = 1$ are selected. LJ cluster are well studied systems and the global minimum is known for clusters up to a few hundred particles, and are therefore a perfect system to test novel search algorithms. We performed global optimizations for different cluster sizes in order to evaluate the performance of the improved search algorithm. A

plot of the potential can be seen in Figure 2.1, and an example of a LJ cluster is shown in Figure 2.2.

2.1.2 Morse Cluster

Morse cluster are very similar to LJ cluster because the energy of the system is also given via a purely distance based potential. It has the form

$$E(r_{ij}) = \epsilon \, e^{\phi(1-r_{ij}/r_0)}(e^{\phi(1-r_{ij}/r_0)} - 2). \tag{2.2}$$

The value ϵ represents the depth of the equilibrium well and r_0 is the distance at which it can be found. Both parameters have been chosen as 1 in this work. The steepness and the width of the potential around the minimum can be influenced via the parameter ϕ. A plot of the Morse potential is shown in Figure 2.3 for the case $\phi = 6$. This case equates to a potential very similar to the LJ one.

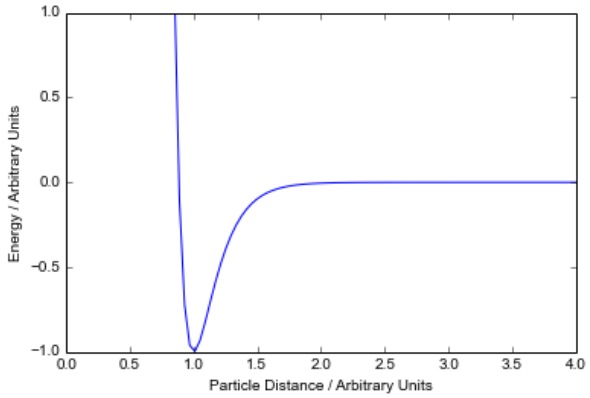

Figure 2.3: Morse Potential $\epsilon = r_0 = 1$, $\phi = 6$

2.1.3 Water Cluster

Clusters of water molecules are more difficult to simulate, compared to the previously presented systems, due to the facts that these consist of two different types of atoms and that an inner structure is present in the molecule.

2.1.3.1 The TIP4P/2005 Water Model

The TIP4P/2005 water model uses a pairwise additive intermolecular potential

$$E_{inter} = \sum_i \sum_{j>i} \left(4\epsilon \left[\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^{6} \right] + \sum_{m \in i} \sum_{n \in j} \frac{q_m q_n}{r_{mn}} \right) \tag{2.3}$$

to simulate the interaction between each pair of water molecules [4, 40]. The two interacting molecules are identified as i and j. It consists of a LJ like potential between the oxygen atoms, with r_{ij} being the distance between these. The values for the parameters ϵ and σ are chosen to approximate the behaviour of water in an optimal way and are given in Table 2.1. The interaction between the partial charge sites in molecule i and j is modelled using the Coulomb interaction between the charge sites. The distance between these sites is given as r_{mn}. The indexes m and n iterate over all charge sites on molecule i and j respectively. Three charge sites are present in each molecule, two of these are located at the positions of the hydrogen atoms and are of the magnitude $\frac{q_M}{2}$. Instead of the two lone electron pairs only one complementing negative charge of the magnitude $-q_M$ is defined at position r_m, which is located on the line connecting the oxygen to the center of mass of the hydrogen atoms. This position can be calculated as

$$r_m = \gamma r_{\mathrm{O}} + (1-\gamma)(r_{\mathrm{H}_1} + r_{\mathrm{H}_2})/2. \tag{2.4}$$

It depends on the position of the atoms $r_{\mathrm{O}}, r_{\mathrm{H}_1}$ and r_{H_2} in the molecule, and a parameter γ, which is also given in Table 2.1.

Up to this point, the model describes the interaction of rigid water molecules. To add the intramolecular flexibility, a potential

$$E_{intra} = \sum_i \left[V_\theta(\theta_i) + V_{\mathrm{OH}}(r_{i1}) + V_{\mathrm{OH}}(r_{i2}) \right] \tag{2.5}$$

is added. The bond angle θ_i is described via a harmonic potential

$$V_\theta(\theta_i) = \frac{1}{2} k_\theta (\theta_i - \theta_{eq})^2, \tag{2.6}$$

depending on the equilibrium bond angle θ_{eq} and a parameter k_θ. A quartic expansion of a Morse potential is chosen for the simulation of the O-H bond stretch

$$V_{OH}(r) = D_r \left[\alpha_r^2 (r - r_{eq})^2 - \alpha_r^3 (r - r_{eq})^3 + \frac{7}{12} \alpha_r^4 (r - r_{eq})^4 \right]. \tag{2.7}$$

It depends on the equilibrium distance r_{eq} and two parameters D_r and α. The full Morse potential is not used in this model due to the fact that it would allow for dissociation to happen. The energy and structure for the global minimum for water clusters simulated using this potential are known and listed in the Cambridge Cluster Database [96].

Table 2.1: Parameter for the TIP4P/2005 water model[40]

Parameter	Value
ϵ	0.1852 kcal mol^{-1}
σ	3.1589 Å
q_M	1.1128 $\lvert e \rvert$
γ	0.73612
D_r	116.09 kcal mol^{-1}
α_r	2.287 Å^{-1}
r_{eq}	0.9419 Å
k_θ	87.58 kcal mol^{-1} rad^{-2}
θ_{eq}	107.4 °

2.1.3.2 The TIP5P Water Model

In TIP5P water, the two negative charges that are present in the water molecule are considered as individual points. These charges are located at the position of the lone electron pairs. The potential of this rigid water model is thus given as [60]

$$E = \frac{e^2}{4\pi\epsilon_0} \sum_{m \in i} \sum_{n \in j} \frac{q_m q_n}{r_{mn}} + 4\epsilon \left[\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^{6} \right]. \tag{2.8}$$

The used values for the parameters are given in Table 2.2. For the TIP5P water cluster, a local minimum with an energy only slightly higher than the global minimum can be found. These structures are shown in Figure 2.4.

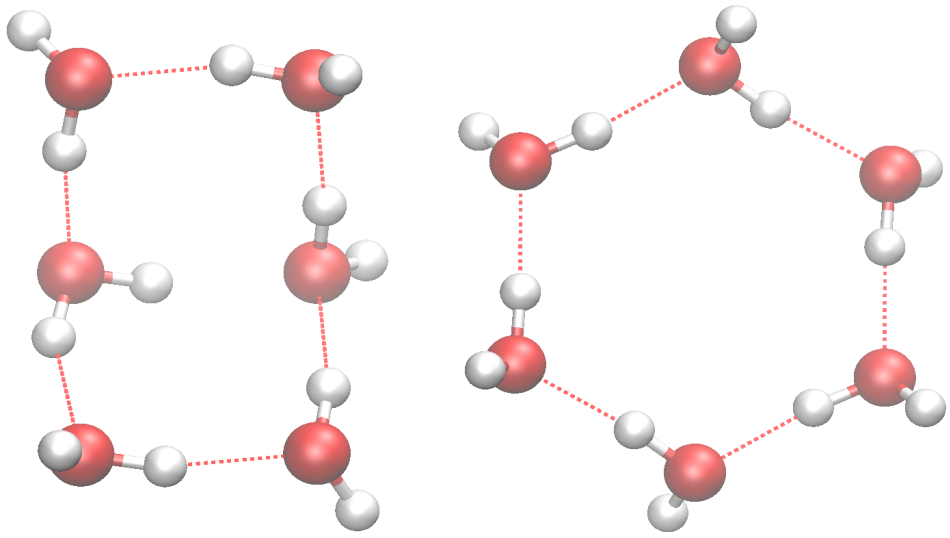

Figure 2.4: The two identified minima for the six rigid water molecules interacting via the TIP5P potential. The structure shown on the left has an energy of $E \approx -195.8\text{kJmol}^{-1}$, whereas the structure on the right is the global minimum known from the Cambridge Cluster Database [96]. Its energy is at $E \approx -197.9\text{kJmol}^{-1}$. [98]

Table 2.2: Parameter for the TIP5P water model[60]

Parameter	Value
ϵ	0.66944 kJ mol^{-1}
σ	3.12 Å
q_M	0.241 $\lvert e \rvert$

2.2 Density Functional Theory

An alternative approach is to calculate the potential using quantum chemical methods, for example density functional theory (DFT). Here the energy and the forces are calculated from first principles. This generally leads to more accurate results, but requests significantly larger computational work.

The starting point for DFT, like for most quantum chemical calculations, is the stationary Schrödinger equation [85] for a system of M nuclei and N electrons

$$\widehat{\mathrm{H}}\Psi_i(\mathbf{x}_1, \mathbf{x}_2, \cdots \mathbf{x}_N, \mathbf{R}_1, \mathbf{R}_2, \cdots \mathbf{R}_M) = E_i\Psi_i(\mathbf{x}_1, \mathbf{x}_2, \cdots \mathbf{x}_N, \mathbf{R}_1, \mathbf{R}_2, \cdots \mathbf{R}_M). \quad (2.9)$$

Here Ψ_i is the exact wavefunction of the i-th state of the system, Ψ_0 is the wavefunction of the ground state, which is supposed to be non degenerate. E_i is the systems total energy in this state. Ψ_i depends on the coordinates of the nuclei $\mathbf{R}_A$, as well as the spatial and spin coordinates ($\mathbf{x}_i \equiv (\mathbf{r}_i, \mathbf{s}_i)$) of the electrons. It contains all information about the system. The Hamiltonian in atomic units

$$\widehat{\mathrm{H}} = \underbrace{-\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2}_{\hat{\mathrm{T}}_e} \underbrace{-\frac{1}{2}\sum_{A=1}^{M}\frac{1}{m_A}\nabla_A^2}_{\hat{\mathrm{T}}_k} \underbrace{-\sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{\mathbf{r}_{iA}}}_{\hat{\mathrm{V}}_{ke}} + \underbrace{\sum_{i=1}^{N}\sum_{j<i}^{N}\frac{1}{\mathbf{r}_{ij}}}_{\hat{\mathrm{V}}_{ee}} + \underbrace{\sum_{A=1}^{M}\sum_{B<A}^{M}\frac{Z_A Z_B}{\mathbf{r}_{AB}}}_{\hat{\mathrm{V}}_{kk}} \tag{2.10}$$

is the operator corresponding to the total energy of the system. The indices A and B denote the M nuclei of the system, i and j identify the N electrons. The Hamiltonian depends on the mass of each nucleus m_A and its electric charge Z_A, and consists of kinetic energy contributions $\hat{\mathrm{T}}$ as well as potential energy contributions $\hat{\mathrm{V}}$.

Due to the fact that electrons are at least 2000 times lighter than the nuclei, and move therefore much faster, one can treat the dynamics of these two separately. Fixed positions for the nuclei can be assumed and an electronic Hamiltonian is defined as

$$\widehat{\mathrm{H}}_{el} = \hat{\mathrm{T}}_e + \hat{\mathrm{V}}_{ee} + \hat{\mathrm{V}}_{ke}. \tag{2.11}$$

This so-called Born-Oppenheimer approximation [15] leads to a parametric dependence of the energy on the positions of the nuclei: $E_{el}(\mathbf{r}; \mathbf{R})$. The position of the nuclei is thus no longer a variable in the calculation, and for each given configuration exactly one energy value can be computed.

To solve the Schrödinger equation, one employs the variational principle. It states that the expectation value of the Hamiltonian for a given trial wave function is always greater or equal to the expectation value for the true ground state wave function. It thus serves as an upper bond for the exact energy. A trial wave function Ψ_{trial} is chosen and the functional $E[\Psi]$ is minimized using the variational principle

$$\left\langle \Psi_{\text{trial}} \middle| \widehat{\mathrm{H}} \middle| \Psi_{\text{trial}} \right\rangle = E_{\text{trial}} \geq E_0 = \left\langle \Psi_0 \middle| \widehat{\mathrm{H}} \middle| \Psi_0 \right\rangle \tag{2.12}$$

to get the best possible approximation to the real wave function.

The Electron Density

The central property of density functional theory (DFT) is the local electron density $\rho(\mathbf{r})$, which is the integral of the many-body wave function over all spin coordinates and all but one spatial coordinate of all electrons

$$\rho(\mathbf{r}) = N \int \cdots \int |\Psi(\mathbf{x}_1, \mathbf{x}_2, \cdots \mathbf{x}_N)|^2 \, \mathrm{d}\mathbf{x}_2 \cdots \mathrm{d}\mathbf{x}_N. \tag{2.13}$$

This density denotes the probability of finding any electron with arbitrary spin in the volume element $\mathrm{d}\mathbf{r}^3$. The electron density has to vanish at infinity ($\rho(\mathbf{r} \to \infty) = 0$) and the integral over the density yields the total number of electrons ($\int \rho(\mathbf{r})\mathrm{d}\mathbf{r} = N$) in the system.

2.2.0.3 First Hohenberg-Kohn Theorem

In 1964 Hohenberg and Kohn [45] stated two fundamental theorems, which are the foundations on which density functional theory resides. The first theorem addresses the external potential V_{ext}, in which the observed electrons move.

> The external potential $V_{\mathrm{ext}}(\mathbf{r})$ is (to within a constant) a unique functional of $\rho(\mathbf{r})$; since, in turn $V_{\mathrm{ext}}(\mathbf{r})$ fixes $\widehat{\mathrm{H}}$ we see that the whole many-particle ground state is a unique functional of $\rho(\mathbf{r})$.

This states that there cannot be different potentials V_{ext} that lead to the same ground state electron density ρ_0, hence every ground state electron density belongs to exactly one external potential. From this information exactly one Hamiltonian can be set up. This leads to the ground state wave function of the system.

$$\rho_0 \to \{N, Z_A, R_A\} \to \widehat{\mathrm{H}} \to \Psi_0 \Rightarrow E_0 \text{ (and all other properties).} \tag{2.14}$$

The functional for the energy $E_0[\rho_0]$ has the following form:

$$E_0\left[\rho_0\right] = T\left[\rho_0\right] + E_{ee}\left[\rho_0\right] + E_{Ne}\left[\rho_0\right]. \tag{2.15}$$

Only the potential energy due to nuclei-electron attraction $E_{Ne}\left[\rho_0\right] = \int \rho_0(\mathbf{r}) V_{Ne} \mathrm{d}\mathbf{r}$ is system-dependent, while the kinetic energy $T[\rho_0]$ and the potential energy due to

electron-electron repulsion $E_{ee}[\rho_0]$ do not depend on the system and can therefore be collected in the system-independent Hohenberg-Kohn functional $F_{HK}[\rho_0]$. This gives

$$E_0[\rho_0] = \int \rho_0(\mathbf{r}) V_{Ne} \mathrm{d}\mathbf{r} + F_{HK}[\rho_0] \tag{2.16}$$

with the unknown Hohenberg-Kohn functional

$$F_{HK}[\rho] = T[\rho] + E_{ee}[\rho] = \left\langle \Psi \middle| \hat{T} + \hat{V}_{ee} \middle| \Psi \right\rangle, \tag{2.17}$$

which is the core of density functional theory. The functional $F_{HK}[\rho]$ contains the kinetic energy $T[\rho]$ and the electron-electron interaction $E_{ee}[\rho]$. At least one part of $E_{ee}[\rho]$, the classical Coulomb interaction $J[\rho]$ is well known

$$E_{ee}[\rho] = \frac{1}{2} \int\int \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{r_{12}} \mathrm{d}\mathbf{r}_1 \mathrm{d}\mathbf{r}_2 + E_{\mathrm{ncl}}[\rho] = J[\rho] + E_{\mathrm{ncl}}[\rho]. \tag{2.18}$$

The non classical electron interaction $E_{\mathrm{ncl}}[\rho]$, consisting of the self interaction correction as well as the exchange-correlation, is unknown. Unfortunately functionals for these properties are not known, and therefore approximations are needed.

Finding explicit expressions for $E_{\mathrm{ncl}}[\rho]$ and $T[\rho]$ is one of the main challenges in density functional theory.

2.2.0.4 Second Hohenberg-Kohn Theorem

The second Hohenberg-Kohn theorem states

> The functional $F_{HK}[\rho]$ gives the lowest energy if the input is the real ground state electron density.

This resembles the variation principle (see section 2.2) From equation (2.15) an energy is obtained for every trial electron density $\tilde{\rho}$. This $\tilde{\rho}$ has to fulfill some boundary conditions ($\tilde{\rho}(\mathbf{r}) \geq 0$, $\int \tilde{\rho}(\mathbf{r})\mathrm{d}\mathbf{r} = N$) and is attributed to some external potential $\tilde{V}_{\mathrm{ext}}$. The obtained energy is an upper boundary for the real ground state energy E_0[1].

$$E_0 \leq E[\tilde{\rho}] = T[\tilde{\rho}] + E_{Ne}[\tilde{\rho}] + E_{ee}[\tilde{\rho}]. \tag{2.19}$$

[1] This variational ansatz is only valid if the exact F_{HK} is used. Since this functional is unknown, density functional theory can yield energies lower than E_0.

Only with the real ground state electron density ρ_0 the real ground state energy E_0 can be obtained [51].

2.2.1 The Kohn-Sham Approach

In 1965, Kohn and Sham presented a way to put the Hohenberg-Kohn theorems to work and suggested approaches to find explicit forms for the functionals $T[\rho(\mathbf{r})]$ and $E_{\mathrm{ncl}}[\rho(\mathbf{r})]$ [54]. They approximated the true kinetic energy using the kinetic energy of a reference system of non-interacting electrons which move in an external effective potential.

2.2.1.1 The Non-Interacting Reference System

Kohn and Sham considered, like in Hartree-Fock theory [93], a Slater determinant Φ_{SD} (see equation (2.21)) consisting of N spin-orbitals instead of the real wave function. The spin-orbitals χ_i in Φ_{SD} are chosen in a way that E is minimal.

This is the exact wave function of a system of non-interacting fermions and therefore one may build a non-interacting reference system and a Hamiltonian which contains an effective, local potential $V_S(\mathbf{r})$

$$\widehat{\mathrm{H}}_S = -\frac{1}{2}\sum_i^N \nabla_i^2 + \sum_i^N V_S(\mathbf{r}_i). \tag{2.20}$$

This Hamiltonian contains electron-electron interactions only in a mean-field fashion, and has a Slater determinant as ground state wave function

$$\Phi_S = \frac{1}{\sqrt{N!}}\begin{vmatrix} \varphi_1(\mathbf{x}_1) & \varphi_2(\mathbf{x}_1) & \cdots & \varphi_N(\mathbf{x}_1) \\ \varphi_1(\mathbf{x}_2) & \varphi_2(\mathbf{x}_2) & \cdots & \varphi_N(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_1(\mathbf{x}_N) & \varphi_2(\mathbf{x}_N) & \cdots & \varphi_N(\mathbf{x}_N) \end{vmatrix}. \tag{2.21}$$

The one-electron equations

$$\hat{\mathrm{f}}^{KS}\varphi_i = \varepsilon_i\varphi_i \tag{2.22}$$

may now be solved with the one-electron Kohn-Sham operator

$$\hat{\mathrm{f}}^{KS} = -\frac{1}{2}\nabla^2 + V_S(\mathbf{r}). \tag{2.23}$$

The effective potential $V_S(\mathbf{r})$ has to be introduced into this operator to establish a connection between the non-interacting reference system and the real system. It has to be chosen in a way that the ground-state density of the non-interacting system equals to the ground state electron density of the real system

$$\rho_S(\mathbf{r}) = \sum_{i}^{N(Occ)} \sum_{s} |\varphi_i(\mathbf{r}, s)|^2 = \rho_0(\mathbf{r}). \tag{2.24}$$

The kinetic energy of this system is given by

$$T_S = -\frac{1}{2}\sum_{i}^{N} \left\langle \varphi_i \left| \nabla^2 \right| \varphi_i \right\rangle. \tag{2.25}$$

This energy T_S is smaller than the true kinetic energy, and therefore one has to find an appropriate term for the small remainder. With this new term the functional $F_{HK}[\rho]$ (see equation (2.17)) is written as

$$F_{HK}[\rho] = T_S[\rho] + J[\rho] + E_{XC}[\rho], \tag{2.26}$$

with the exchange-correlation energy functional $E_{XC}[\rho]$. This functional is the part of $F_{HK}[\rho]$ which still remains unknown. It is obtained as the difference between the kinetic energy of the real system and the kinetic energy of the non-interacting reference system plus the difference between the electron-electron interaction and the classical Coulomb interaction

$$E_{XC}[\rho(\mathbf{r})] \equiv (T[\rho] - T_S[\rho]) + (E_{ee}[\rho] - J[\rho]) = T_C[\rho] + E_{\mathrm{ncl}}[\rho]. \tag{2.27}$$

The exchange-correlation energy functional $E_{XC}[\rho]$ contains all parts of the energy functional which are unknown, and all many-body effects.

To determine the orbitals of the non-interacting reference system and to get a Slater determinant which yields the electron density of the real system, one obtains from the variational principle

$$\left(-\frac{1}{2}\nabla^2 + \left[\int \frac{\rho(\mathbf{r}_2)}{r_{12}}\mathrm{d}\mathbf{r}_2 + V_{XC}(\mathbf{r}_1) - \sum_{A}^{M}\frac{Z_A}{r_{1A}}\right]\right)\varphi_i = \left(-\frac{1}{2}\nabla^2 + V_{\mathrm{eff}}(\mathbf{r}_1)\right)\varphi_i = \varepsilon_i\varphi_i.$$

(2.28)

Comparison of this equation with equation (2.23), shows that V_{eff} equals V_S

$$V_S(\mathbf{r}) \equiv V_{\text{eff}}(\mathbf{r}) = \int \frac{\rho(\mathbf{r}_2)}{r_{12}} \mathrm{d}\mathbf{r}_2 + V_{XC}(\mathbf{r}_1) - \sum_A^M \frac{Z_A}{r_{1A}}. \tag{2.29}$$

If one has the exact exchange-correlation potential to insert into equation (2.29) one could solve the one-electron equations to get the orbitals and the ground state electron density and energy etc. Unfortunately V_{XC} is defined via the unknown functional E_{XC} as

$$V_{XC} \equiv \frac{\delta E_{XC}}{\delta \rho} \tag{2.30}$$

and is therefore unknown too. Approximations for E_{XC} will be given in the next subsection.

2.2.2 Approximations for the Exchange-Correlation Functional

Since all the remaining, unknown parts of the functional are collected in $E_{XC}[\rho]$, and no explicit expression for this functional is known, one needs to develop certain approximations to solve the Kohn-Sham equations.

2.2.2.1 Local (Spin) Density Approximation

In the local density approximation (LDA) one takes the exchange and correlation energy of a hypothetical uniform electron gas. Therefore $E_{XC}[\rho]$ is written in the following form

$$E_{XC}^{LDA}[\rho] = \int \rho(\mathbf{r}) \varepsilon_{XC}^{LDA}(\rho(\mathbf{r})) \mathrm{d}\mathbf{r}. \tag{2.31}$$

The quantity $\varepsilon_{XC}(\rho(\mathbf{r}))$, which is the exchange and correlation energy of a particle in a homogeneous electron gas of the density ρ, is weighted with the probability of finding

an electron at this location. The exchange and correlation parts of $\varepsilon_{XC}(\rho(\mathbf{r}))$ are viewed separately

$$\varepsilon_{XC}(\rho(\mathbf{r})) = \varepsilon_X(\rho(\mathbf{r})) + \varepsilon_C(\rho(\mathbf{r})). \tag{2.32}$$

For the exchange the so-called Slater exchange, which was derived by Bloch and Dirac in the late 1920's [13, 26, 87], is assumed

$$\varepsilon_X = -\frac{3}{4}\left(\frac{3\rho(\mathbf{r})}{\pi}\right)^{\frac{1}{3}}. \tag{2.33}$$

Unfortunately there are no closed-form expressions for the correlation part $\varepsilon_C(\rho(\mathbf{r}))$ but good numerical expressions can be derived by Monte-Carlo simulations for the electron gas [18].

2.2.2.2 Generalized Gradient Approximation

The generalized gradient approximation (GGA) is a logical next step beyond the LDA. Instead of taking just the value of the electron density at one position, the gradient of the density $\nabla\rho(\mathbf{r})$ is also considered. The exchange-correlation function E_{XC} is written as

$$E_{XC}^{GGA}[\rho] = \int \rho(\mathbf{r})\varepsilon_{XC}^{GGA}(\rho(\mathbf{r})\nabla\rho(\mathbf{r}))\mathrm{d}\mathbf{r}. \tag{2.34}$$

For cases with slowly varying electron density this generally improves the quality of the calculations. Problems may arise for systems with faster varying density, but modern functionals can take care of these problems [73].

2.2.3 Solving the Kohn-Sham Equations

The general procedure for solving the Kohn-Sham equations is the self consistent field (SCF) method. It is an iterative process which is necessary due to the fact that the result of this calculations (the electron density) is needed to generate the effective potential which is needed as an input in the Kohn-Sham equations. Some kind of basis set is

used to construct the orbitals as a superposition of atomic orbitals. The electron density is calculated from this. The resulting effective potential is afterwards used to solve the one particle Kohn-Sham equations. In the last step, a new density is calculated from the results and the Kohn-Sham equations are solved again with this new potential. This procedure has to be repeated until convergence in the density or total energy is achieved.

2.3 Molecular Dynamics

Molecular Dynamics (MD) is one of the most powerful methods for molecular simulations. It is used to evaluate expectation values for macroscopic thermodynamic properties of the system at finite temperature. The movement of the atoms in a given system is calculated iterativeley by solving Newton's equations of motion. The forces acting on the atoms can be calculated using the already presented methods of either classical potentials or ab-initio methods like DFT.

A number of global optimization methods rely on MD for the propagation of the system from one local minimum to the next. Here, one is not interested in the calculation of macroscopic properties of the system, instead only changes in the geometry on short time scales are relevant. Due to its nature, MD is an efficient tool for the discovery of new trial structures. Its inherent advantages are the preference of low lying structures and the favoured crossing of low lying energy barriers. A short introduction to MD is given here.

In MD, one tries to solve the classical equations of motion for the particles in the investigated system. For a simple atomic system of N particles with the coordinates $R = (R_1, R_2, \ldots, R_N)$ these equations are:

$$F_I = M_I \ddot{R}_I. \tag{2.35}$$

The force F_I is given as the derivative of the potential energy $U(R)$

$$F_I = -\frac{\partial}{\partial R_I} U. \tag{2.36}$$

The forces which affect the particles in the system have therefore to be calculated. In classical MD, this calculation is done using force fields, which describe the interactions between the atoms, whereas in ab-initio MD the forces are generated from electronic structure calculations. The momenta of the simple atomic system, which we consider here are defined as $P = (P_1, P_2, \ldots, P_N)$ and can be used to write the kinetic energy as

$$E_{\text{kin}} = \frac{1}{2}\sum_{I=1}^{N} M_I \dot{R}_I^2 = \sum_{I=1}^{N} \frac{P_I^2}{2M_I}. \tag{2.37}$$

In the Hamilton formalism, the equations of motion are a system of coupled differential equations

$$\ddot{R}_I = \frac{\dot{P}_I}{M_I} \qquad \text{and} \qquad \dot{P}_I = F_I. \tag{2.38}$$

This system can be transformed into a difference equation. To do this, R_I is expanded in a Taylor expansion as

$$R_I(t+\Delta t) = R_I(t) + \dot{R}_I(t)\Delta t + \frac{F_I(t)}{2M_I}(\Delta t)^2 + \frac{\partial^3 R_I(t)}{3!\partial t^3}(\Delta t)^3 + O((\Delta t)^4), \tag{2.39}$$

and

$$R_I(t-\Delta t) = R_I(t) - \dot{R}_I(t)\Delta t + \frac{F_I(t)}{2M_I}(\Delta t)^2 - \frac{\partial^3 R_I(t)}{3!\partial t^3}(\Delta t)^3 + O((\Delta t)^4). \tag{2.40}$$

The difference between these two equations is

$$R_I(t+\Delta t) - R_I(t-\Delta t) = 2\dot{R}_I(t)\Delta t + O((\Delta t)^3), \tag{2.41}$$

which leads us to

$$\dot{R}_I(t) = \frac{R_I(t+\Delta t) - R_I(t-\Delta t)}{2\Delta t} + O((\Delta t)^3). \tag{2.42}$$

This equation is the so-called Verlet [94] algorithm. A more often used approach is the velocity Verlet algorithm [92] in which R_I and $\dot{R}_I$ can be evaluated in the same iteration. At first the positions R_I at $t + \Delta t$ are evaluated

$$R_I(t+\Delta t) = R_I(t) + \dot{R}(t)\Delta t + \frac{F_I(t)}{2M_I}(\Delta t)^2. \tag{2.43}$$

In the next step the new forces F_I are calculated from the interaction potential as

$$F_I(t + \Delta t) = F_I \left[R_I(t + \Delta t) \right]. \tag{2.44}$$

To calculate the velocities $\ddot{R}$ we insert $t \to t + \Delta t$ into (2.40) to get

$$\ddot{R}_I(t) = \ddot{R}_I(t + \Delta t) - \ddot{R}_I(t + \Delta t)\Delta t + \frac{F_I(t + \Delta t)}{2M_I}(\Delta t)^2, \tag{2.45}$$

which is equivalent to

$$R_I(t + \Delta t) = R_I(t) + \frac{F_I(t) + F_I(t + \Delta t)}{2M_I}\Delta t. \tag{2.46}$$

This algorithm is now used to move the atoms for a certain time step Δt. A large number of repetitions of this procedure generates a progression of positions and velocities in the system. This is normally referred to as a trajectory. The size of the time-step in the MD calculation can not be arbitrarily large. It generally has to be rather small, for example the order of

$$\Delta t_{\max} \leq \frac{1}{20}\frac{1}{f_{\max}} \approx 10^{-15} s, \tag{2.47}$$

where $f_{\max}$ refers to the highest vibrational frequency in the system.

The velocity Verlet algorithm is the de facto standard for MD simulations. The calculation of exact trajectories is generally impossible, but one is generally interested in system properties which are obtained from statistical mean values, where the exact movements of the particles are not relevant. In GO algorithms there is no interest in these properties at all, instead one is only interested in the propagation of the particles to push the system towards a new, previously unknown structure. Therefore MD in GO is often initialized with rather high initial velocities and stopped as soon as a new structure is found, long before the system has reached the equilibrium.

3 Structure Prediction

In this chapter we address the question how the prediction of stable structures can be performed. This is an important task in chemistry, as the properties of the investigated material depend on the stable structure, which is given by the global minimum of the PES.

Structure prediction can be performed on a local or global level. Local techniques minimize the energy of the system in order to discover the nearest local minimum on the PES. Global methods, on the other hand, attempt to find the globally lowest minimum. This global minimum thus represents the stable structure of a material at the chosen external conditions. In this chapter local optimizations are introduced in the beginning, and global minimization methods are presented thereafter.

Many of the presented global methods rely on a local minimization techniques in parts of the algorithm. The GO only has to direct the system into the attraction region of a minimum, as the exact minimum can then be discovered via a local optimization. Furthermore, the PES can be transformed using local minimizations. By assigning the energy of the nearest minimum to every point on the PES, areas of attraction around each local minimum become areas of equal energy. This greatly reduces the complexity of the given problem.

The PES of relevant systems is typically very high dimensional, with an exponentially increasing number of local minima for increasing system sizes. This fact makes it rather difficult to find the global minimum, and unfortunately there is generally no guarantee that it has been found. Minima of lower energy may be available and separated by high energy barriers.

3.1 Local Minimizations

The most popular local optimization methods, namely steepest descent (SD), conjugate gradient (CG) and the quasi-Newton methods introduced by Broyden, Fletcher, Goldfarb and Shanno (BFGS and L-BFGS) [16, 32, 39, 69, 86] will be presented here.

3.1.1 The Steepest Descent Method

The goal of SD is to find the minimum of a quadratic test function

$$\phi(\mathbf{x}) = \frac{1}{2}\mathbf{A}\mathbf{x}^2 - b\mathbf{x}, \tag{3.1}$$

where $\mathbf{A}$ is a symmetric positive definite matrix. The goal of the minimization is to find an extremum (minimum) $\mathbf{x}^*$ of the function ϕ. Extrema of a function can be found as points where the gradient of the function is zero

$$0 = \nabla\phi(\mathbf{x}^*) = \mathbf{A}\mathbf{x}^* - b. \tag{3.2}$$

In the steepest descent method one follows the negative gradient of the function $-\nabla\phi(\mathbf{x})$ towards the minimum. This direction yields the vector of the direction in which the function decreases the most, and is generally followed until the function value decreases no longer. While moving down a valley towards the local minimum this method tends to perform many small steps. Each step is performed into a direction orthogonal to the previous one. The search thus moves in a rather inefficient zigzag fashion, until it reaches the minimum.

3.1.2 The Conjugate Gradient Method

A set of vectors $\{\mathbf{p}_0, \mathbf{p}_1, \ldots, \mathbf{p}_{n-1}\}$ is called conjugate with respect to a matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$ if

$$\mathbf{p}_i^T \mathbf{A}\mathbf{p}_j = 0 \text{ for all } i \neq j. \tag{3.3}$$

This set of vectors spans the complete space of $\mathbb{R}^n$. This allows us to minimize $\phi(\mathbf{x})$ in at most n steps by minimizing along all n directions of the individual vectors [44]. Additionally, we are able to express the difference vector between a guess $\mathbf{x}_0$ for the exact solution $\mathbf{x}^*$ as a linear combination of these vectors:

$$\mathbf{x}^* - \mathbf{x}_0 = \sigma_0 \mathbf{p}_0 + \sigma_1 \mathbf{p}_1 + \ldots + \sigma_{n-1} \mathbf{p}_{n-1} \tag{3.4}$$

The parameters σ_k correspond to the step length α_k needed to minimize a quadratic function ϕ along the axis $\mathbf{x}_k + \alpha_k \mathbf{p}_k$. Hence it follows for the search step

$$\mathbf{x}^* = \mathbf{x}_0 + \alpha_0 \mathbf{p}_0 + \alpha_1 \mathbf{p}_1 + \ldots + \alpha_{n-1} \mathbf{p}_{n-1}. \tag{3.5}$$

This method, which only works for the here assumed quadratic example, first finds the minimum along the first dimension of the solution space, than along the next, and so on until the minimum is found. The set of conjugate vectors is not known at the beginning and has to be generated. In the conjugate gradient method each direction is generated from the gradient $-r_k = \nabla\phi$ and the previous search direction as

$$\mathbf{p}_k = -r_k + \beta_k \mathbf{p}_{k-1}. \tag{3.6}$$

Here, β_k is generated from the condition (3.3) as

$$\beta_k = \frac{r_k \mathbf{A} p_{k-1}}{\mathbf{p}_{k-1}^T \mathbf{A} \mathbf{p}_{k-1}} = \frac{r_k^T r_k}{r_{k-1}^T r_{k-1}}. \tag{3.7}$$

Instead of solely choosing the gradient as the minimization direction, CG incorporates the previous gradients to optimize the search direction.

3.1.3 Quasi-Newton Methods

In the quasi-Newton minimization, instead of just using the previous gradient directions, an approximation of the Hessian matrix is generated and updated during the course of the minimization. This information can be used for a fast and reliable discovery of the local minimum. The approximate inverse $\mathbf{B}_i$ of the Hessian matrix $\mathbf{H}$ is therefore updated in every step of the algorithm, and converges to the true inverse

$$\lim_{i \to \infty} \mathbf{B}_i = \mathbf{H}^{-1}. \tag{3.8}$$

The search step, starting from the current point $\mathbf{x}_i$, is then given as

$$\Delta\mathbf{x} = \mathbf{x} - \mathbf{x}_i = -\mathbf{H}^{-1} \cdot \nabla f(\mathbf{x}_i), \tag{3.9}$$

where $\nabla f(\mathbf{x}_i)$ is the gradient of the to be minimized function $f(\mathbf{x}_i)$. Using the difference between the coordinates $\mathbf{x}_{i+1} - \mathbf{x}_i$ and the difference between the gradients $\nabla f(x_{i+1}) - \nabla f(\mathbf{x}_i)$, the inverse Hessian matrix $\mathbf{B}_i$ is then updated using the following formula[75]

$$\begin{aligned}\mathbf{B}_{i+1} = \mathbf{B}_i &+ \frac{(\mathbf{x}_{i+1} - \mathbf{x}_i) \otimes (\mathbf{x}_{i+1} - \mathbf{x}_i)}{(\mathbf{x}_{i+1} - \mathbf{x}_i) \cdot (\nabla f(\mathbf{x}_{i+1}) - \nabla f(\mathbf{x}_i))} \\ &- \frac{[\mathbf{B}_i \cdot (\nabla f(\mathbf{x}_{i+1}) - \nabla f(\mathbf{x}_i))] \otimes [\mathbf{B}_i \cdot (\nabla f(\mathbf{x}_{i+1}) - \nabla f(\mathbf{x}_i))]}{(\nabla f(\mathbf{x}_{i+1}) - \nabla f(\mathbf{x}_i)) \cdot \mathbf{B}_i \cdot (\nabla f(\mathbf{x}_{i+1}) - \nabla f(\mathbf{x}_i))} \\ &+ [(\nabla f(\mathbf{x}_{i+1}) - \nabla f(\mathbf{x}_i)) \cdot \mathbf{B}_i \cdot (\nabla f(\mathbf{x}_{i+1}) - \nabla f(\mathbf{x}_i))]\, \mathbf{u} \otimes \mathbf{u},\end{aligned} \tag{3.10}$$

where $\mathbf{u}$ is defined as

$$\begin{aligned}\mathbf{u} = &\frac{(\nabla \mathbf{x}_{i+1} - \mathbf{x}_i)}{(\mathbf{x}_{i+1} - \mathbf{x}_i) \cdot (f(\mathbf{x}_{i+1}) - \nabla f(\mathbf{x}_i))} \\ &- \frac{\mathbf{B}_i \cdot (\nabla f(\mathbf{x}_{i+1}) - \nabla f(\mathbf{x}_i))}{(\nabla f(\mathbf{x}_{i+1}) - \nabla f(\mathbf{x}_i)) \cdot \mathbf{B}_i \cdot (\nabla f(\mathbf{x}_{i+1}) - \nabla f(\mathbf{x}_i))}.\end{aligned} \tag{3.11}$$

In the low memory version of BFGS, called L-BFGS, only the vectors $(\mathbf{x}_{i+1} - \mathbf{x}_i)$ and $(\nabla f(\mathbf{x}_{i+1}) - \nabla f(\mathbf{x}_i))$ from the last number of steps are saved. L-BFGS uses these vectors instead of the complete Hessian matrix in the minimization. The Hessian matrix has not to be saved in this approach, which reduces the needed memory. This also decreases the utility of L-BFGS for our approach, because only a low rank version of the Hessian matrix is generated from the last few steps and it is not readily available at the end of the minimization. To get the approximation of the Hessian matrix from L-BFGS it has to be generated from $(\mathbf{x}_{i+1} - \mathbf{x}_i)$ and $(\nabla f(\mathbf{x}_{i+1}) - \nabla f(\mathbf{x}_i))$.

3.2 Methods for Global Optimization

Now we turn towards the question of how to perform a GO. A naive approach is to sample every point of a very fine-grained grid on the PES. This is not feasible for most

systems due to the exponential scaling of the complexity of the PES. More sophisticated approaches that only sample a limited number of structures have to be used.

The simplest approach is a search algorithm, where random structures are generated as starting points and optimized. This leads to a greatly reduced number of needed energy calculations compared to the completely random sampling. If one starts enough local optimizations at a sufficient number of random points, there is a reasonable high probability to even find the global minimum of the system. This is, of course, still a rather inefficient way to search the global minimum, because a high number of tries is still needed.

3.2.1 Basin Hopping

In basin hopping [27, 95, 97, 101] a transformed PES is explored using MC moves. This transformation is done using local optimizations, and the energy of the closest local minimum is assigned to each point on the PES. The acceptance of each suggested MC move is thus not dependant on the energy at the suggested destination, but instead on the energy of the nearest local minimum to this point. The complexity of the PES is reduced greatly by this approach.

3.2.2 Tabu Search

In tabu search [90], the system is moved out of the local minimum along the direction with the lowest ascent in energy, followed by a local optimization to identify the next minimum. If the system ends up in a structure which has already been visited a tabu is applied to the generation of the search direction to prevent it from following the exact same direction again.

3.2.3 Activation Relaxation Technique

A similar approach is the activation relaxation technique (ART) [61, 67], where a local search is performed in a way that the system is moved out of the current minimum across the nearest saddle point on the PES. From this point the configuration then is relaxed to the next minimum.

3.2.4 Metadynamics

In metadynamics, the goal of moving the system out of one minimum of the PES into the next is achieved via the addition of Gaussian potentials to the PES to slowly fill the current minimum [55, 63, 72]. These potentials push the system out of the current minimum towards the next one. In order to apply these potentials the dimensionality of the problem has to be reduced. Generally only the size and form of the unit cell are chosen as collective variables. The minimum energy structure of the atoms within the chosen cell can then be found using MD.

3.2.5 Simulated Annealing

Another popular example for search methods is simulated annealing (SA) [11, 24, 50, 64, 65, 71, 83, 101]. In this method a MC or MD simulation is performed. This method tries to emulate the process of physical annealing by starting with a high kinetic energy, so that the system is able to cross most barriers. The temperature is slowly decreased during the simulation, thus decreasing the probability of jumps between minima. With this set-up the simulation explores vast areas of the PES in the beginning and then slowly directs the system towards a minimum where it ends up once the temperature is low enough. Unfortunately there is no way to guarantee that this approach ends up in the global minimum, as local minima with a large attraction area are very likely to be found. Another problem in this approach is the speed with which the temperature is modified. The progression of this parameter is directly responsible for the success of the SA.

3.2.6 Temperature Accelerated MD

In temperature accelerated MD [88], a MD simulation of the system is carried out at an increased temperature to increase the probability of escaping the current minimum of the PES. The system is however constrained to the current minimum and only the escape probability for the different possible transition pathways are determined. They are then scaled back to the low temperature and the one with the highest probability is carried out.

3.2.7 Evolutionary Algorithm

The so-called genetic or evolutionary algorithms (EA) [17, 22, 23, 35, 38, 46, 58, 66, 68, 70] employ a rather different approach. In contrary to all previously presented methods EA does not explore the PES by moving the system from one minimum to another. Instead an approach inspired by evolution in populations of animals or humans is employed.

The method starts with a population of randomly generated structures. Each of these structures is optimized to the closest local minimum and they are then ranked according to a fitness criterion, which is generally the total energy at the minimum.

A new generation of trial structures is then generated via modifications and combinations of the trial structures with the best fitness. This process is repeated again and again. This leads to an accumulation of the best structures in the population until the global minimum is found.

The possibilities of generating these new structures "children" from the chosen "parents" are either based on mutation or child generation.

- *Mutation:* New structures are generated by modifying the lowest energy structures. For example the position of a few or all atoms in the cell can be altered, atom types can be exchanged between two positions or the complete cell can be deformed by stretching, compressing or shearing.

- *Child generation:* So called child structures can be generated by combining a number of structures from the previous generation. Generally one would take a slice from each of two parent structures and combine these to one new structure.

The new set of structures is then again optimized to find the best structures. The ground state structure is found with high probability after a number of repetitions of this procedure.

4 The Artificial Bee Colony Method

In this chapter the artificial bee colony (ABC) method is presented. This global optimization method is based on swarm intelligence and employs the intelligent communication of multiple random walkers which each individually explore the PES in an efficient way. An extension to this method is developed and tested with a number of model systems.

4.1 Introduction to ABC

ABC is inspired by the foraging behaviour of bees [48, 82, 98]. This method employs a number of walkers that are used to systematically explore the PES of the system. In the original ABC approach, three different types of walkers are defined

- *Employees* are located at the different known local minima of the PES and explore random search steps from their position, but only accept the new location if it is of lower energy compared to the old one.

- *Onlookers* explore the local surroundings of the employee walkers to help with the exploration of this area, namely accelerating the local optimization.

- *Scouts* move, completely unconditionally, to randomly chosen positions on the PES to find new local minima to explore.

This approach thus uses swarm intelligence to couple local with global searches and hereby prevents the system from getting trapped in these local minima. The energy of the system can be evaluated simultaneously at the different points where the walkers

reside, and an intelligent communication scheme between these walkers is used to guide the search.

4.2 Extensions of ABC

To evaluate the fitness of every discovered minimum on the PES, a fitness function $f(\mathbf{x})$ is needed. We define it as follows

$$f(E(\mathbf{x})) = \begin{cases} 1 - E(\mathbf{x}), & E(\mathbf{x}) < 0 \\ (1 + E(\mathbf{x}))^{-1}, & E(\mathbf{x}) \geq 0, \end{cases} . \tag{4.1}$$

Here $\mathbf{x}$ is a vector that consists of all structural parameters, like atomic positions, of the investigated system. The maxima of this function represent minima of the energy of the PES additionally it is non-negative and strictly monotonically decreasing with respect to the energy.

A set of walkers $\{A_i\}_{i=1}^{N_A}$ are now employed to perform the optimization. Each walker is assigned to a structural configuration $\mathbf{x}_i$ with the fitness f_i

$$A_i = (\mathbf{x}_i, f_i), \tag{4.2}$$
$$\mathbf{x}_i = (\mathbf{R}_1^{(i)}, \mathbf{R}_2^{(i)}, \ldots, \mathbf{R}_{N_P}^{(i)}), \tag{4.3}$$
$$E_i = E(\mathbf{x}_i), \tag{4.4}$$
$$f_i = f(E_i), \tag{4.5}$$

where $\mathbf{R}_j^{(i)}$ is the position of atom j in configuration $\mathbf{x}_i$. Contrary to the initial ABC method, no rigorous division into scouts, employees and onlookers is present among the walkers in this approach. Instead, all walkers are of the same type, but two possible search modes are available for every walker. Depending on the demands of the search each one can either perform a free or a local move. The free move is similar to the scout type and places the walker at a completely random point on the PES. The local move, on the other hand, chooses a point in the vicinity of the walker and accepts the move to this new point only if the fitness of the walker is improved through this move.

The ABC algorithm is initialized with all walkers doing one single step in the free search mode, in order to distribute them as randomly as possible on the PES. Afterwards,

each walker switches to the local search mode and remains in this state until no further improvement of the fitness can be found for some amount of local search attempts. The walker, which basically is stuck in a funnel on the PES, switches back to the global search mode, so it can escape this local minimum and start anew at a different point on the PES. Specifically, every time a local search step fails to improve the fitness for a number of times N_L, the displacement factor r_s is reduced. This factor is a parameter for the maximum length of the vector which is used to move the current walker, and if it falls below a certain value, the free move is performed the next time this specific walker does a search step.

This approach resembles a completely random search attempt with added local optimizations. The advantage of ABC over the random approach is caused by the fact that information about the structure of the PES is exchanged between the walkers. It is shared in order to efficiently direct the focus onto promising areas. To accomplish this goal, this method is divided into two different phases, called employee and onlooker phase. During the so-called employee phase each walker either performs one local or free move depending on the current state it is in. Successful walkers are allowed to perform an additional move in the second part of the method, called onlooker phase. Walkers are considered successful if they are in promising areas of the PES. The move probability in the onlooker phase is determined as

$$p_i = \frac{f_i}{\sum_{j=1}^{N_A} f_i} \tag{4.6}$$

for each individual walker. This procedure leads to an increased sampling rate in areas with a higher fitness, which greatly facilitates the discovery of minima. An overview over the method is shown in Figure 4.1.

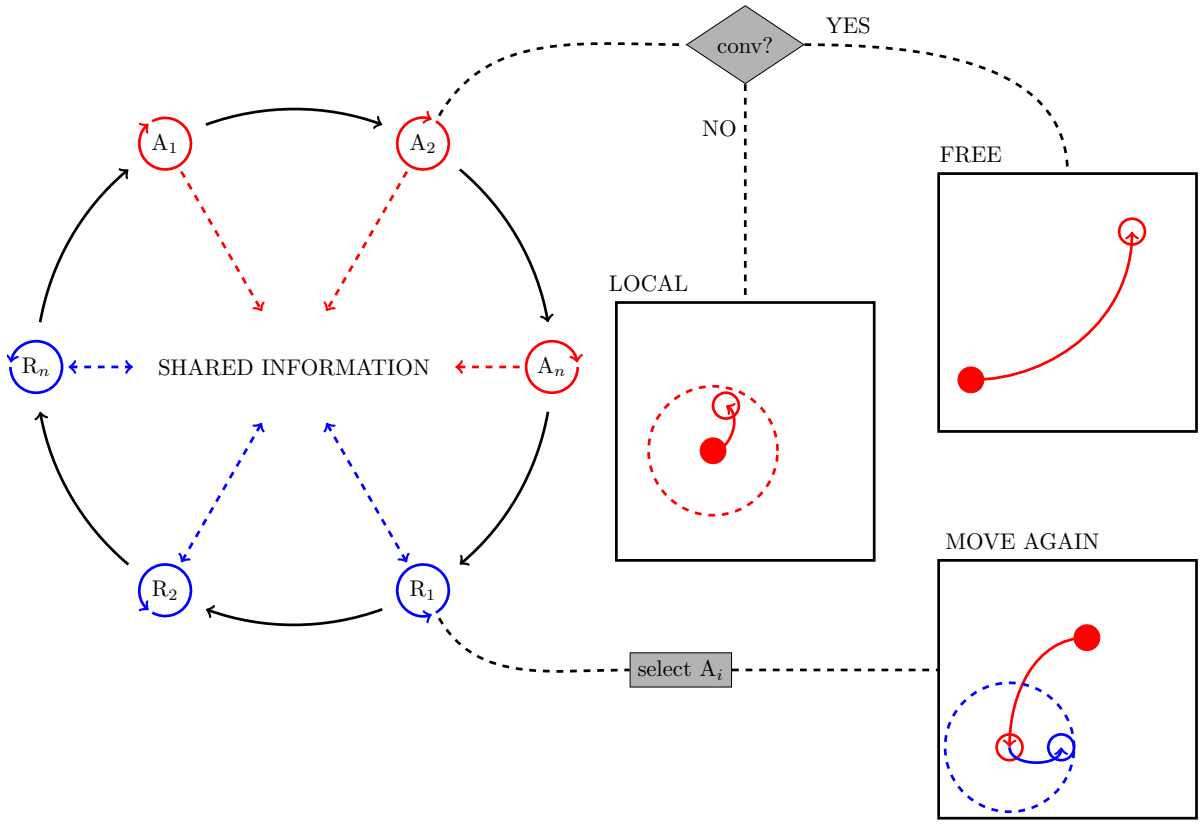

Figure 4.1: Overview of the ABC algorithm [98].

4.3 Results

This algorithm has been applied to different systems of molecular clusters. These systems, namely Morse clusters, Lennard-Jones clusters, and TIP5P water clusters were introduced in Section 2.1. The different settings for the parameters with respect to the number of particles N_P are shown in Table 4.1. The numbers of walkers in the swarm and onlooker state are given by N_A and N_O respectively. The number of allowed local search steps before a reduction of r_s is performed is given as N_L. The number of reductions of r_s before a free move is enforced is given as N_D. All clusters are initialized with the particles in a rather restricted box, defined by the length B_0, but are allowed to expand freely during the simulation.

Due to the high stochastic component in this method all results shown here are the average number of needed energy calculations (single point calculations, SPC) over a

Table 4.1: Settings for the ABC algorithmn with respect to the number of particles N_P in the investigated system.

Agents	N_A	$2 \cdot N_P$	Displacement Factor	r_s	0.01
Onlooker Steps	N_O	$20 \cdot N_P$	Initialization Box	B_0	$0.02 \cdot \sqrt[3]{N_P}$
Limit	N_L	$5 \cdot N_P$	Divisions	N_D	7

high number of simulations.

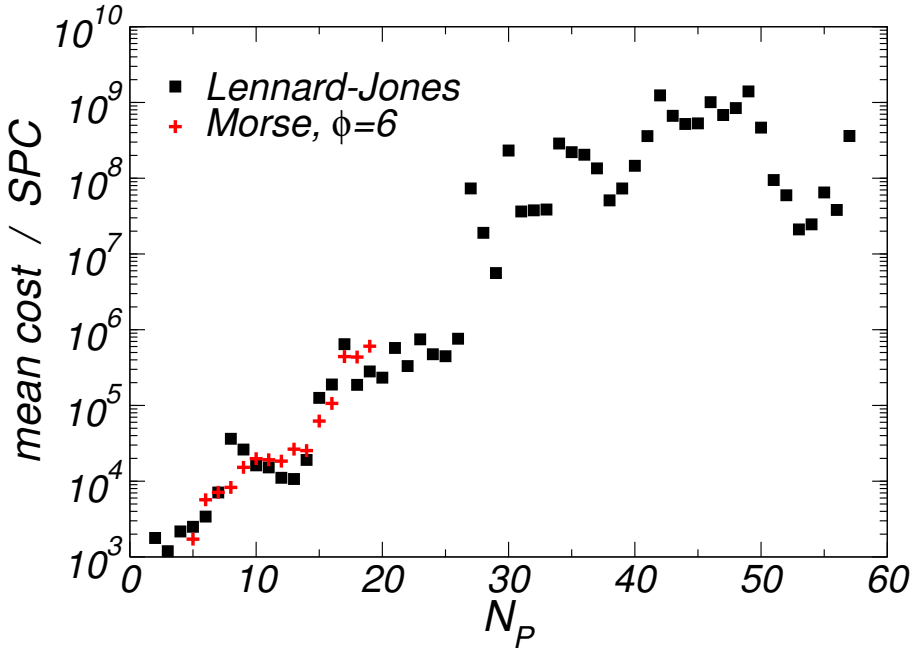

Figure 4.2: Number of needed calculations with respect to cluster size for the LJ clusters and the Morse clusters with $\Phi = 6$ [98].

Results for Morse cluster up to 25 particles and LJ cluster up to 57 particles are shown in Figure 4.2. For the cases with $N_P \leq 25$, results are averaged over 1000 distinct calculations, for the larger systems 200 distinct runs have been performed. The results shown indicate a roughly exponential scaling of calculation time with respect to the system size, although huge fluctuations are present in the calculations for higher cluster sizes. Compared to clusters of similar size, the LJ cluster with 38 particles, which is

generally considered a relatively complicated system, shows no increased calculation time in the ABC method. The PES of this system shows a the steep double funnel geometry, which poses a difficult challenge for most GO methods.

Similar results have been found for water clusters, where again 1000 independent runs have been performed for the smaller systems ($2, 3, 4, 5$ or 7 molecules) and between 31 and 242 for the rest. These results are shown in Figure 4.3. With the exception of the six molecule cluster an almost perfect exponential scaling of the method with respect to the system size can be found.

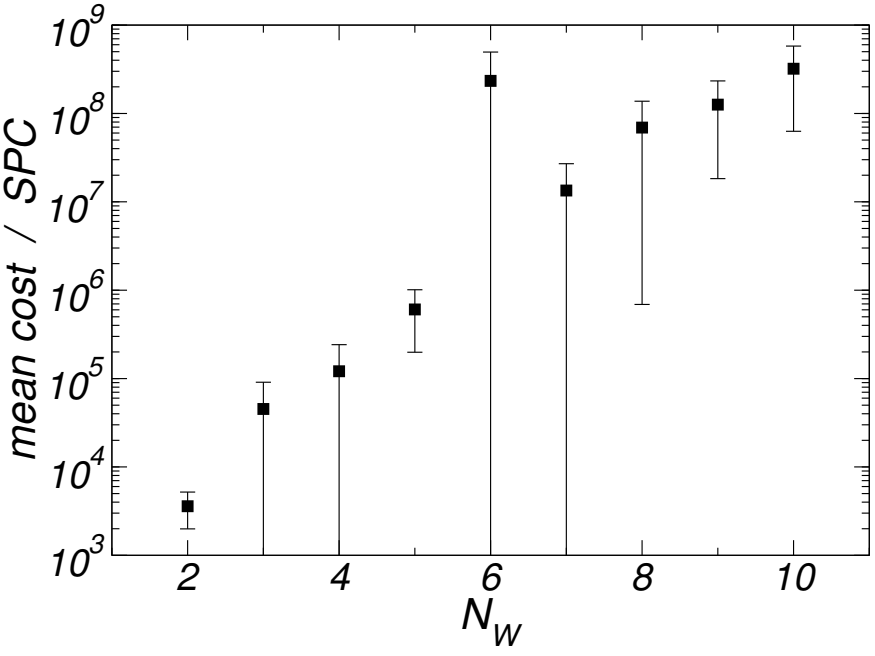

Figure 4.3: Number of needed calculations with respect to cluster size for the TIP5P water clusters [98].

The six molecule cluster is known to be hard to optimize if the TIP5P potential is used. This is due to the fact that two minima of relatively similar energy are present. The ABC algorithm generally finds the higher energy minimum after a time expected if one interpolates the results for the other cluster sizes. The minimum, which is lower in energy, however is rather difficult to find for the ABC method. Both structures are

shown in Figure 2.4 in subsection 2.1.3.

The described issue nicely illustrates the previously mentioned complexity of the GO problem. One of these two water clusters can be found relatively easy and might seem to be the global minimum. Only a method that is capable of escaping the attraction basin of this minimum and explore different parts of the PES is able to find the lower energy structure. Thus this process requires significantly longer simulation times.

4.4 Conclusion

The presented modification of the ABC method has been successfully applied to three different model systems. The algorithm was able to find the global minimum in all cases and scales exponentially with system size. ABC manages the structure discovery without the usage of gradients. The information provided by the gradients is expected to be useful to further improve the algorithm.

5 Extension of the Minima Hopping Method

In this chapter, an approach to find the global minimum of Lennard-Jones clusters using an improved Minima Hopping (MH) is presented.

MH is a powerful method to systematically sample the PES of a system and find the global minimum efficiently. In MH, the discovery of new structures is driven by short MD simulations at an elevated temperature [7, 36] This method has already been applied successfully on various systems like Lennard-Jones and silicon clusters [37, 41, 100].

The main idea behind the enhancements to this method, that we are going to present in this chapter, is the usage of information about the local surroundings at each minimum. This data is to be used in the determination of the search direction. The idea is to extract this local information, specifically the direction of the lowest ascent, from the Hessian matrix of the system. The chance of finding new, lower energy structures is higher if we cross lower barriers. We use this to improve the MH, so that the ground state structure can be found faster. This approach can be used instead of the random approach to achieve a fast and accurate determination of the global minimum in the MH method.

5.1 The Bell-Evans-Polanyi-Principle

The basis of the MH method is the Bell-Evans-Polanyi-principle [12, 31], which states that, in a chemical reaction, the reaction enthalpy is inversely proportional to the activa-

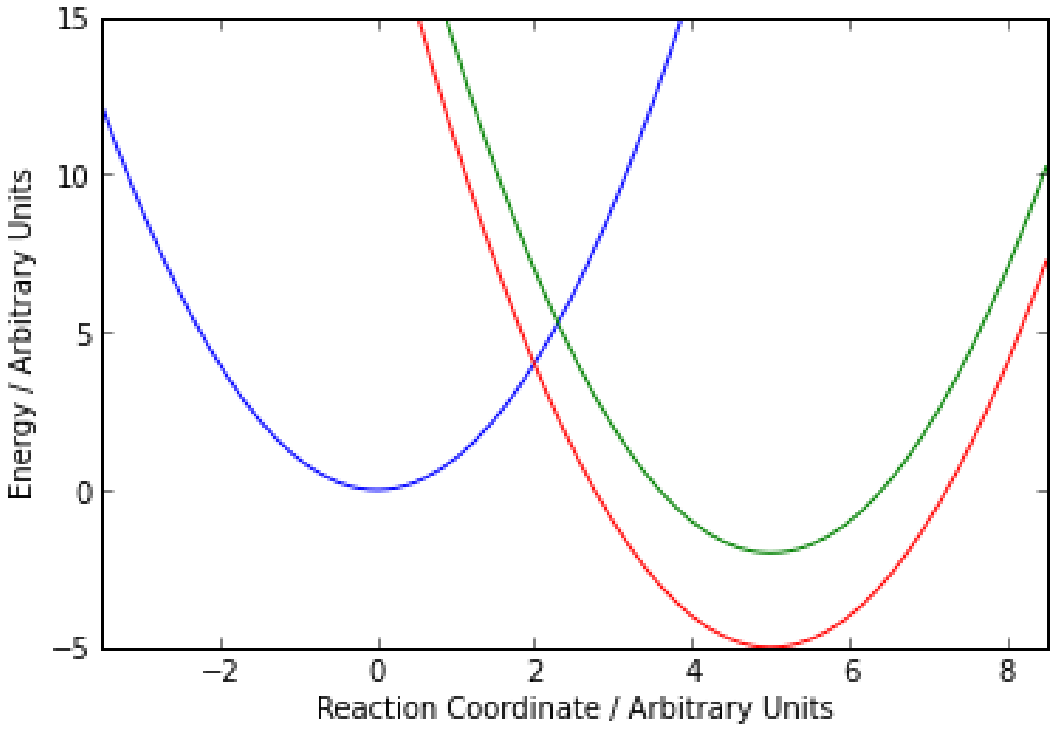

Figure 5.1: Sketch of the general idea behind the Bell-Evans-Polanyi-principle. The barrier (given by the point where two curves intersect) between the blue and the red curve is lower compared to the barrier between the blue and the green curve. This arises due to the lower energy of the minimum described by the red curve.

tion energy, i.e. the height of the energy barrier between the reactants and the products on the PES. A sketch of this principle is shown in Figure 5.1.

Even though we are not considering actual chemical reactions, and are not interested in the transition state, this statement is of great relevance. It implies that the chance of finding structures with a lower total energy is higher, if the simulation preferably directs the system towards crossing these lower energy barriers.

Global optimization approaches in which the Bell-Evans-Polanyi-principle is used, are methods like tabu search [90], metadynamics, and minima hopping. These methods generally use MD, Monte-Carlo (MC) or other methods to direct the system from one minimum to the next in order to sample the PES systematically. The expectation is that an exhaustive search will lead to the discovery of the global minimum of the system.

5.2 Introduction to Minima Hopping

The central part of the MH method is the MD escape. This escape step consists of a short MD run at elevated temperature, which is employed to leave the current minimum and find a new one, preferably a previously unknown one with a lower total energy. An acceptance algorithm which evaluates the new found minimum is employed afterwards. MD achieves the goal of pushing the system to new minima, according to the Bell-Evans-Polany-principle, over the separating energy barriers, while preferably crossing comparatively lower ones.

The desired outcome of this procedure is the discovery of a new minimum that has not been visited before. Consequently, two undesired outcomes are possible: The search may end up in the minimum from which it started, or in one that has already been found in a previous search step.

In order to perform this MD escape, it has to be initialized by assigning a set of velocities and directions to the particles in the system. In MH, these vectors are drawn from a Gaussian distribution and scaled to a given energy E_{kin}. It generally might seem advantageous to choose a rather high value for E_{kin} to increase the chances of leaving

the current minimum, as barriers with an energy lower than E_{kin} are preferably crossed during the MD simulation. This however would enable the MD simulation to cross most barriers, and not only the low lying ones. This is disadvantageous for our approach since we want to preferably cross lower barriers. Hence a balanced value for E_{kin} has to be found if we want to reach the goal of directing the search towards promising regions of the PES.

During the MD run the energy of the system is continuously checked so that we can determine if the system has crossed a barrier into a new region on the PES. As soon as the system has passed a predetermined number of these barriers the MD run is stopped. Afterwards the geometry of the system is relaxed to determine the total energy of the new-found local minimum. Any local minimization scheme like Conjugate Gradient (CG) or quasi-Newton methods (see section 3.1) can be used here. Repeated visits of previously found minima can still happen with this approach, and therefore a history of all found structures has to be kept in MH. As mentioned earlier, every new found minimum is compared against this list to determine if it is already known. It is sufficient to keep track of the energies of all found structures as this is the feature which is used for the distinction of the different minima.

The kinetic energy for the next MD simulation is adjusted depending on the outcome of the previous search step. If a previously known minimum is revisited, the kinetic energy is increased by multiplying E_{kin} with a parameter $\beta_2 > 1$ to give the system the chance to explore more distant areas of the energy surface which are separated by rather high barriers. Similarly, the energy is increased if the search ends in the minimum from which it has started ($E_{\text{kin}} = E_{\text{kin}} \cdot \beta_1, \beta_1 > 1$). Consequently, if a new minimum is found the kinetic energy is decreased ($E_{\text{kin}} = E_{\text{kin}} \cdot \beta_3, \beta_3 < 1$), which leads to a better sampling of the area close to this minimum. This also helps to create a balance between a thorough exploration of areas around metastable structures and an exhaustive sampling of the complete potential energy surface.

If a new minimum is found, there is still the possibility that its energy is too high, and it is therefore undesired. Consequently, each new minimum is accepted or rejected based upon a simple energy threshold. If the energy of the new found minimum is lower than the energy of the start structure, or its increase is less than a predefined threshold E_{diff},

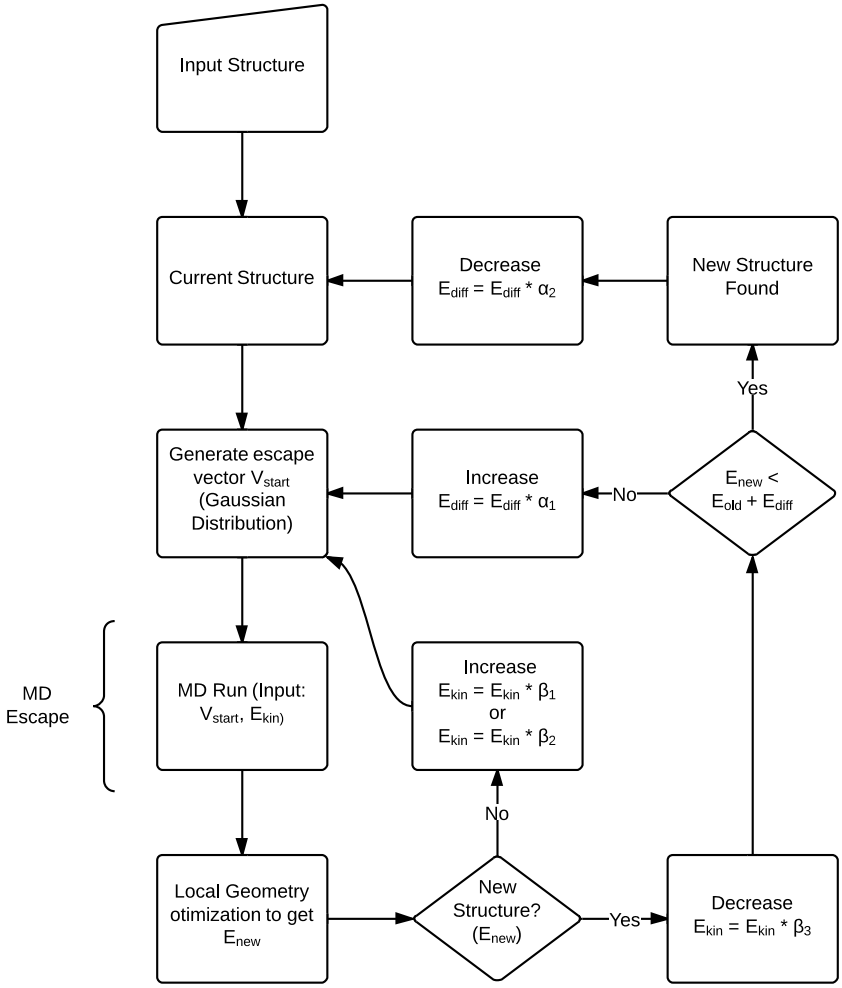

Figure 5.2: Flowchart of the MH algorithm.

it is accepted. This threshold is constantly adjusted during the simulation. Specifically, it is reduced after every accepted move ($E_{\text{diff}} = E_{\text{diff}} \cdot \alpha_2, \alpha_2 < 1$) and increased every time a minimum is rejected ($E_{\text{diff}} = E_{\text{diff}} \cdot \alpha_1, \alpha_1 > 1$). This procedure, an overview of which is depicted in Figure 5.2, prevents the system from moving towards structures with a higher energy and directs it towards regions of lower energy, but also prevents it from being stuck if a wide funnel area around a rather deep minimum exists on the PES.

After a sufficient number of rejected escape tries, almost every move will be accepted and the system can leave the current minimum. Additionally, an acceptance rate of about 50% of all proposed minima is achieved by this method. The values for $\alpha_1 = \frac{1}{\alpha_2} = 1.05$ and $\beta_1 = \beta_2 = \frac{1}{\beta_3} = 1.02$ were taken from Ref. [36].

5.3 Local Optimization

In this work we want to exploit additional information about the system to improve the search. Specifically, we want to use the information about the local surroundings of each minimum, given by an approximation to the Hessian matrix. If a quasi-Newton method is chosen for the local minimization within the MH, this matrix, which contains all partial second derivatives at the minimum, is generated without additional computational effort.

Therefore, we implemented the BFGS [75] and L-BFGS [56, 69] algorithm into the MH method and compared them to the CG minimization, which was used in the initial approach [36], and will serve as a reference. Calculations have been done for LJ clusters in the size range from 60 to 74 particles.

For each cluster size a set of 200 different starting structures has been generated by randomly placing LJ-particles into appropriately sized simulation boxes. This set of configurations is used throughout this work as a starting point for all simulations and all shown results are the average over these 200 independent runs.

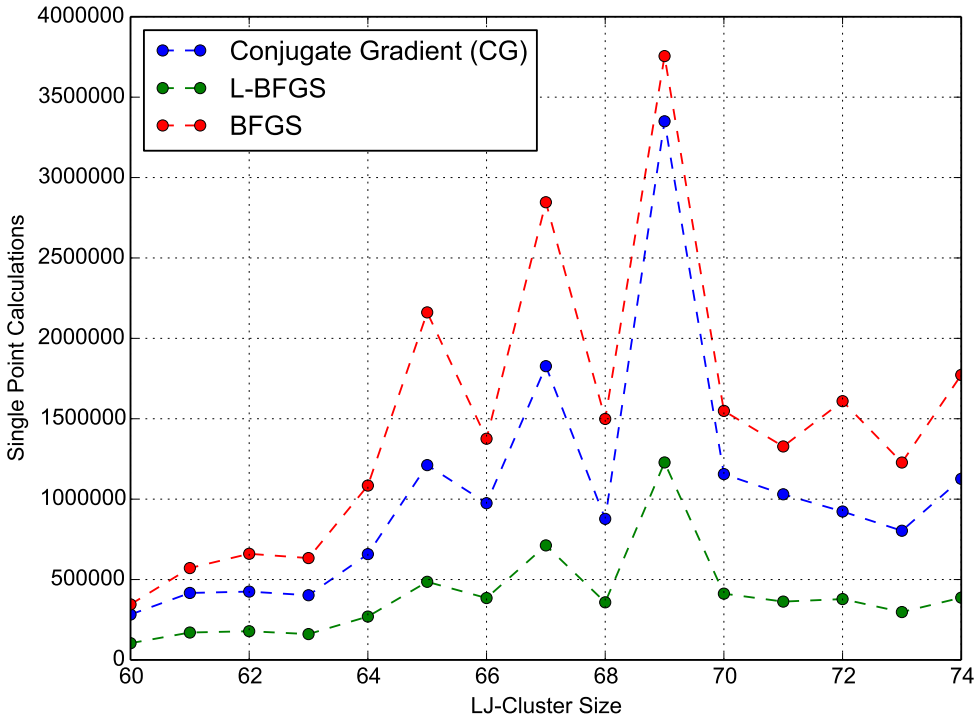

Figure 5.3: Comparison of the used minimization methods in the MH algorithm. Averaged number of performed SPC over 200 runs, started from 200 randomly generated structures. The search direction in the escape part is generated from a Gaussian distribution in all cases.

We started with an evaluation of the performance of the different minimizers, results of which are presented in Figure 5.3. The BFGS algorithm needs the highest number of single point energy and force calculations (SPC), followed by the CG minimization. The L-BFGS method uses a significantly lower number of SPCs compared to the other two methods and gives by far the best performance.

Table 5.1: Detailed results for the use of different minimization methods in MH. Shown for the GO of the 67 particles LJ-cluster.

Minimization Method	Total Number of used SPC	Visited Structures	Barrier Height E_{kin}	Acceptance Criterion E_{diff}
CG	1827025	1975	0.609	0.612
L-BFGS	712283	1795	0.617	0.620
BFGS	2846505	1426	0.521	0.526

Detailed results for the calculations for the cluster with 67 particles are given in Table 5.1. Here again the big improvement attained by L-BFGS can be seen. It needs only about 40% of the number of SPCs needed by the approach with CG. The BFGS method, which needs the highest number of SPC finds the global minimum after the lowest total number of visited structures. This indicates that this method needs even more SPCs per visited structure compared to the other approaches.

A detailed evaluation of the results shows that the highest difference in calculation time is caused by the BFGS minimization, which needs significantly more SPCs to find the local minimum than the other two local optimizations. Especially the line search within the BFGS is responsible for a huge number of calculations. An improvement of this part might be able to bring the performance close to the other methods, and probably even surpass it.

BFGS additionally provides the lowest values for the characteristic parameters of the global search part of the algorithm (E_{kin} and E_{diff}). Values obtained here are are mostly independent of the used minimization method, since these parameters are only relevant in the escape step, and the chosen local minimization method should have no influence on this part of the algorithm. The goal is to reduce the resulting mean values for

these parameters, because this indicates that the algorithm was able to cross low energy barriers and find mainly structures with lower energy.

The average value for the kinetic energy E_{kin} corresponds to the height of the crossed barriers. The acceptance criterion, on the other hand, E_{diff} is the allowed maximal increase in total energy up to which a new proposed structure is accepted. Both parameters are dynamically adjusted during the run so that about half of all moves result in a new, accepted minimum.

5.4 Optimization of the Search Direction

The search direction, which is the vector that is used to initialize the MD escape in the escape part of the MH, is essential for the performance of the method. If it is possible to initialize every search step so that it will end up in a minimum with lower energy, the search will find the global minimum reliably after a low number of steps, at least for cases where the PES is not overly complicated.

The starting point for the search direction is a random draw from a Gaussian distribution. This completely unbiased approach is a reasonable starting point and generally works, as illustrated by the results in the previous section.

5.4.1 Softening

A first improvement to the search direction is softening. During softening the randomly generated escape direction is moved towards low curvature directions using an iterative method [84].

To achieve this, a point p is chosen at a distance d along the randomly drawn search vector. The gradient of the system is calculated at p and a new search direction is generated by moving p along a vector perpendicular to the initial search vector. Using this approach, the gradient is used to change the search vector towards an escape along

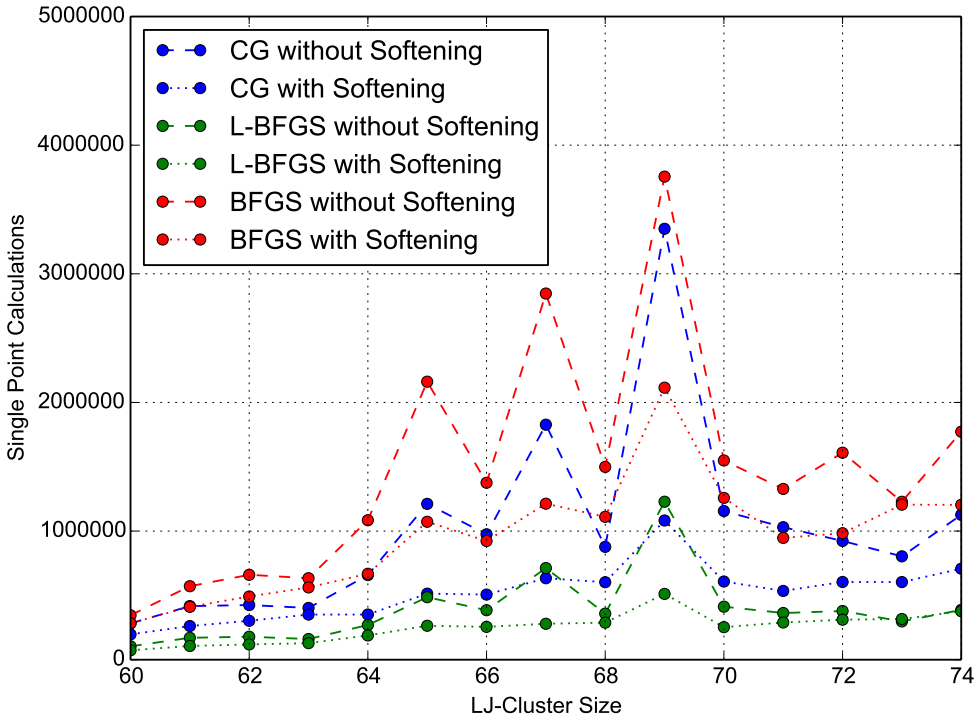

Figure 5.4: Influence of softening on MH for the different minimizations. 20 softening steps have been applied in all cases. The number of needed SPC is significantly reduced upon the application of softening.

a direction where the ascent of the potential energy is lower, thus improving the search. One can repeat this procedure several times, each time starting from the current local minimum and the slightly improved direction, until the starting direction for the MD search is optimized. It is important not to overdo softening since this would lead to a case where the same starting direction is chosen for every escape attempt starting from the same minimum. This process adjusts the randomly chosen direction with respect to the local surroundings and is motivated by the idea to start the search in a direction where the barrier between the current and the next minimum is lower.

To evaluate the influence of different numbers of repetitions of the softening approach a comparison between 0, 10, 20 and 30 softening steps is shown for the BFGS minimization in Figure 5.5. The results show that this approach greatly improves the performance

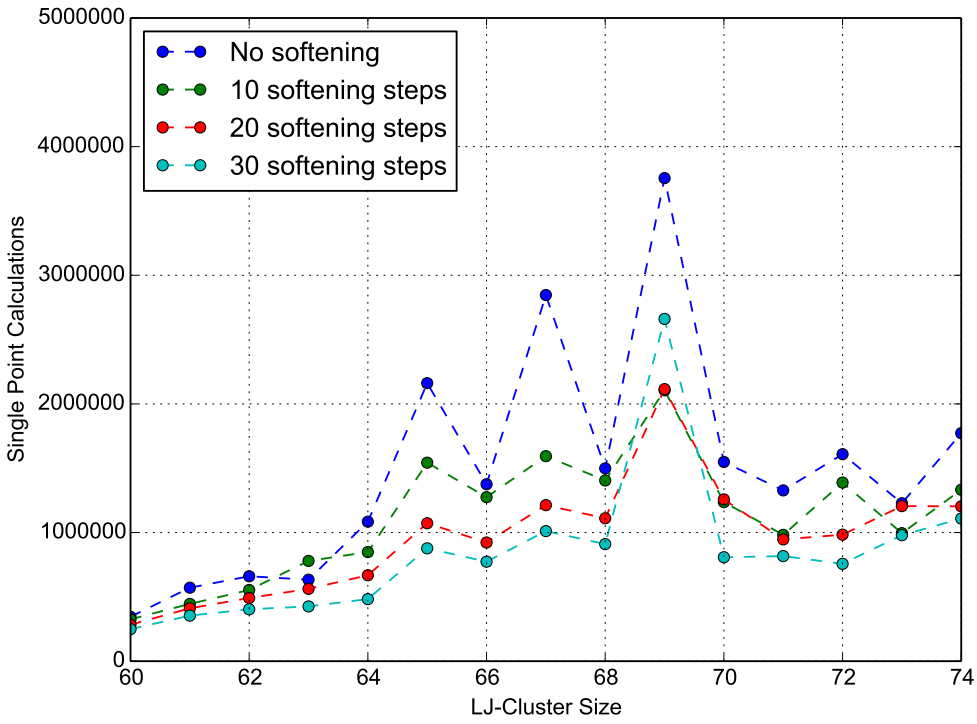

Figure 5.5: Influence of different number of repetitions of the softening approach on the performance of MH. All calculations done with the BFGS minimization.

of the MH in the beginning but there is almost no difference between 20 and 30 repetitions.

Additionally, calculations with 20 softening steps have been performed for all three used minimization methods. The performance of these different approaches is shown in Figure 5.4. Softening is very advantageous in all cases and the general improvement achieved is in the same relative order independent of the used minimization method.

Detailed results for the 67 particle cluster are shown in Table 5.2. Notable here is the fact that the number of visited structures, as well as the average values for the barrier height and acceptance, depend almost exclusively on the application of softening. Softening greatly influences the escape part of the algorithm. The increased efficiency of the search is significant in all observed parameters. Especially the fact that the average height of the crossed barriers E_{kin} is decreased by about 50 % shows that softening results in the desired outcome of preferably crossing lower barriers. The acceptance criterion E_{diff} is also decreased by about half, this indicates that way less unfavourable moves, where the total energy is increased, have to be accepted, and more of the proposed moves find a new minimum.

Table 5.2: Detailed results for the application of softening on the different minimization methods in MH. Shown for the GO of the 67 particle LJ-cluster.

Minimization Method	Softening Steps	Total Number of used SPC	Visited Structures	Barrier Height E_{kin}	Acceptance Criterion E_{diff}
CG	0	1827025	1975	0.609	0.612
CG	20	633477	856	0.289	0.291
L-BFGS	0	712283	1795	0.617	0.620
L-BFGS	20	279394	866	0.299	0.300
BFGS	0	2846505	1426	0.521	0.526
BFGS	10	1593681	884	0.346	0.349
BFGS	20	1212920	710	0.316	0.320
BFGS	30	1011112	615	0.312	0.315

Minimahopping with the L-BFGS local minimization and the application of softening shows a greatly improved performance compared to the initial approach. The number of needed SPCs is reduced to about 15% of the initial value.

Regardless of the great increase of performance provided by softening, we decided to perform the analysis of new approaches without this improvement. This has the advantage that it allows us to examine the influence of new approaches in greater detail. Upon the application of softening two competing methods, where only small differences in the choice of the search direction are present, might show the same result. Softening basically tends to smooth out small differences between applied modifications, because different proposed search directions can be changed so they end up pointing into the same valley on the PES. Calculations with softening generally give better results, so it can still be applied after the other approaches have been tried and evaluated, to improve the best performing methods even further.

5.4.2 Hessian Matrix

Instead of modifying the randomly generated search vector via softening we want to choose this vector as optimal as possible. Our idea is to use information extracted from the Hessian matrix for the improvement of the starting direction in the escape step. The Hessian matrix $\mathbf{H} = (H_{ij})_{i,j=1}^{n}$ of a function $f(x)$ is the matrix of all partial second derivatives, thus each element of this matrix is given as

$$H_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}. \tag{5.1}$$

The eigenvectors of this matrix represent the vibrational modes of the investigated system. This basically implies that, at a local minimum, the eigenvector corresponding to the lowest of the eigenvalues denotes the softest mode of the Hessian matrix and therefore points the system into the direction of the softest ascent out of this minimum. This mode thus represents a favourable direction for the start of the search step, according to the Bell-Evans-Polany-principle. The usage of the Hessian in the MH algorithm should therefore be useful for exploring the PES as efficiently as possible, and enhancing the performance of the global search. In addition to that we can add a white noise term afterwards to preserve a random component in the search.

Unfortunately, the computation of the complete Hessian matrix requires a high number of second derivatives, the calculation of which is expensive. By choosing a quasi-Newton

scheme (namely the BFGS or L-BFGS method, see chapter 3.1) for the local geometry optimization step an approximate Hessian is generated during the minimization, at no additional cost. The quasi-Newton requires only first derivatives to be calculated. The resulting matrix can be employed for the proposed changes to MH.

A local minimization and thus the approximation to the Hessian matrix is generated in any case. In order to improve the performance of the MH method different possibilities on how the provided information from the Hessian matrix can be applied to determine the starting search direction are available.

5.4.3 Lowest Mode of the Hessian Matrix

At first we tested the straightforward approach, where the escape direction is chosen as only the softest mode of the Hessian. This approach was implemented in conjunction with the generation of the Hessian via the BFGS and the L-BFGS method. Specific results are shown in Figure 5.6 and Figure 5.7 respectively. Additionally the influence of softening on these approaches is shown.

In the case of the L-BFGS minimization the use of the lowest Hessian direction needs significantly more SPC steps than in the case where this direction is chosen randomly. In this approach a number of calculations for larger clusters without softening failed to finish all 200 runs within the given time. The additional use of softening mitigates this problem partially but it still offers no improvement over the original approach. Here the improvement provided via softening is huge, this is due the fact that whenever the Hessian mode is really unfavourable, and the system tries this same approach again and again, softening is really helpful by providing improvement that prevents this problem.

In the BFGS case this approach, with only one mode of the Hessian, still gives comparatively poor results, but the relative increase in needed SPC is significantly smaller than in the L-BFGS case. This can be attributed to the fact that the BFGS Hessian is a better approximation to the true Hessian matrix, and therefore the resulting choice of the search direction leads to a higher chance of finding a new minimum.

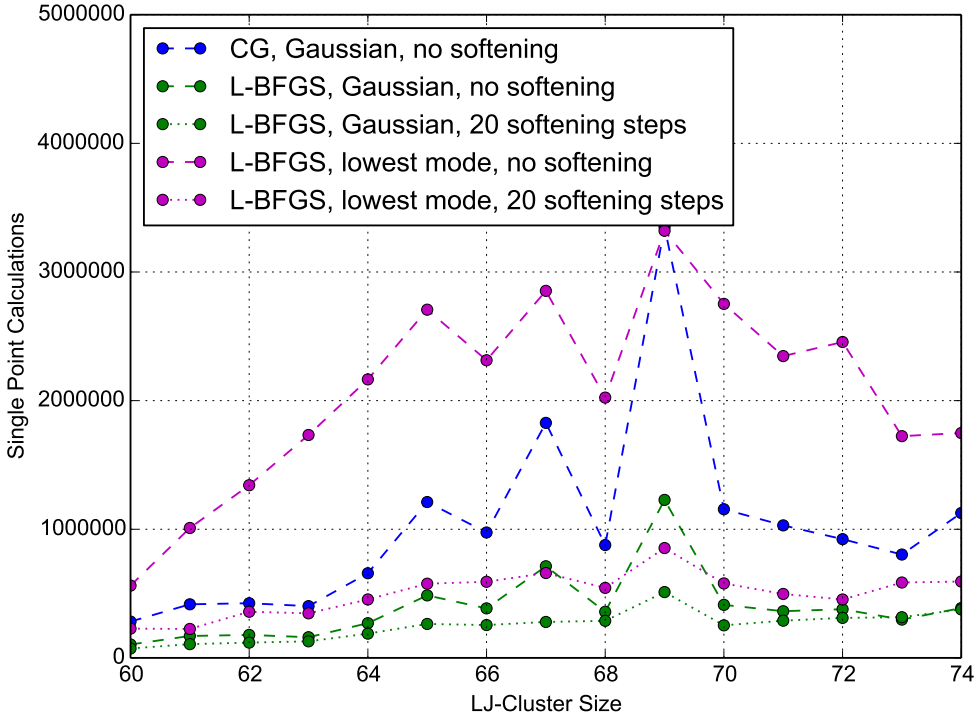

Figure 5.6: Comparison of the different choices for the starting direction in the MD step. Lowest mode of the L-BFGS Hessian matrix compared to a purely random direction. Initial approach with the CG minimization shown as comparison.

Table 5.3: Detailed results for usage of the lowest mode of the Hessian matrix in MH. Shown for the GO of the 67 particle LJ-cluster

Minimization Method	Search Direction	Softening Steps	Number of used SPC	Visited Structures	E_{kin}	E_{diff}
L-BFGS	Lowest mode	0	No comparable results			
L-BFGS	Gaussian	0	712283	1795	0.617	0.620
L-BFGS	Lowest mode	20	659686	2161	0.201	0.203
L-BFGS	Gaussian	20	279394	866	0.299	0.300
BFGS	Lowest mode	0	5767591	2965	0.462	0.468
BFGS	Gaussian	0	2846505	1426	0.521	0.526
BFGS	Lowest mode	20	1528103	902	0.291	0.295
BFGS	Gaussian	20	1212920	710	0.316	0.320

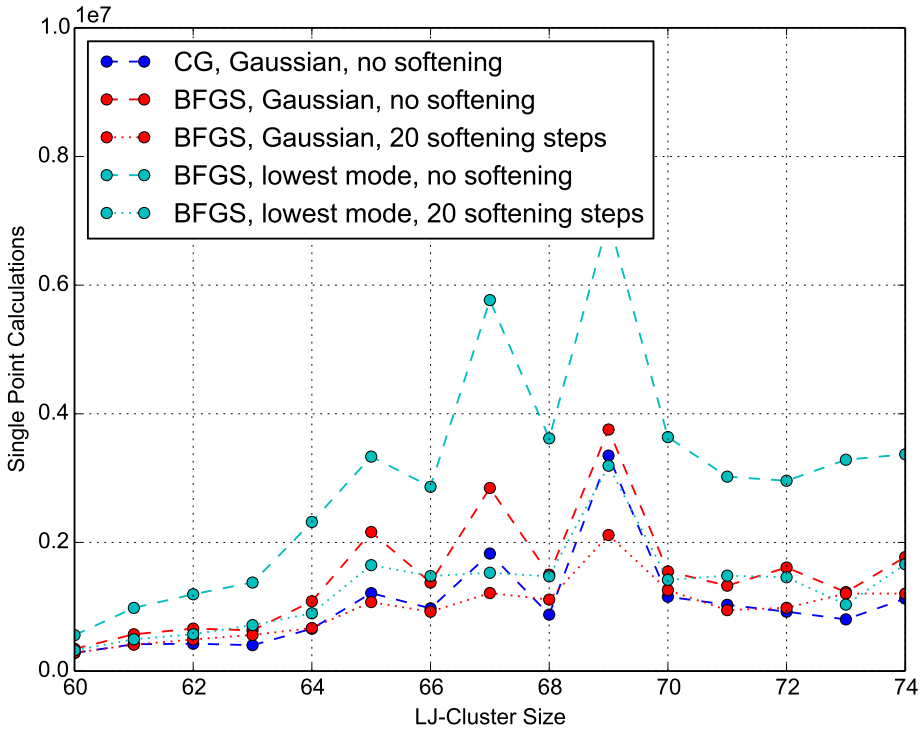

Figure 5.7: Comparison of the different choices for the starting direction in the MD step. Lowest mode of the BFGS Hessian matrix compared to a purely random direction. Initial approach with the CG minimization shown as comparison.

Detailed results for the 67 particle LJ-cluster are shown in Table 5.3. The application of the lowest mode of the Hessian matrix manages to achieve lower values for E_{kin} and E_{diff} compared to the simulations for the same approach with a random direction. The application of softening has however an even greater effect on the values for E_{kin} and E_{diff}. Although the random and the lowest-mode approach show a huge difference in the number of needed SPC, E_{kin} and E_{diff} are very similar if the same number of softening steps is applied.

The problem posed by a static choice of the search direction is clearly present in both cases. In case of an unsuccessful search attempt, the system will try the same escape direction time and time again, merely increasing the kinetic energy E_{kin} of the MD and the acceptance threshold E_{diff} until these parameters have been increased enough to accept any proposed escape. This defeats the purpose of MH and is similar to an (overly complicated) low mode search. By checking the mean values for the energy parameters at the end of the simulation these problems can be detected.

Furthermore the lowest mode alone is not a good choice as the search vector it may only represent just a small rotation of a group which has no influence on the overall structure of the system. In such a case this approach does not help to find the optimal direction for the escape, because it does not help to bring the system to a significant different state. A mixture of low lying modes will probably be a better approach in this case.

This approach provides no improvements compared to the initial approach, but shows that the BFGS Hessian is a better approximation than the one generated by L-BFGS.

5.4.4 Linear Combination of Multiple Low Lying Modes of the Hessian Matrix

The next step is to generate the direction for the MD escape by a linear combination of a number of low lying modes extracted from the Hessian matrix. Since this requires an accurate Hessian, we only implemented this approach in conjunction with the BFGS minimization.

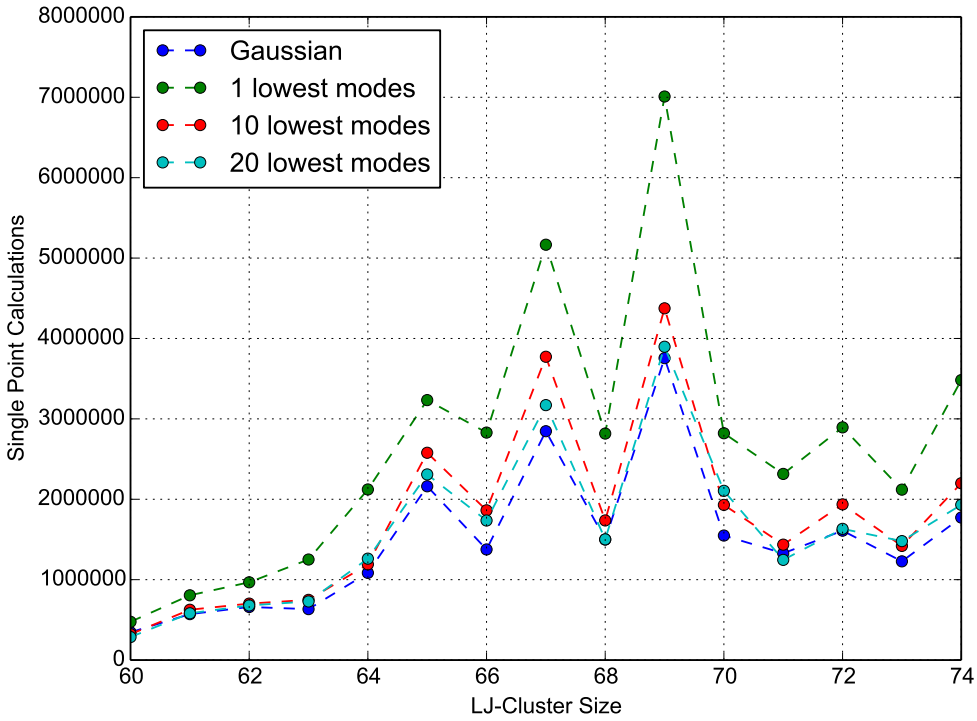

Figure 5.8: Performance of the method with the usage of the lowest mode vs a linear combination of the lowest 10 and 20 modes of the Hessian matrix. The BFGS minimization is used and no softening is applied.

Calculations with linear combinations of different number of lowest modes have been done. The resulting initial guess $\mathbf{v}$ is generated as

$$\mathbf{v} = \sum_{n=1}^{N} \frac{\mathbf{v}_n}{N}. \tag{5.2}$$

Here $\mathbf{v}_n$ denotes the vector generated from the n-th lowest mode. Results are shown in Figure 5.8. By choosing the sum of the lowest 10 or 20 modes, performance similar to the case where the direction was chosen at random can be attained.

Again, detailed results for the 67 particle LJ-cluster are shown in Table 5.4. All parameters converge towards the value of the Gaussian approach, this indicates that the utilization of an increasing number of Hessian modes mitigates the previously stated problem of approaches without a random component. We refer to the fact that these try to cross the same barrier again and again, until E_{kin} and E_{diff} are high enough to allow any transition. The randomness in the original approach is a rather good method to prevent repeated visiting of structures in the vicinity of one found structure and push the system into new regions on the PES.

This approach uses the fact each one of these low lying modes represents an excitation in a different part of the system, and these individual components combined result in a good initial guess for the search direction. Still, the initial, purely random, approach leads to a very similar performance. The logical step forward therefore lies in a combination of the two approaches, executed such that each method compensates the shortcomings of the other.

Table 5.4: LJ-cluster, 67 particle, no softening

Search Direction	Total Number of used SPC	Visited Structures	Barrier Height E_{kin}	Acceptance Criterion E_{diff}
1 mode	5767591	2965	0.462	0.468
10 modes	3772464	1897	0.507	0.512
20 modes	3171139	1592	0.514	0.519
Gaussian	2846505	1426	0.521	0.526

5.4.5 Search Direction as a Mixture of Hessian Matrix Information and a Random Factor

As the next step we wanted to add a certain degree of randomness to the search direction to counteract the previous mentioned problems. The idea of this approach is to again take the linear combination of the N lowest Hessian modes, but this time a weighting factor w_n is drawn for every mode $\mathbf{v}_n$. The factor w_n is generated as a random number in the range from -1 to 1. This approach uses the Hessian information but no longer has the issue of repeatedly trying the same escape until it is accepted. Every new escape attempt from the same minimum will generate a new set of weighting factors and thus a new escape direction.

$$\mathbf{v} = \sum_{n}^{N} \frac{w_n}{N} \cdot \mathbf{v}_n. \tag{5.3}$$

Results are shown in shown in Figure 5.9 and Table 5.5.

If the here generated initial direction is additionally mixed with the Gaussian direction from the initial approach a slightly improved performance can be attained compared to the completely random approach. The gain is very small and might be countermanded by the additional work that is needed for extracting the Hessian modes. These results are shown in Figure 5.9 and Table 5.5.

The approaches where the search direction is not mixed with a vector generated from a Gaussian distribution show again lower values in the characteristic parameters E_{kin} and E_{diff}. This proves that these methods on average cross lower barriers and have to accept less unfavourable moves.

5.4.6 Search Direction as a Mixture of Hessian Information and a Variable Factor

We now consider a combination of all formerly described approaches with the objective to obtain an even more efficient way of computing the global minimum. Here, the first

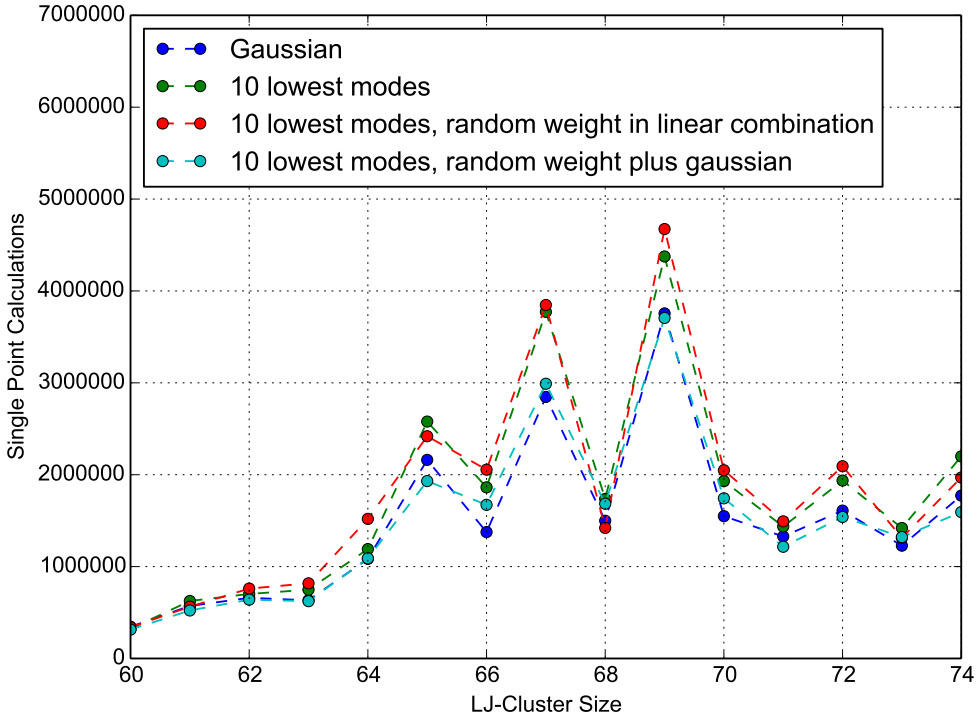

Figure 5.9: Performance of the method with the usage of the 10 lowest modes compared to a randomly weighted linear combination of the same 10 modes and an approach where this is additionally combined with the Gaussian direction from the initial approach. The BFGS minimization is used and no softening is applied.

Table 5.5: LJ-cluster, 67 particle, no softening

Search Direction	Total Number of used SPC	Visited Structures	Barrier Height E_{kin}	Acceptance Criterion E_{diff}
Gaussian	2846505	1426	0.521	0.526
10 modes	3772464	1897	0.507	0.512
10 modes, random weight	3847023	1940	0.501	0.503
10 modes + Gaussian	2990028	1496	0.521	0.524

escape attempt out of each and every minimum is done by using only the direction from the Hessian matrix, similar to the method presented in subsection 5.4.4. For each subsequent escape attempt, starting from the same minimum, the initial guess is composed from the randomly generated escape vector $\mathbf{v}_{\text{random}}$ and the Hessian direction $\mathbf{v}_{\text{Hessian}}$ as follows:

$$\mathbf{v} = \alpha \mathbf{v}_{\text{random}} + (1 - \alpha)\mathbf{v}_{\text{Hessian}}. \tag{5.4}$$

The vector $\mathbf{v}_{\text{Hessian}}$ is either chosen as only the one lowest mode or a linear combination of the N lowest modes, and $\mathbf{v}_{\text{random}}$ is drawn from a Gaussian distribution. The parameter α is defined as follows

$$\alpha = \begin{cases} 0 & \text{if } a = 0 \\ \min\left(1, b \cdot c^a\right) & \text{else} \end{cases}. \tag{5.5}$$

Here, a denotes how many previous escape attempts have been started from this minimum, b and c are parameters that determine how fast α is increased. The value of α is zero for the first search attempt out of every minimum, so that only $\mathbf{v} = \mathbf{v}_{\text{Hessian}}$ is chosen as the search direction in this case. If this search is not successful, α is increased exponentially to raise the amount of $\mathbf{v}_{\text{random}}$ that is introduced into the escape vector. At some point α will reach one, and consequently only the random search direction is applied for each further escape attempt from this minimum.

Table 5.6: LJ-cluster, 67 particle, no softening

Search Direction	Total Number of used SPC	Visited Structures	Barrier Height E_{kin}	Acceptance Criterion E_{diff}
Gaussian	2846505	1426	0.521	0.526
$b = 0.01, c = 1.4$	4639592	2386	0.453	0.456
$b = 0.1, c = 1.3$	3676727	1865	0.483	0.486
$b = 0.1, c = 1.3$, 5 modes	3655347	1840	0.495	0.498

In Figure 5.10 and Table 5.6 the results for different values for b and c are shown. It can be seen that the best results are obtained if the calculation adds the random part as early as possible, and, as with previous results, if a higher number of Hessian modes is used.

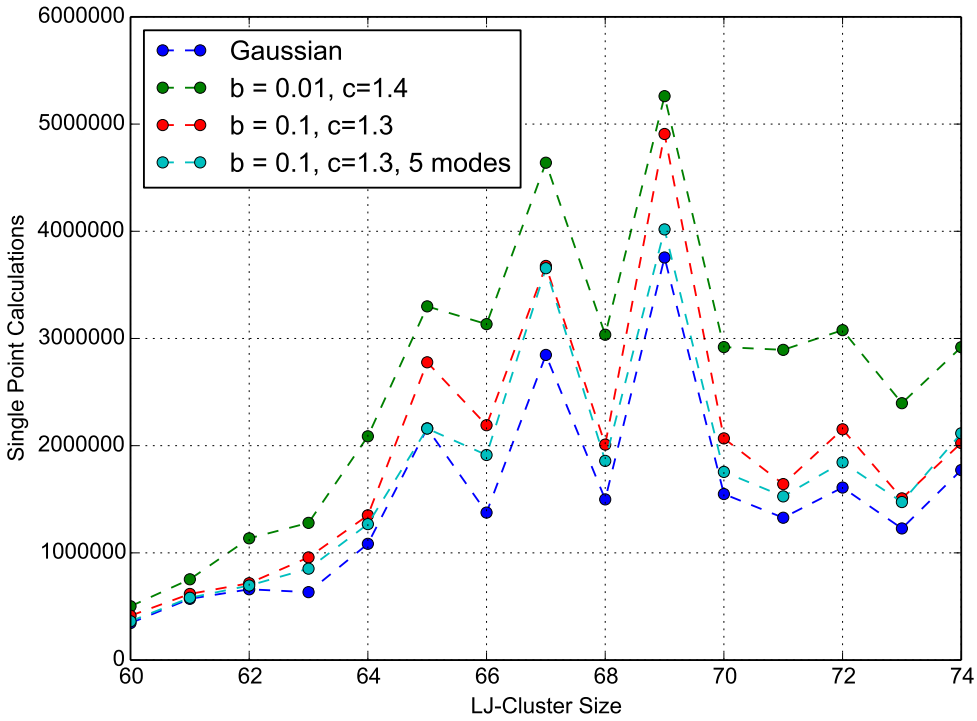

Figure 5.10: Performance of the method where an increasing portion of the Gaussian vector is included into the search direction after each unsuccessful move. The BFGS minimization is used and no softening is applied.

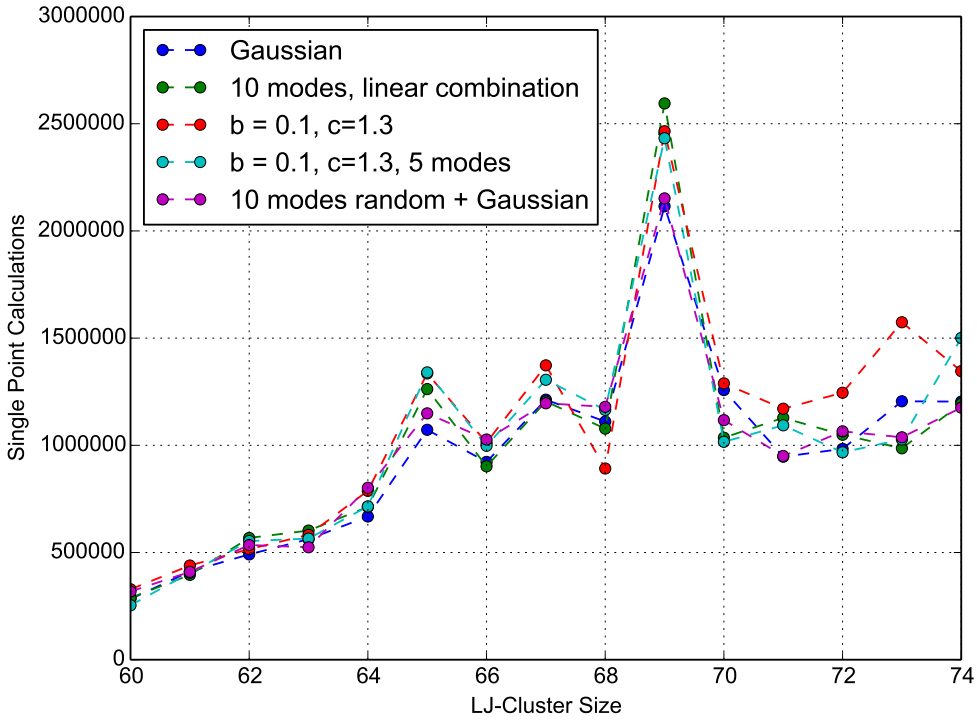

Figure 5.11: Comparison of the best performing methods with the addition of softening.

This approach offers a comparable efficiency to the approach with only the Gaussian vector as the search direction. The barrier height E_{kin} and acceptance criterion E_{diff} are again lower whenever the Hessian information is used. This is proof that our proposed goal of crossing lower lying barriers is achieved. Noteworthy is the fact that the best performing method visits a significantly higher number of trial structures before finding the global minimum. Possible explanations for this behaviour are discussed later (see section 5.5).

5.4.7 Best Performing Methods including Softening

A selection of the best performing methods has been chosen and calculations with the addition of 20 softening steps have been performed for these approaches. A comparison of the results for these calculations is presented in Figure 5.11. As shown previously, the

Table 5.7: LJ-cluster, 67 particle, 20 softening steps

Minimization Method	Total Number of used SPC	Visited Structures	Barrier Height E_{kin}	Acceptance Criterion E_{diff}
Gaussian	1212920	710	0.316	0.320
10 modes	1204411	711	0.312	0.317
Mixed approach $b = 0.1, c = 1.3$	1372927	805	0.307	0.309
Mixed approach $b = 0.1, c = 1.3$, 5 Modes	1305609	765	0.309	0.312
Mixed approach Gaussian + 10 modes	1196007	702	0.318	0.320

softening greatly decreases the number of needed SPCs for all implemented methods. The needed SPCs, the barrier height E_{kin} and the acceptance criterion E_{diff} for the 67 particles LJ-cluster are shown in and Table 5.7. For these values the same trends as in previous calculations without softening can be found.

5.5 Conclusions

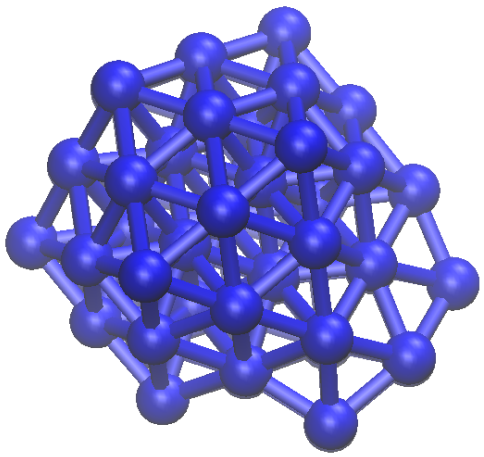

Figure 5.12: 60 particles LJ-cluster.

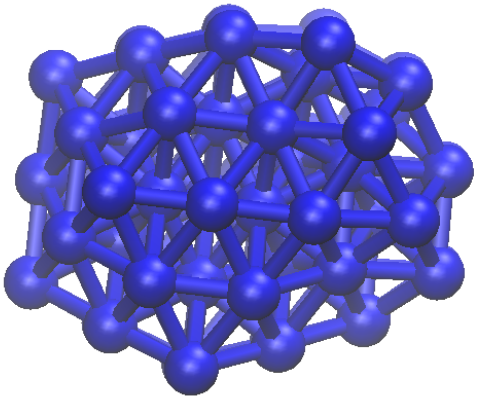

Figure 5.13: 67 particles LJ-cluster.

The here presented results indicate that the usage of the local surroundings of each minimum has the potential to improve the calculations. Although the starting approach is already very good the inclusion of the information provided by the Hessian matrix is indeed able to increase the performance of the MH algorithm. This improvement is rather small and can be easily negated by few detrimental factors. An overview of the results is given in Figure 5.15 and Table 5.8. As examples, the discovered minimal energy structures for the LJ-cluster with 60, 67 and 74 particles are shown in Figures 5.12 - 5.14.

Table 5.8: LJ-cluster, 67 particle

Minimization Method	Total Number of used SPC	Visited Structures	E_{kin}	E_{diff}
BFGS, Gaussian, no softening	2846505	1426	0.521	0.526
BFGS, Mixed approach, no softening $b = 0.1, c = 1.3$, 5 Modes	3655347	1840	0.495	0.498
BFGS, Gaussian, 20 softening	1212920	710	0.316	0.320
BFGS, Mixed approach, 20 softening $b = 0.1, c = 1.3$, 5 Modes	1305609	765	0.309	0.312
L-BFGS, Gaussian, 20 softening	279394	866	0.299	0.300

We were able to apply the information extracted from the local surroundings of each minimum in order to preferably cross lower barriers during the simulation. This is proven by the decrease of the values for the average barrier height E_{kin} and mean acceptance

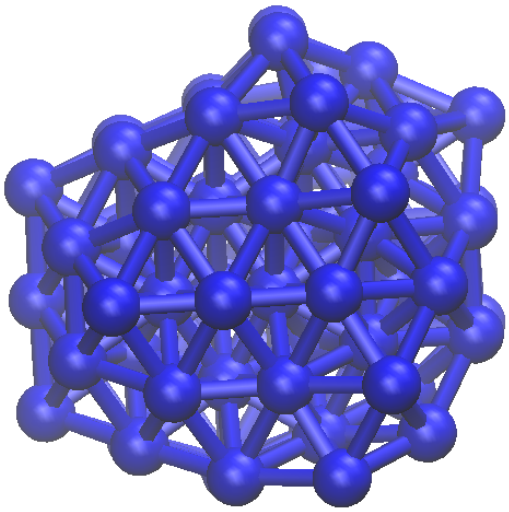

Figure 5.14: 74 particles LJ-cluster.

criterion E_{diff} whenever the Hessian information is used in a calculation.

It is noteworthy that the sampling efficiency is rather indifferent to the choice of the search vector, as long as at least a small stochastic element is present, or a linear combination of modes from the Hessian matrix is chosen. Even though the total performance might be similar, the different implemented approaches however provide valuable insight how the simulations work in general. As stated before, the shown results are an average over 200 simulations with different random seeds, therefore the performance might be negatively influenced by cases where the usage of the Hessian matrix causes problems. For example the generated Hessian approximation might be not sufficiently accurate. If in only a few of these 200 simulations the Hessian mode points into a disadvantageous directions, the increased computational work needed for these particular runs may already completely negate the improvements achieved in the other runs. Another problem are cases where the system alternates between a number of low lying modes. If a number of minima with similar energy exist close to each other, the Hessian matrix based approaches will cause the system to jump between these minima, because the lowest ascent out of the minimum points always into the direction of these other minima with similar energy. This fulfils our requirement of crossing only low lying barriers, but does not provide any progress towards finding the global minimum. A random approach has a way higher chance of leaving this area of the PES and thus discover new minima. It is

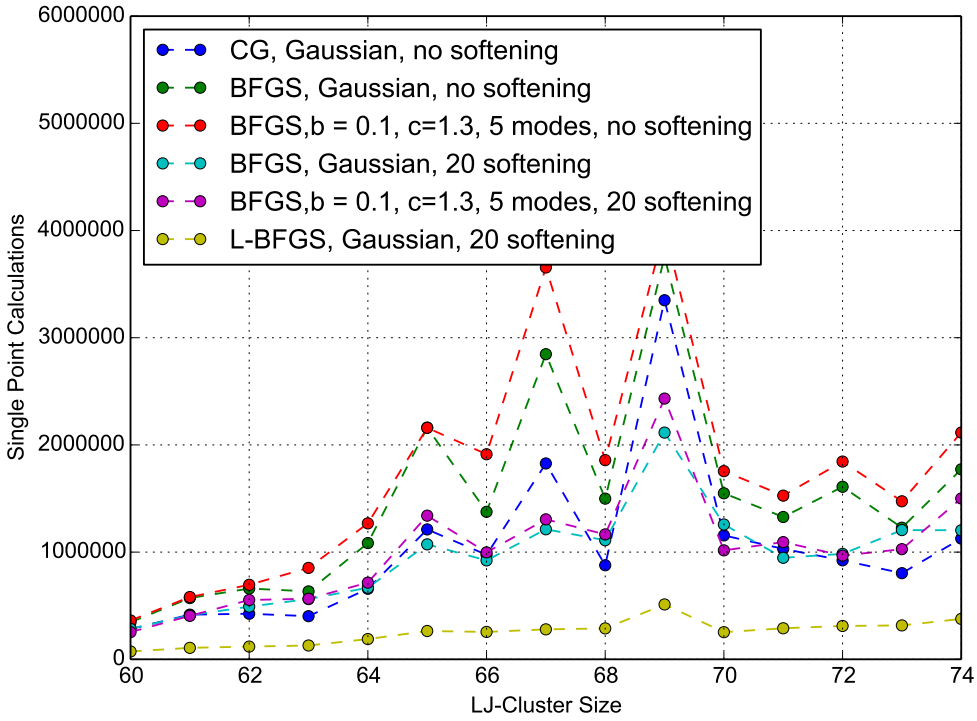

Figure 5.15: Overview over the best performing approaches within th MH scheme.

noticeable that the best performing method visits a huge number of minima before the global minimum is found, this indicates that this problem happens here.

Another defining factor for the performance of the method is the local optimization. The biggest increase of the overall performance of the MH algorithm was attained by implementing the L-BFGS minimization into the method. The performance of this method significantly increases the speed of the local minimization part of the method, and the runs performed with it are significantly faster than the ones with the BFGS and the CG minimization. Calculations with the L-BFGS optimization combined with softening deliver the best performance of all analysed methods.

6 Metallization in SiH_{12}

6.1 Introduction

Hydrogen, even though it is the simplest element in the periodic table, is of high interest in many fields of research. Already in 1935 it has been predicted by Wigner and Huntington [99] to become a metallic conductor under high pressure. They assumed that it would transform from its molecular form to an atomic solid upon the application of the said pressure. The existence of metallic hydrogen has implications in many fields of research. It is very possible that it is a high-T_c superconductor. Additionally, it is relevant in astrophysics to explain the huge magnetic fields present at gas giants like Jupiter and Saturn. These planets are made up from light elements like hydrogen and helium and possess some of the strongest planetary magnetic fields in the solar system. The existence of a metallic hydrogen core may provide a convincing explanation for this phenomenon [6].

6.1.1 Metallization in Hydrogen

The phase diagram of hydrogen and the question of its metallization is a much discussed topic. The melting line of hydrogen shows a negative slope [81], which suggests the possibility of the existence of liquid metallic hydrogen at low finite temperatures [25, 28, 30]. Additionally, it might even be possible to attain a metallic phase in molecular hydrogen if the bandgap closes due to the application of very high pressure [10, 76]. Unfortunately the structures of hydrogen under higher pressure (> 150 GPa) are still unknown [42, 57],

and therefore the question if it might become metallic is still open [21, 43, 62]. The problem of the determination of the structure stems from the fact that X-ray diffraction on the light hydrogen atoms is not possible. Hence, a huge number of possible structures have been predicted for hydrogen. Most of these are very similar in energy, making it difficult to determine which one represents the global minimum. A metallic state has been found in most of these systems [5, 10, 19, 20, 47, 49, 52, 76, 89].

6.1.2 Hydrogen Rich Compounds

Another promising attempt to achieve metallization in hydrogen is the introduction of different elements into the system. Hydrides of group 14 elements are one of the most studied systems in this regard. Silane (SiH_4) is already known to show metallic behaviour under pressure [29]. Similar results have been found if more hydrogen is added to the system. In 2009 Strobel et al. reported the discovery of the high pressure system $SiH_4(H_2)_2$ [9, 91]. In their experiments the sample shows darkening, which is a sign of a metal-insulator transition, already at around 35 GPa . Simulations find metallization in this system at much higher pressure, namely at around 164 GPa [77]. Inspired by these results we are interested in the properties of a system with a even further increased hydrogen content. Therefore, we are examining the electronic structure of SiH_{12} to discover its structure and metallization pressure.

6.2 Results and Discussion

6.2.1 High Pressure Structure of SiH_{12}

The evolutionary algorithm as implemented in the USPEX package [70] was used to determine the minimum energy structure of SiH_{12} at different pressures. Simulations have been performed at 25 GPa, 50 GPa, 100 GPa 200 GPa, and 300 GPa, with four formula units of SiH_{12} per unit cell. The local minimization and energy calculations needed to evaluate the structures in the evolutionary algorithm have been conducted using the

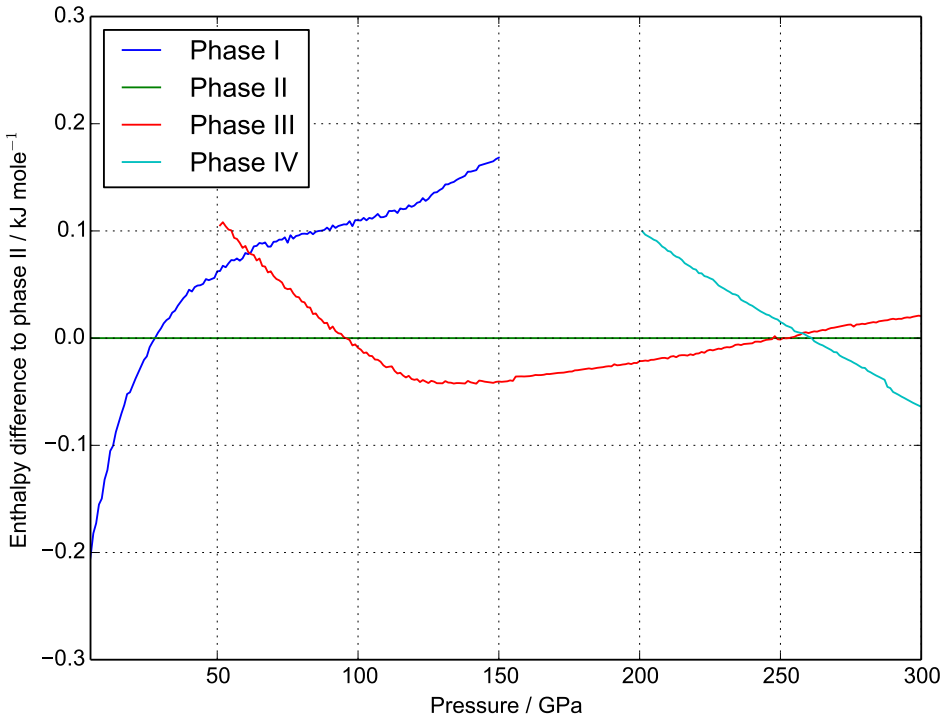

Figure 6.1: Enthalpy difference to phase II in SiH_{12} in the pressure range from 0 to 300 GPa.

Quantum ESPRESSO code [34] with the PBE exchange-correlation functional [73]. A kinetic energy cut-off of 612 eV has been chosen for the calculations. The discovered lowest energy structures were expanded to higher and lower pressures by a stepwise change in the applied pressure in the simulation. A local structure optimization has been performed after every step.

We find four different minimum energy structures for SiH_{12} in the pressure range from 0 to 300 GPa. The first modification is stable up to a pressure of 28 GPa, where it transforms to phase II. This modification is the lowest energy structure in the pressure range from 28 to 95 GPa. At even higher pressures the system changes into a third modification which is the stable one until the system changes to a phase IV above 260 GPa. The enthalpy differences between the different phases of SiH_{12} are shown in Figure 6.1. It is shown relative to the enthalpy of phase II. We present now one example structure for each of the found modifications to discuss the differences.

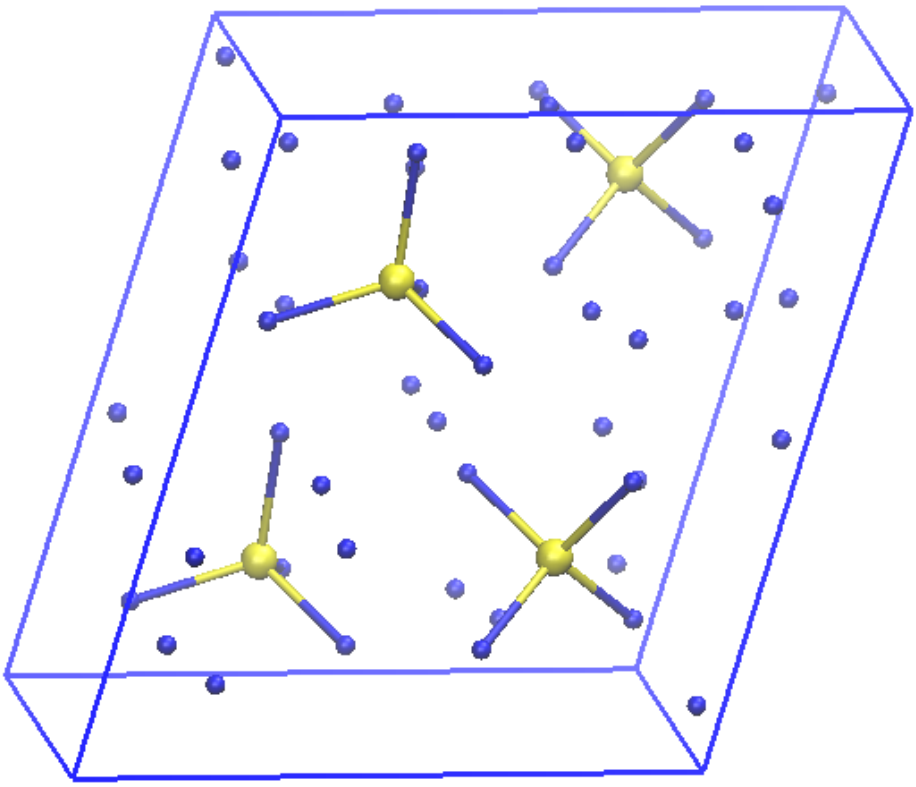

Figure 6.2: Phase I of SiH_{12} at 25 GPa. This phase is found to be stable from 0 to 28 GPa.

6.2.2 Phase I

In the lowest pressure range we find elemental hydrogen surrounding individual SiH_4 molecules. This is not surprising since this is the molecular configuration one would expect at ambient pressure. The structure found at 25 GPa is shown in Figure 6.2.

6.2.3 Phase II

With slightly increased pressure the structure shifts to a phase where an interconnection between the silane starts to form. This silane layer is separated from hydrogen. The silicon containing layer is composed of exactly four formula units of SiH_4, and the re-

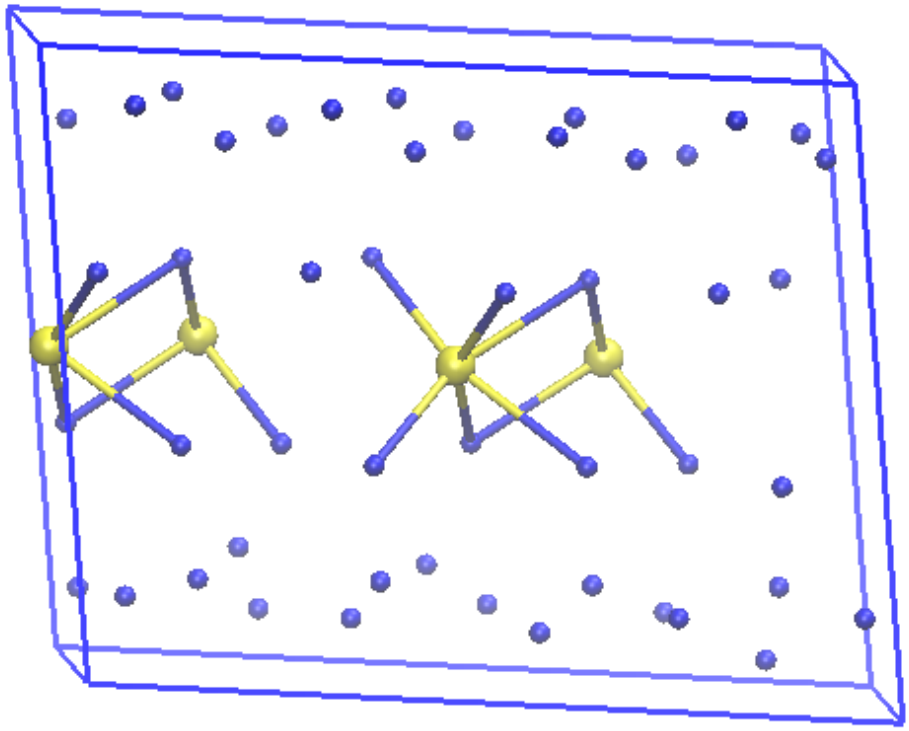

Figure 6.3: Phase II of SiH_{12} at 50 GPa. This phase is found to be stable from 28 to 95 GPa.

maining hydrogen exists in two layers where an ordered arrangement of H_2 molecules is present. The substructure of the silicon atoms is not perfectly symmetric, a top view of six unit cells can be seen in Figure 6.4.

6.2.4 Phase III

The modification found to be stable between 95 and 260 GPa is very similar to phase II and shown in Figure 6.5. It is of slightly higher symmetry. We find it to occupy space group 14 ($P2_1/c$). Each silicon is coordinated by eight hydrogen atoms, which in turn are coordinating two silicon atoms each. Again the silane layer contains four formula units of SiH_4, whereas the rest of the hydrogen forms a hydrogen layer. The silicon sub-lattice is shown in Figure 6.6, it shows a perfectly square symmetry.

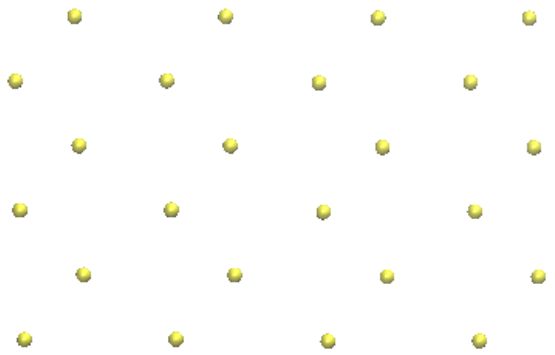

Figure 6.4: Top view of the silicon sub lattice of phase II.

6.2.5 Phase IV

The modification of SiH_{12} we find at the highest pressures now shows a highly symmetric structure where every silicon atom is 16-fold coordinated by hydrogen. It occupies space group 11 ($P2_1/m$).

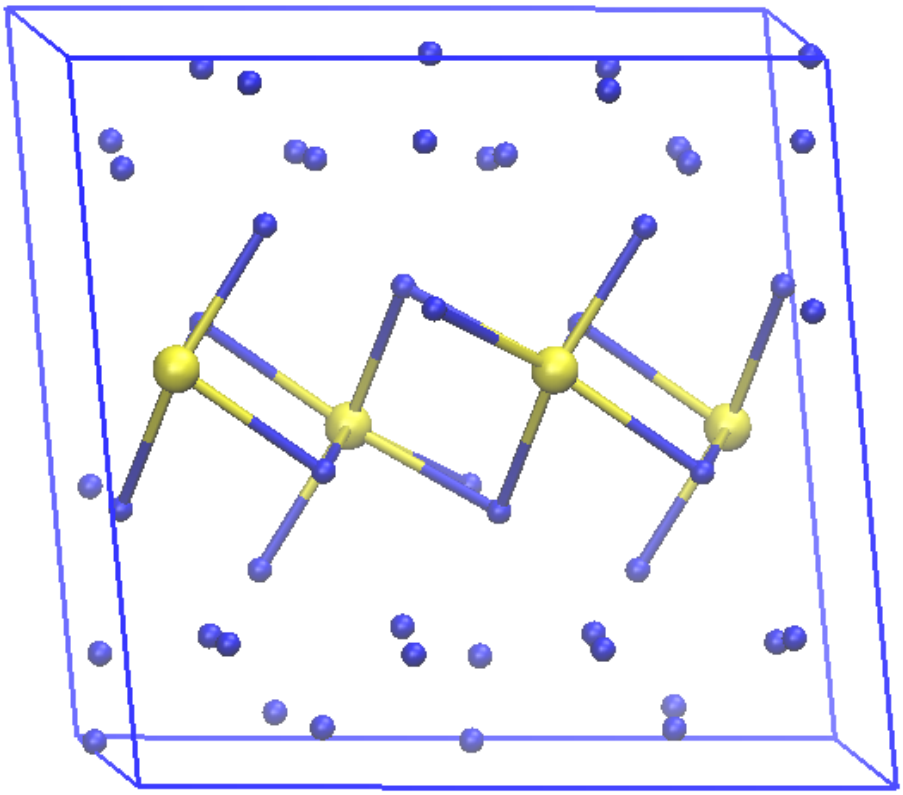

Figure 6.5: Phase III of SiH_{12} at 150 GPa. This phase is found to be stable from 95 to 260 GPa.

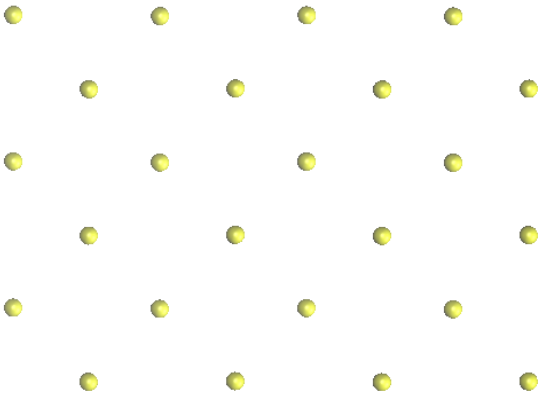

Figure 6.6: Top view of the silicon sub lattice of phase III.

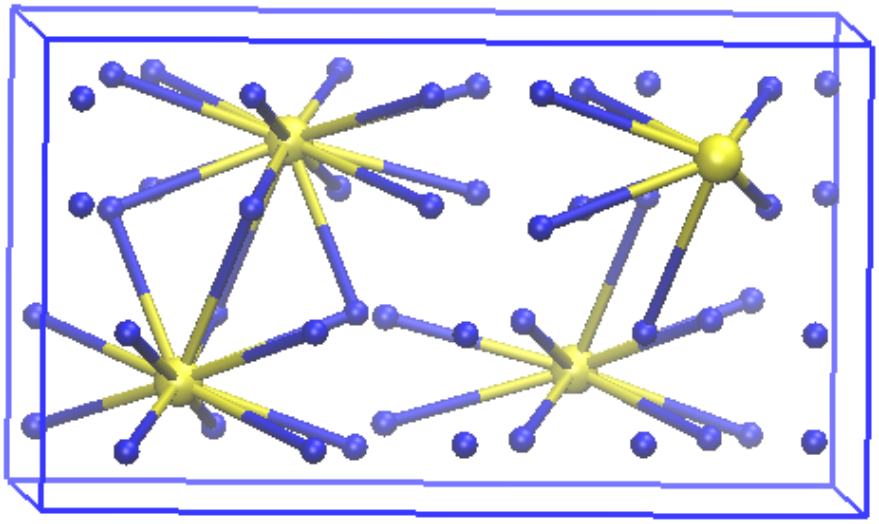

Figure 6.7: Phase IV of SiH_{12} at 300 GPa. This phase is found to be stable above 260 GPa

6.2.6 Electronic Structure of SiH_{12}

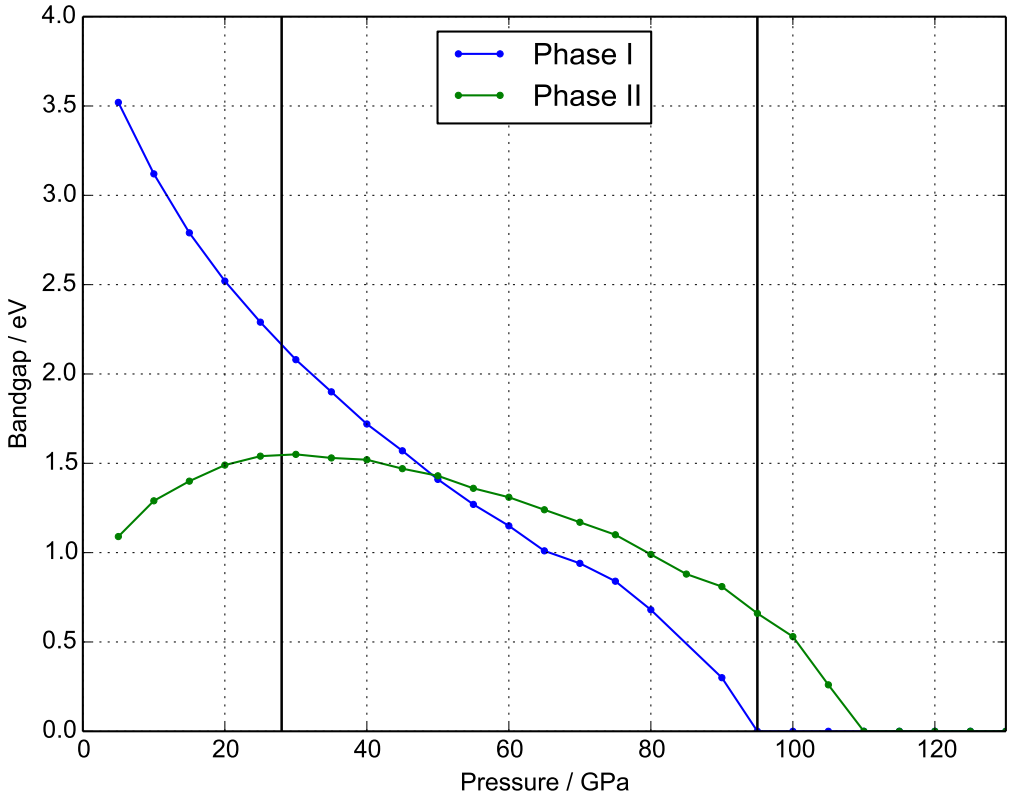

Figure 6.8: Bandgap of the different found modifications of SiH_{12} with respect to the pressure. The vertical lines indicate the pressures where a transformation from one phase to the next is found. Phase I, which is stable below 28 GPa shows the largest bandgap at low pressures. In the range between 28 and 95 GPa , phase I and II behave similar.

We calculated the band structure and density of states for the found structures. These calculations have been performed using the same exchange-correlation functional and energy cut-off. The bandgaps with respect to the applied pressure are shown in Figure 6.8. Phase I and II show a metallization at around 95 - 110 GPa. Phase III and IV are found to have no band gap in the observed pressure range. It is important to note that phase III is the stable structure at pressures over 95 GPa, the pressure where the bandgap in phase II closes, this indicates that the SiH_{12} system might show metallization with the transformation to phase III. Phase I, which is stable below 28 GPa shows the largest

bandgap at low pressures, this is expected, since it shows a structure composed from SiH_4 and H_2 molecules.

Unfortunately, the fundamental band gap is generally underestimated at a local and semi local level of theory in DFT [74]. Since these calculations have only be performed using the PBE functional we have to keep in mind that the real metallization pressure is probably higher. The bandgap generally is underestimated by $\sim$50%. In calculations using hybrid DFT [89] the calculated bandgap is increased, and therefore also a higher metallization pressure is found. The difference in the metallization pressure, between these levels of theory, in $SiH_4(H_2)_2$ has been reported as 20 GPa. It is reasonable to assume that an increase in a similar order can be found for SiH_{12}.

6.2.7 Density of States

The density of states of the different SiH_{12} phases have been calculated. In Figure 6.9 and Figure 6.10 the DOS of phase I and II is shown at the point where the system transforms between these phases. The same is true for the transition between phase II and III, here the DOS at the transition point is shown in Figure 6.11 and Figure 6.12. The DOS of phase II at this point is the one at the point where the bandgap in the system closes due to the increased pressure. Both structures shows a deep minimum in the DOS at the Fermi level, indicating that the system might still be metallic at this point, especially if we consider the underestimation of the bandgap in DFT calculations. The DOS at the transition between phase III and IV is shown in Figure 6.13 and Figure 6.14. Here the system is clearly already metallic in both cases.

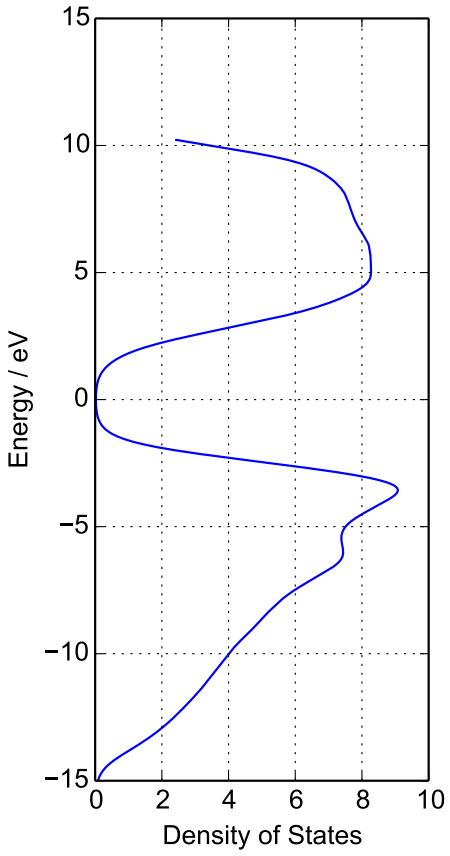

Figure 6.9: Density of states of phase I at 30 GPa.

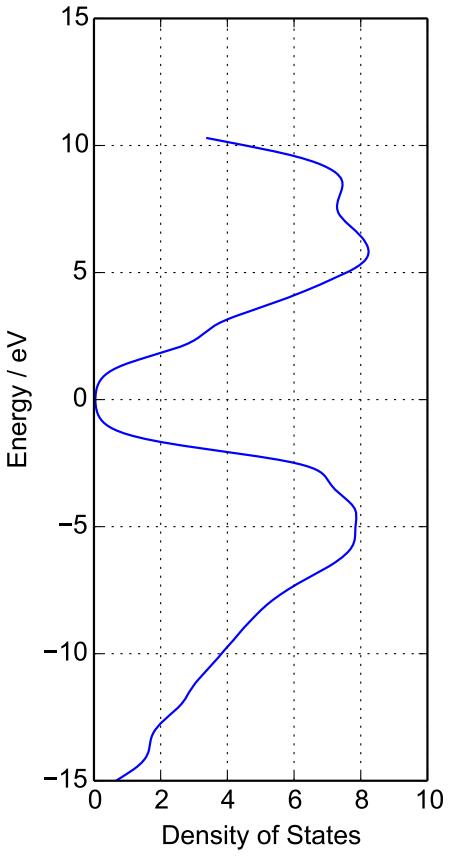

Figure 6.10: Density of states of phase II at 30 GPa.

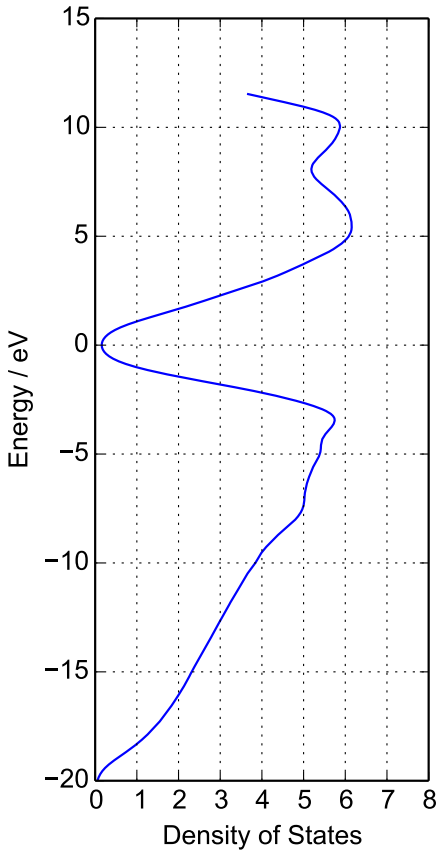

Figure 6.11: Density of states of phase II at 100 GPa.

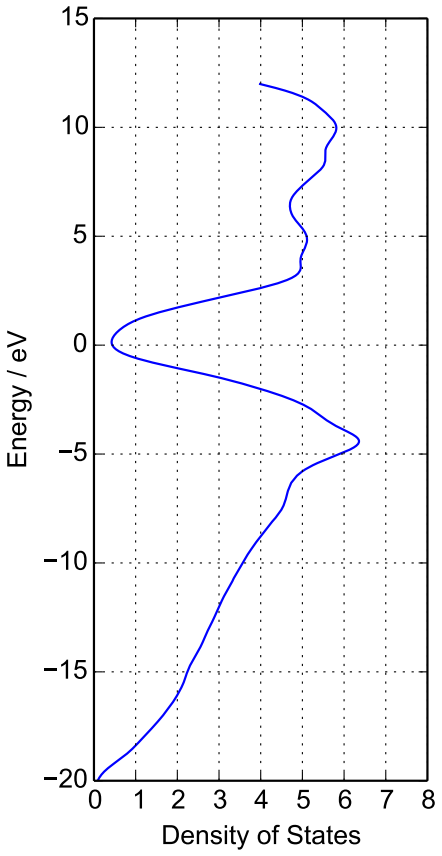

Figure 6.12: Density of states of phase III at 100 GPa.

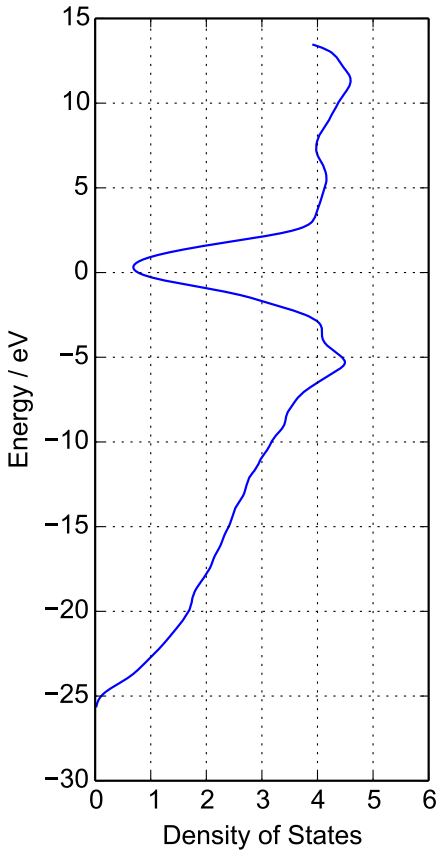

Figure 6.13: Density of states of phase III at 260 GPa.

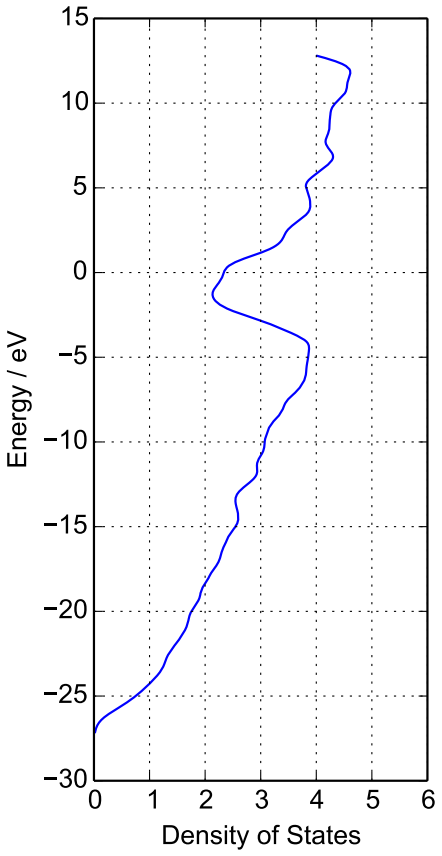

Figure 6.14: Density of states of phase IV at 260 GPa.

6.3 Conclusions

A metallization in SiH_{12} has been found at pressures around 100 GPa. This value is very low, even considering the underestimation of bandgap and metallization pressure by the here applied DFT method. Regardless of the higher hydrogen content these values are lower than the previously reported metallization pressure in $SiH_4(H_2)_2$, and provide evidence that compounds with rather high hydrogen contents pose promising systems for further analysis.

Bibliography

[1] http://en.wikipedia.org/wiki/File:Potential_Energy_Surface_and_Corresponding_Reaction_Coordinate_Diagram.png. Accessed: 2014-08-05.

[2] http://en.wikipedia.org/wiki/Diamond_anvil_cell#mediaviewer/File:Diamond_Anvil_Cell_-_Cross_Section.svg. Accessed: 2014-09-19.

[3] http://upload.wikimedia.org/wikipedia/commons/f/f7/Jupiter_interior.png. Accessed: 2014-10-24.

[4] Abascal, J. L. F., and Vega, C. A general purpose model for the condensed phases of water: Tip4p/2005. *J. Chem. Phys. 123*, 23 (2005), 234505.

[5] Alavi, A. What is the nature of the a-phase of solid hydrogen? *Philos. T. Roy. Soc. A* (1998), 263–266.

[6] Alavi, A., Parrinello, M., and Frenkel, D. Ab initio calculation of the sound velocity of dense hydrogen: implications for models of jupiter. *Science 269*, 5228 (1995), 1252–1254.

[7] Amsler, M., and Goedecker, S. Crystal structure prediction using the minima hopping method. *J. Chem. Phys. 133*, 22 (2010), 224104.

[8] Ashcroft, N. W. Metallic hydrogen: A high-temperature superconductor? *Phys. Rev. Lett. 21* (1968), 1748–1749.

[9] Ashcroft, N. W. Pressing some boundaries in mendeleev's chart. *Physics 2* (2009), 65.

[10] Barbee, T. W., Cohen, M. L., and Martins, J. L. Theory of high-pressure phases of hydrogen. *Phys. Rev. Lett. 62* (1989), 1150–1153.

[11] Behler, J., Martoňák, R., Donadio, D., and Parrinello, M. Pressure-induced phase transitions in silicon studied by neural network-based metadynamics simulations. *Phys. Status Solidi (b) 245*, 12 (2008), 2618–2629.

[12] Bell, R. P. The theory of reactions involving proton transfers. *Proc. R. Soc. Lond. 154* (1936), 414–429.

[13] Bloch, F. Bemerkung zur Elektronentheorie des Ferromagnetismus und der elektrischen Leitfähigkeit. *Z. Phys. 57*, 7-8 (1929), 545–555.

[14] Bonev, S. A., Schwegler, E., Ogitsu, T., and Galli, G. A quantum fluid of metallic hydrogen suggested by first-principles calculations. *Nature 431*, 7009 (2004), 669–672.

[15] Born, M., and Oppenheimer, R. Zur Quantentheorie der Molekeln. *Ann. Phys. 389*, 20 (1927), 457–484.

[16] Broyden, C. G. The convergence of a class of double-rank minimization algorithms 1. general considerations. *IMA J. Appl. Math. 6*, 1 (1970), 76–90.

[17] Bush, T. S., Catlow, C. R. A., and Battle, P. D. Evolutionary programming techniques for predicting inorganic crystal structures. *J. Mater. Chem. 5* (1995), 1269–1272.

[18] Ceperley, D. M., and Alder, B. J. Ground state of the electron gas by a stochastic method. *Phys. Rev. Lett. 45* (1980), 566–569.

[19] Ceperley, D. M., and Alder, B. J. Ground state of solid hydrogen at high pressures. *Phys. Rev. B 36* (1987), 2092–2106.

[20] Chacham, H., and Louie, S. G. Metallization of solid hydrogen at megabar pressures: A first-principles quasiparticle study. *Phys. Rev. Lett. 66* (1991), 64–67.

[21] Chen, N. H., Sterer, E., and Silvera, I. F. Extended infrared studies of high pressure hydrogen. *Phys. Rev. Lett. 76* (1996), 1663–1666.

[22] Daven, D., Tit, N., Morris, J., and Ho, K. Structural optimization of Lennard-Jones clusters by a genetic algorithm. *Chem. Phys. Lett. 256*, 1 (1996), 195–200.

[23] Deaven, D. M., and Ho, K. M. Molecular geometry optimization with a genetic algorithm. *Phys. Rev. Lett. 75* (1995), 288–291.

[24] Deem, M. W., and Newsam, J. M. Determination of 4-connected framework crystal structures by simulated annealing. *Nature 342* (1989), 260.

[25] Deemyad, S., and Silvera, I. F. Melting line of hydrogen at high pressures. *Phys. Rev. Lett. 100* (2008), 155701.

[26] Dirac, P. A. M. Note on exchange phenomena in the Thomas atom. *Math. Proc. Cambridge 26* (1930), 376–385.

[27] Doye, J. P. K., and Wales, D. J. Thermodynamics of global optimization. *Phys. Rev. Lett. 80* (1998), 1357–1360.

[28] Eremets, M., and Trojan, I. Evidence of maximum in the melting curve of hydrogen at megabar pressures. *JETP Lett. 89*, 4 (2009), 174–179.

[29] Eremets, M. I., Trojan, I. A., Medvedev, S. A., Tse, J. S., and Yao, Y. Superconductivity in hydrogen dominant materials: Silane. *Science 319* (2008), 1506–1509.

[30] Eremets, M. I., and Troyan, I. A. Conductive dense hydrogen. *Nat. Mater. 10* (2011), 927–931.

[31] Evans, M. G., and Polanyi, M. Inertia and driving force of chemical reactions. *Trans. Faraday Soc. 34* (1938), 11–24.

[32] Fletcher, R. A new approach to variable metric algorithms. *Comput. J. 13*, 3 (1970), 317–322.

[33] Frenkel, D., and Smit, B. *Understanding Molecular Simulation: From Algorithms to Applications*. Computational science series. Elsevier Science, 2001.

[34] Giannozzi, P., Baroni, S., Bonini, N., Calandra, M., Car, R., Cavazzoni, C., Ceresoli, D., Chiarotti, G. L., Cococcioni, M., Dabo, I., Dal Corso, A., de Gironcoli, S., Fabris, S., Fratesi, G., Gebauer, R., Gerstmann, U., Gougoussis, C., Kokalj, A., Lazzeri, M., Martin-Samos, L., Marzari, N., Mauri, F., Mazzarello, R., Paolini, S., Pasquarello, A., Paulatto, L., Sbraccia, C., Scandolo, S., Sclauzero, G., Seitsonen, A. P., Smogunov, A., Umari, P., and Wentzcovitch, R. M. QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials. *J. Phys.-Condens. Mat. 21*, 39 (2009), 395502 (19pp).

[35] Glass, C. W., Oganov, A. R., and Hansen, N. Uspex-evolutionary crystal structure prediction. *Comput. Phys. Commun. 175*, 11 (2006), 713–720.

[36] Goedecker, S. Minima hopping: An efficient search method for the global minimum of the potential energy surface of complex molecular systems. *J. Chem. Phys. 120*, 21 (2004), 9911–9917.

[37] Goedecker, S., Hellmann, W., and Lenosky, T. Global minimum determination of the born-oppenheimer surface within density functional theory. *Phys. Rev. Lett. 95* (2005), 055501.

[38] Goldberg, D. E. *Genetic Algorithms in Search, Optimization and Machine Learning*, 1st ed. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1989.

[39] Goldfarb, D. A family of variable-metric methods derived by variational means. *Math. Comp. 24* (1970), 23–26.

[40] Habershon, S., Markland, T. E., and Manolopoulos, D. E. Competing quantum effects in the dynamics of a flexible water model. *J. Chem. Phys. 131*, 2 (2009), 024501.

[41] Hellmann, W., Hennig, R. G., Goedecker, S., Umrigar, C. J., Delley, B., and Lenosky, T. Questioning the existence of a unique ground-state structure for si clusters. *Phys. Rev. B 75* (2007), 085411.

[42] Hemley, R. J., and Mao, H. K. Phase transition in solid molecular hydrogen at ultrahigh pressures. *Phys. Rev. Lett. 61* (1988), 857–860.

[43] Hemley, R. J., Mao, H.-k., Goncharov, A. F., Hanfland, M., and Struzhkin, V. Synchrotron infrared spectroscopy to 0.15 ev of h_2 and d_2 at megabar pressures. *Phys. Rev. Lett. 76* (1996), 1667–1670.

[44] Hestenes, M. R., and Stiefel, E. Methods of conjugate gradients for solving linear systems. *J. Res. Nat. Bur. Stand. 49* (1952), 409–436.

[45] Hohenberg, P., and Kohn, W. Inhomogeneous electron gas. *Phys. Rev. 136*, 3B (1964), B864.

[46] Holland, J. *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. University of Michigan Press, 1975.

[47] Johnson, K. A., and Ashcroft, N. Structure and bandgap closure in dense hydrogen. *Nature 403*, 6770 (2000), 632–635.

[48] Karaboga, D., and Basturk, B. A powerful and efficient algorithm for numerical function optimization: artificial bee colony (abc) algorithm. *J. Global Optim. 39*, 3 (2007), 459–471.

[49] Kaxiras, E., Broughton, J., and Hemley, R. J. Onset of metallization and related transitions in solid hydrogen. *Phys. Rev. Lett. 67* (1991), 1138–1141.

[50] Kirkpatrick, S., Gelatt, C. D., and Vecchi, M. P. Optimization by simulated annealing. *Science 220*, 4598 (1983), 671–680.

[51] Koch, W., and Holthausen, M. C. *A Chemist's Guide to Density Functional Theory*, 2. a. ed. Wiley-VCH, 2001.

[52] Kohanoff, J., Scandolo, S., de Gironcoli, S., and Tosatti, E. Dipole-quadrupole interactions and the nature of phase III of compressed hydrogen. *Phys. Rev. Lett. 83* (1999), 4097–4100.

[53] Kohn, W. Nobel lecture: Electronic structure of matter—wave functions and density functionals. *Rev. Mod. Phys. 71* (1999), 1253–1266.

[54] Kohn, W., and Sham, L. J. Self-Consistent equations including exchange and correlation effects. *Phys. Rev. 140*, 4A (1965), A1133.

[55] Laio, A., and Parrinello, M. Escaping free-energy minima. *P. Natl. Acad. Sci. 99*, 20 (2002), 12562–12566.

[56] Liu, D., and Nocedal, J. On the limited memory method for large scale optimization. *Math. Program. B 45*, 3 (1989), 503–528.

[57] Lorenzana, H. E., Silvera, I. F., and Goettel, K. A. Evidence for a structural phase transition in solid hydrogen at megabar pressures. *Phys. Rev. Lett. 63* (1989), 2080–2083.

[58] M. Woodley, S., D. Battle, P., D. Gale, J., and Richard A. Catlow, C. The prediction of inorganic crystal structures using a genetic algorithm and energy minimisation. *Phys. Chem. Chem. Phys. 1* (1999), 2535–2542.

[59] Maddox, J. Crystals from first principles. *Nature 335* (1988), 201.

[60] Mahoney, M. W., and Jorgensen, W. L. A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. *J. Chem. Phys. 112*, 20 (2000), 8910–8922.

[61] Malek, R., and Mousseau, N. Dynamics of Lennard-Jones clusters: A characterization of the activation-relaxation technique. *Phys. Rev. E 62* (2000), 7723–7728.

[62] Mao, N., and Hemley, R. Optical studies of hydrogen above 200 gigapascals-evidence for metallization by band overlap. *Science 244*, 19 (1989), 1462–1465.

[63] Martoňák, R., Laio, A., and Parrinello, M. Predicting crystal structures: The parrinello-rahman method revisited. *Phys. Rev. Lett. 90* (2003), 075503.

[64] Martoňák, R., Donadio, D., Oganov, A. R., and Parrinello, M. Crystal structure transformations in SiO_2 from classical and ab initio metadynamics. *Nat. Mater. 5*, 8 (2006), 623–626.

[65] Martoňák, R., Laio, A., Bernasconi, M., Ceriani, C., Raiteri, P., Zipoli, F., and Parrinello, M. Simulation of structural phase transitions by metadynamics. *Z. Kristallogr. 220* (2005), 489–498.

[66] Mestres, J., and Scuseria, G. E. Genetic algorithms: A robust scheme for geometry optimizations and global minimum structure problems. *J. Comput. Chem. 16*, 6 (1995), 729–742.

[67] Mousseau, N., and Barkema, G. T. Traveling through potential energy landscapes of disordered materials: The activation-relaxation technique. *Phys. Rev. E 57* (1998), 2419–2424.

[68] Niesse, J. A., and Mayne, H. R. Global geometry optimization of atomic clusters using a modified genetic algorithm in space-fixed coordinates. *J. Chem. Phys. 105*, 11 (1996), 4700–4706.

[69] Nocedal, J. Updating quasi-newton matrices with limited storage. *Math. Comput. 35* (1980), 773–782.

[70] Oganov, A. R., and Glass, C. W. Crystal structure prediction using ab initio evolutionary techniques: Principles and applications. *J. Chem. Phys. 124*, 24 (2006), 244704.

[71] Pannetier, J., J.Bassas-Alsina, Rodriguez-Carvajal, J., and Caignaert, V. Prediction of crystal structures from crystal chemistry rules by simulated annealing. *Nature 346*, 6282 (1990), 345.

[72] Parrinello, M., and Rahman, A. Crystal structure and pair potentials: A molecular-dynamics study. *Phys. Rev. Lett. 45* (1980), 1196–1199.

[73] Perdew, J. P., Burke, K., and Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett. 77*, 18 (1996), 3865.

[74] Perdew, J. P., and Levy, M. Physical content of the exact kohn-sham orbital energies: Band gaps and derivative discontinuities. *Phys. Rev. Lett. 51* (1983), 1884–1887.

[75] Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. *Numerical Recipes 3rd Edition: The Art of Scientific Computing*, 3 ed. Cambridge University Press, New York, NY, USA, 2007.

[76] Ramaker, D. E., Kumar, L., and Harris, F. E. Exact-exchange crystal Hartree-Fock calculations of molecular and metallic hydrogen and their transitions. *Phys. Rev. Lett. 34* (1975), 812–814.

[77] Ramzan, M., Lebègue, S., and Ahuja, R. Electronic structure and metalization of a silane-hydrogen system under high pressure investigated using density functional and gw calculations. *Phys. Rev. B 81* (2010), 233103.

[78] Richardson, C. F., and Ashcroft, N. W. High temperature superconductivity in metallic hydrogen: Electron-electron enhancements. *Phys. Rev. Lett. 78* (1997), 118–121.

[79] Richters, D., and Kühne, T. Liquid methane at extreme temperature and pressure: Implications for models of uranus and neptune. *JETP Lett. 97*, 4 (2013), 184–187.

[80] Richters, D., and Kühne, T. D. Self-consistent field theory based molecular dynamics with linear system-size scaling. *J. Chem. Phys. 140*, 13 (2014), 134109.

[81] Scandolo, S. Liquid-liquid phase transition in compressed hydrogen from first-principles simulations. *P. Natl. Acad. Sci. 100*, 6 (2003), 3051–3053.

[82] Schiffmann, C., and Sebastiani, D. Artificial bee colony optimization of capping potentials for hybrid quantum mechanical/molecular mechanical calculations. *J. Chem. Theory Comput. 7*, 5 (2011), 1307–1315.

[83] Schön, J. C., and Jansen, M. Auf dem weg zur syntheseplanung in der festkörperchemie: Vorhersage existenzfähiger strukturkandidaten mit verfahren zur globalen strukturoptimierung. *Angew. Chem. Ger. Edit. 108*, 12 (1996), 1358–1377.

[84] Schönborn, S. E., Goedecker, S., Roy, S., and Oganov, A. R. The performance of minima hopping and evolutionary algorithms for cluster structure prediction. *J. Chem. Phys. 130*, 14 (2009), 144108.

[85] Schrödinger, E. An undulatory theory of the mechanics of atoms and molecules. *Phys. Rev. 28*, 6 (1926), 1049.

[86] Shanno, D. F. Conditioning of quasi-newton methods for function minimization. *Math. Comp. 24* (1970), 647–656.

[87] Slater, J. C. A simplification of the Hartree-Fock method. *Phys. Rev. 81* (1951), 385–390.

[88] Sørensen, M. R., and Voter, A. F. Temperature-accelerated dynamics for simulation of infrequent events. *J. Chem. Phys. 112*, 21 (2000), 9599–9606.

[89] Städele, M., and Martin, R. M. Metallization of molecular hydrogen: Predictions from exact-exchange calculations. *Phys. Rev. Lett. 84* (2000), 6070–6073.

[90] Stepanenko, S., and Engels, B. Gradient tabu search. *J. Comput. Chem. 28* (2007), 601–611.

[91] Strobel, T. A., Somayazulu, M., and Hemley, R. J. Novel pressure-induced interactions in silane-hydrogen. *Phys. Rev. Lett. 103* (2009), 065701.

[92] Swope, W. C., Andersen, H. C., Berens, P. H., and Wilson, K. R. A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. *J. Chem. Phys. 76*, 1 (1982), 637–649.

[93] Szabo, A., and Ostlund, N. S. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. Dover Publications, 1996.

[94] Verlet, L. Computer "experiments" on classical fluids. I. thermodynamical properties of Lennard-Jones molecules. *Phys. Rev. 159* (1967), 98–103.

[95] Wales, D. J., and Doye, J. P. K. Global optimization by basin-hopping and the lowest energy structures of Lennard-Jones clusters containing up to 110 atoms. *J. Phys. Chem. A 101*, 28 (1997), 5111–5116.

[96] Wales, D. J., Doye, J. P. K., Dullweber, A., Hodges, M. P., Calvo, F. Y. N. F., Hernández-Rojas, J., and Middleton, T. F. The cambridge cluster database. `http://www-wales.ch.cam.ac.uk/home/www/httpd/htdocs/CCD.html`. Accessed: 2014-09-30.

[97] Wales, D. J., and Scheraga, H. A. Global optimization of clusters, crystals, and biomolecules. *Science 285*, 5432 (1999), 1368–1372.

[98] Wehmeyer, C., Falk von Rudorff, G., Wolf, S., Kabbe, G., Schärf, D., Kühne, T. D., and Sebastiani, D. Foraging on the potential energy surface:

A swarm intelligence-based optimizer for molecular geometry. *J. Chem. Phys. 137*, 19 (2012), 194110.

[99] Wigner, E., and Huntington, H. B. On the possibility of a metallic modification of hydrogen. *J. Chem. Phys. 3*, 12 (1935), 764–770.

[100] Willand, A., Gramzow, M., Alireza Ghasemi, S., Genovese, L., Deutsch, T., Reuter, K., and Goedecker, S. Structural metastability of endohedral silicon fullerenes. *Phys. Rev. B 81* (2010), 201405.

[101] Woodley, S. M., and Catlow, R. Crystal structure prediction from first principles. *Nat. Mater. 7* (2008), 937 – 946.

Declaration

I hereby declare that I wrote the dissertation submitted without any unauthorized external assistance and used only sources acknowledged in the work. All textual passages which are appropriated verbatim or paraphrased from published and unpublished texts as well as all information obtained from oral sources are duly indicated and listed in accordance with bibliographical rules. In carrying out this research, I complied with the rules of standard scientific practice as formulated in the statutes of Johannes Gutenberg-University Mainz to insure standard scientific practice.

Mainz, 29. 10. 2014

Daniel Schärf