



Andreas Jeindl, BSc.

# **Influence of Molecule Size on Surface Polymorph Formation: An ab-initio Study with Machine Learning**

## **MASTERARBEIT**

zur Erlangung des akademischen Grades

Diplom-Ingenieur

Masterstudium Technische Physik

eingereicht an der

**Technischen Universität Graz**

Betreuer

Dr.techn. Oliver T. Hofmann  
Univ.-Prof. Dr. Peter Hadley

Institut für Festkörperphysik

## **EIDESSTATTLICHE ERKLÄRUNG**

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe. Das in TUGRAZonline hochgeladene Textdokument ist mit der vorliegenden Masterarbeit identisch.

---

Datum

---

Unterschrift

## Acknowledgment

First of all, I want to thank all the people that helped me throughout the process of writing this thesis. I would especially like to express my gratitude towards my supervisor Oliver Hofmann, who led the way, but still allowed me to go towards the direction that fascinated me the most.

Without the preliminary work and help of my co-workers, especially Michael, but also Lukas and Alex, this thesis would not be what it is today, thank you for that. The fruitful discussions with the members of the advanced materials modeling team and the other members of the Institute of Solid State Physics are also experiences I do not want to miss.

For the many CPU hours spent, I want to thank the Vienna Scientific Cluster and the Argonne Leadership Computing Facility. Financial support by the Austrian Science Fund was received with great appreciation.

The student days would not have been the same without my dearest colleagues: Berni, Cesi, Sascha, Lukas, Flo, Sarah, Mari and many more. Thank you for the countless hours spent learning, discussing and having fun.

Finishing my studies could not have been accomplished without my family. Thank you for supporting me unreservedly over the years.

Finally, I want to wholeheartedly thank the person that has accompanied me throughout most of the years at university: Thank you Lisa.

Abstract

# Influence of Molecule Size on Surface Polymorph Formation: An *ab-initio* Study with Machine Learning

Andreas Jeindl

*Institute of Solid State Physics, Graz University of Technology*

A variety of properties of metal-organic interfaces are determined by the polymorph that the molecules assume on the (metal) surface. Intriguingly, some properties of such surface-induced phases even exceed those of the bulk crystal. Thus, for the engineering of functional interfaces it is crucial to understand and predict the polymorphs molecules assume on a given surface. Still, there are surprisingly few systematic studies that allow to infer reliable relationships between the molecular structure and the kinds of interface polymorphs that form. The reason for this is, that in experiments often kinetically trapped phases form, which prohibit the systematic investigation of thermodynamic relationships. Computational studies are prohibited by the "configurational explosion" which leads to an intractably vast number of possible polymorphs.

In this thesis, the influence of molecule size on the formation of surface polymorphs is investigated systematically with the example of acenequinones on Ag(111). Starting from the small benzoquinone, the analysis is continued via anthraquinone to the larger pentacenequinone. An exhaustive structure search is performed utilizing the SAMPLE approach, which uses a combination of physically-motivated coarse graining of the configurational potential energy surface and machine learning in the form of Bayesian linear regression to overcome the configurational explosion. The adjustments and improvements to the method, developed in the course of this thesis to enable the prediction of larger molecules with functional groups, are presented. The influence of the size of single molecules on their interaction with the surface is elucidated through the investigation of preferred adsorption sites for single molecules on the substrate. To complete the picture, the low energy polymorphs and surface patterns for benzoquinone and anthraquinone are shown. The driving forces for their formation are illuminated by representation of the complex interactions via molecule-substrate and molecule-molecule contributions.

**Einfluss der Molekülgröße auf die Bildung von Oberflächenpolymorphen:  
Eine *ab-initio* Studie mit maschinellem Lernen**

Andreas Jeindl

*Institut für Festkörperphysik, Technische Universität Graz*

Eine Vielzahl der physikalischen Eigenschaften von dünnen Filmen organischer Materialien auf Metallen werden von der Anordnung der Moleküle auf der Oberfläche dominiert. Bemerkenswerterweise sind einige Eigenschaften von solchen oberflächeninduzierten Phasen besser als jene im Molekülkristall. Zur Entwicklung funktioneller Grenzflächen ist ein Verständnis des Zusammenhangs zwischen Molekülstruktur und dem Polymorph, das sich auf der Oberfläche formt, daher von großem Interesse. Leider existieren jedoch nur sehr wenige systematische Studien, die einen klaren Zusammenhang zwischen Molekülstruktur und Grenzflächenpolymorph schließen lassen. Die Gründe dafür sind einerseits, dass bei experimentellen Studien oftmals kinetisch stabilisierte Phasen entstehen, welche die Untersuchung der thermodynamischen Zusammenhänge verhindern. Simulationen andererseits sind durch die riesige Anzahl möglicher Strukturen, auch bekannt als Konfigurationsexplosion, beschränkt.

Ziel dieser Arbeit ist die Untersuchung des Zusammenhangs zwischen der Molekülgröße und der Polymorphe, die sich auf der Oberfläche bilden. Als Mustersystem dafür dienen Chinone unterschiedlicher Größe auf der Silber (111) Oberfläche. Zur Überwindung der Konfigurationsexplosion wird die SAMPLE Methode vorgestellt, welche eine physikalisch motivierte Diskretisierung der Potentialoberfläche mit Bayesscher Statistik kombiniert. Die Anpassungen und Weiterentwicklungen der Methode, welche im Rahmen dieser Arbeit für die korrekte Vorhersage großer Moleküle mit funktionellen Gruppen entwickelt wurden, werden erläutert. Die Abhängigkeit der dominierenden Wechselwirkungen zwischen Molekül und Oberfläche von der Größe der Moleküle wird anhand der präferierten lokalen Adsorptionsgeometrien einzelner Moleküle dargestellt. Abgerundet wird die Betrachtung mit den energetisch günstigsten Polymorphen und Oberflächenstrukturen von Benzochinon- und Anthrachinon-Monolagen. Die treibenden Kräfte für die Bildung jener Strukturen werden durch Aufteilung der Energien in Molekül-Substrat und Molekül-Molekül Wechselwirkungen veranschaulicht.

*I may not have gone  
where I intended to go,  
but I think I have ended up  
where I needed to be.*

DOUGLAS ADAMS

# Contents

<b>1. Introduction</b>	<b>1</b>
1.1. Quantum Mechanical Basics . . . . .	1
1.1.1. Born Oppenheimer Approximation . . . . .	2
1.1.2. Hartree-Fock Theory . . . . .	3
1.1.3. Basis Sets . . . . .	5
1.1.4. Density Functional Theory . . . . .	7
1.1.5. Exchange-Correlation Functionals . . . . .	8
1.1.6. Errors Inherent to Common Density Functional Approximations . . . . .	11
1.1.7. Periodic Quantum Mechanical Calculations . . . . .	12
1.2. Interactions of Organic Molecules and Surfaces . . . . .	15
1.3. Structure Search . . . . .	16
1.4. Computational Structure Search . . . . .	17
1.4.1. Linear Regression Methods . . . . .	17
1.4.2. Bayesian Linear Regression . . . . .	18
1.5. The SAMPLE Approach . . . . .	19
1.5.1. Step 1: Generation of Local Geometries . . . . .	20
1.5.2. Step 2: Combination to Configurations . . . . .	20
1.5.3. Step 3: Learning of Interactions . . . . .	21
1.5.4. Step 4: Prediction . . . . .	23
<b>2. Making Tractable</b>	<b>25</b>
2.1. Quantum Mechanical Method Used . . . . .	25
2.2. Mixed Basis Set Method . . . . .	25
2.2.1. Application to Ag(111) . . . . .	26
2.3. Local Adsorption Geometry Search . . . . .	29
2.4. Improvements to the Learner . . . . .	30
2.4.1. New Feature Vector . . . . .	30
2.4.2. Choice of Minimal Distances . . . . .	31
2.4.3. New Covariance Kernels . . . . .	32
2.4.4. Improvements of the Prior . . . . .	33
2.4.5. Comparison of the Improvements . . . . .	34
<b>3. Application</b>	<b>35</b>
3.1. Systems of Interest . . . . .	35
3.2. Local Geometry Search . . . . .	36
3.2.1. Search Procedure . . . . .	36
3.2.2. Results for Benzoquinone . . . . .	37
3.2.3. Results for Anthraquinone . . . . .	38
3.2.4. Results for Pentacenequinone . . . . .	39
3.3. Structure Search . . . . .	41
3.3.1. Preparations . . . . .	41
3.3.2. Inspection of the Learning Behavior . . . . .	42
3.3.3. Predicted Configurations . . . . .	44
3.3.4. Explanation of the Best Structures . . . . .	47
<b>Conclusion</b>	<b>50</b>

<b>Appendices</b>	<b>51</b>
A. Convergence of DFT Settings for an Ag Bulk . . . . .	51
B. Convergence Tests for the Adsorption Energy of Pentacenequinone . . . . .	53
C. Computational Details of the Mixed Basis Set Method . . . . .	56
D. Electrostatic Prior . . . . .	60
E. Minimal Distances . . . . .	63
F. Interpolation of the Rotational PES . . . . .	64
G. Local Geometry Rotation Sweeps . . . . .	65
H. Hyper Parameter Tests . . . . .	67
I. Best Structures of Anthraquinone . . . . .	68
<b>List of Figures</b>	<b>70</b>
<b>Bibliography</b>	<b>72</b>

# 1. Introduction

In this chapter the theoretical basics of this thesis are introduced. First the quantum mechanical methodology and its limitations are presented. Then an overview of typical interactions of molecules with surfaces followed by an introduction to surface poly morphs is given. In the last part the structure search method and some approaches to solve the linear model, which is at the core of the method, are explained.

ack

## 1.1. Quantum Mechanical Basics

The following section deals with some quantum mechanical methods, which are necessary for a full *ab-initio* modeling of organic/inorganic interfaces. It is mainly based on the book *Introduction to Computational Chemistry* by Frank Jensen [1] in the third edition. Other sources will be mentioned explicitly.

### Fundamentals

Starting point for quantum mechanical calculations is always the Schrödinger equation [2], whose time independent shorthand operator form is:

$$\mathbf{H}_{tot}|\Psi\rangle = E|\Psi\rangle \quad (1.1)$$

This is an eigenvalue equation, which has no analytic solution for more than two particles. To still find a solution, approximations are applied. The total non-relativistic Hamilton operator can be written as:

$$\mathbf{H}_{tot} = \mathbf{T}_n + \mathbf{T}_e + \mathbf{V}_{ne} + \mathbf{V}_{ee} + \mathbf{V}_{nn} \quad (1.2)$$

Here  $\mathbf{T}$  represents the kinetic and  $\mathbf{V}$  the potential energy operators, the subscript index  $n$  denotes nuclei and  $e$  electrons. For sufficient description of a quantum mechanical system, a suitable wave function  $\Psi$  has to be found. This can be done with the help of the variational principle, which states that the total energy  $E$  obtained by evaluation of the Hamilton operator with any approximate wave function will always be higher than the ground state energy  $E^0$  of the system, unless the wave function is optimal.

$$E = \frac{\langle\Psi|\mathbf{H}|\Psi\rangle}{\langle\Psi|\Psi\rangle} \geq E^0 \quad (1.3)$$

Therefore, trial wave functions can be constructed and improved successively by minimizing the total energy. For normalized wave functions the denominator becomes unity and the equation is reduced to

$$E = \langle\Psi|\mathbf{H}|\Psi\rangle \geq E^0 \quad (1.4)$$

Due to the coupling of nuclei and electrons with themselves and each other, the Schrödinger equation is very hard to solve. Therefore some approximations are made, which are explained in the following pages.

### 1.1.1. Born Oppenheimer Approximation

One first approximation is to separate nuclear and electronic wave functions. Therefore the full Hamilton operator 1.2 is first transformed to the center of mass system:

$$\mathbf{H}_{tot} = \underbrace{\mathbf{T}_n + \mathbf{T}_e + \mathbf{V}_{ne} + \mathbf{V}_{ee} + \mathbf{V}_{nn}}_{=\mathbf{H}_e} - \underbrace{\frac{1}{2M_{tot}} \left( \sum_i^{N_{elec}} \nabla_i \right)^2}_{=\mathbf{H}_{mp}} \quad (1.5)$$

with the electronic Hamiltonian operator  $\mathbf{H}_e$  and the mass-polarization  $\mathbf{H}_{mp}$ . This mass-polarization arises, because for systems with more than two particles a rigorous separation of center-of-mass motion from internal motion is not possible.  $\mathbf{H}_e$  now only depends on the positions of the nuclei, but not on their momenta. If the full set of solutions for the electronic Schrödinger equation is now assumed to be known, equation 1.1 can be rewritten to

$$\mathbf{H}_e(\mathbf{R})\Psi_i(\mathbf{R}, \mathbf{r}) = E_i(\mathbf{R})\Psi_i(\mathbf{R}, \mathbf{r}); \quad i = 1, 2, \dots, \infty \quad (1.6)$$

Because of the hermiticity of the Hamiltonian, the  $\Psi_i$  can be chosen to be orthonormal with respect to each other. Now the total wave function can be written as an expansion in the set of electronic functions, where the expansion coefficients are functions of nuclear coordinates

$$\Psi_{tot}(\mathbf{R}, \mathbf{r}) = \sum_{i=1}^{\infty} \Psi_{ni}(\mathbf{R})\Psi_i(\mathbf{R}, \mathbf{r}) \quad (1.7)$$

Inserting this wave function into the Schrödinger equation leads to

$$\sum_{i=1}^{\infty} (\mathbf{T}_n + \mathbf{H}_e + \mathbf{H}_{mp})\Psi_{ni}(\mathbf{R})\Psi_i(\mathbf{R}, \mathbf{r}) = E_{tot} \sum_{i=1}^{\infty} \Psi_{ni}(\mathbf{R})\Psi_i(\mathbf{R}, \mathbf{r}) \quad (1.8)$$

This equation is then multiplied from the left by a specific electronic wave function  $\Psi_j^*$  and integrated over the electron coordinates. Neglecting the non-adiabatic coupling elements (coupling of different electronic states with  $\nabla_n$ ) and neglecting the mass polarization leads to

$$(\mathbf{T}_n + E_j + \underbrace{\langle \Psi_j | \nabla_n^2 | \Psi_j \rangle}_{=U(\mathbf{R})})\Psi_{nj} = E_{tot}\Psi_{nj} \quad (1.9)$$

$E_j$  is the energy of the j-th electronic wave function.  $U(\mathbf{R})$  is called diagonal correction and is usually a very small, slowly with  $R_i$  varying function and therefore neglected in the Born Oppenheimer approximation, which can now be rewritten as

$$(\mathbf{T}_n + E_j)\Psi_{nj} = E_{tot}\Psi_{nj} \quad (1.10)$$

This equation describes nuclei moving on the potential energy surfaces given by the solution of the electronic Schrödinger equation. As the coupling terms of different electronic states are neglected, the approximation breaks down, as soon as different solutions to the Schrödinger equation come close energetically (in other words when potential energy surfaces overlap).

Now that the nuclear motion is separated, a solution for the electronic Schrödinger equation can be approached.

### 1.1.2. Hartree-Fock Theory

To fulfill the antisymmetry requirement for fermionic wave functions (Pauli principle) the total wave function can be built from Slater determinants with single-electron wave functions (orbitals) as matrix elements. When calculating molecules, these one-electron wave functions, which each are a product of spatial and spin orbitals, are interpreted as molecular orbitals. One further approximation made in Hartree-Fock, is taking only a single determinant for the trial wave function and thereby neglecting electron-electron correlation completely. This leads to a formalism which describes independent electrons moving in the field generated by all particles. Such methods are called mean field methods.

The full electronic Hamiltonian is

$$\mathbf{H}_e = \mathbf{T}_e + \mathbf{V}_{ne} + \mathbf{V}_{ee} + \mathbf{V}_{nn} \quad (1.11)$$

$$= \sum_i^{N_{elec}} \underbrace{\left( -\frac{1}{2} \nabla_i^2 - \sum_A^{N_{nuclei}} \frac{Z_A}{|\mathbf{R}_A - \mathbf{r}_i|} \right)}_{\mathbf{h}_i} + \sum_{j>i}^{N_{elec}} \underbrace{\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}}_{\mathbf{g}_{ij}} + \mathbf{V}_{nn} \quad (1.12)$$

Here the one-electron operator  $\mathbf{h}_i$  describes the motion of a single electron in the field of all nuclei.  $\mathbf{g}_{ij}$  represents the electron-electron repulsion. Inserting the electronic Hamiltonian and the Slater-type wave function into 1.4 leads to an energy of the form

$$E = \sum_{i=1}^{N_{elec}} \langle \psi_i(1) | \mathbf{h}_i | \psi_i(1) \rangle \quad (1.13)$$

$$+ \sum_{i=1}^{N_{elec}} \sum_{j>i}^{N_{elec}} (\langle \psi_i(1) \psi_j(2) | \underbrace{\mathbf{g}_{12}}_{\mathbf{J}_{ij}} | \psi_i(1) \psi_j(2) \rangle - \langle \psi_i(1) \psi_j(2) | \underbrace{\mathbf{g}_{12}}_{\mathbf{K}_{ij}} | \psi_j(1) \psi_i(2) \rangle) + V_{nn} \quad (1.14)$$

$$E = \sum_{i=1}^{N_{elec}} h_i + \sum_{i=1}^{N_{elec}} \sum_{j>i}^{N_{elec}} (J_{ij} - K_{ij}) + V_{nn} \quad (1.15)$$

The operators  $\mathbf{J}_i$  and  $\mathbf{K}_i$  with the corresponding energies  $J_{ij}$  and  $K_{ij}$  are called Coulomb and exchange operators. The exchange operator arises due to the (anti)symmetry of the wave functions of indistinguishable particles upon exchange. As a next step, a set of orthonormal molecular orbitals that minimizes the energy or at least makes it stationary with respect to a change in molecule orbitals needs to be found. Carrying out this optimization with Lagrange multipliers leads to the following set of Hartree-Fock equations

$$\mathbf{F}_i \phi'_i = \sum_j^{N_{elec}} \lambda_{ij} \phi'_j \quad (1.16)$$

$$\mathbf{F}_i = \mathbf{h}_i + \sum_j^{N_{elec}} (\mathbf{J}_j - \mathbf{K}_j) \quad (1.17)$$

where the Lagrange multipliers  $\lambda_{ij}$  are the elements of a Hermitian matrix  $\Lambda$ .  $\mathbf{F}_i$  is the so-called Fock-operator. Transformation of the molecular orbitals to a set of orbitals, which makes  $\Lambda$  diagonal (canonical MOs), leads to a set of pseudo-eigenvalue equations for the orbitals

$$\mathbf{F}_i \phi_i = \epsilon_i \phi_i \quad (1.18)$$

The eigenvalues  $\epsilon_i$  are interpreted as the energies of the molecular orbitals. Equation 1.18 is not a real eigenvalue equation because the Coulomb and exchange elements of the Fock operator still depend on all occupied orbitals. Therefore an iterative self-consistent field (SCF) method must be used to determine all MOs and their energies.

The total energy of the system (equation 1.15) can now also be written with respect to the eigenenergies of the orbitals

$$E = \sum_i^{N_{elec}} \epsilon_i - \frac{1}{2} \sum_{ij}^{N_{elec}} (J_{ij} - K_{ij}) + V_{nn} \quad (1.19)$$

$$\epsilon_i = \langle \psi_i | \mathbf{F}_i | \psi_i \rangle = h_i + \sum_j^{N_{elec}} (J_i - K_i) \quad (1.20)$$

The Coulomb and exchange terms have to be subtracted as they are counted twice in the calculation of the MO eigenvalues.

HF is not the exact solution of the Schrödinger equation as the electron-electron repulsion is only treated in an averaged manner. This is called the *Hartree-Fock limit*.

### Koopmans' Theorem

If it is assumed, that for a system with  $N$  electrons and a system with one electron removed or added from orbital  $k$  the MOs are identical, then the ionization energy or electron affinity is simply given as the orbital energy  $\epsilon_k$ . This is also called "frozen MO" approximation. When expanding the MOs in a set of basis functions this approach is reasonable for ionization energies. Using it for electron affinities is debatable as the physical meaning of the unoccupied orbitals is often questionable, especially when the lowest unoccupied molecular orbital lies above the vacuum level of the system.

### Beyond Hartree-Fock

To describe the electron correlation, neglected by Hartree-Fock, a series of different methods have been developed. All of those methods include electron correlation by adding more Slater determinants to the calculation. This leads to a multideterminant wave function of the form

$$\Psi = a_0 \Phi_{HF} + \sum_{i=1} a_i \Phi_i \quad (1.21)$$

A detailed description of some of those methods can be found in *Introduction to Computational Chemistry*.

### 1.1.3. Basis Sets

For all but very small systems, the Hartree-Fock equations can not be solved by directly mapping out the molecular orbitals in real space. Therefore the orbitals are expanded into basis functions  $\chi_\alpha$

$$\phi_i = \sum_{\alpha}^{N_{basis}} c_{\alpha i} \chi_{\alpha} \quad (1.22)$$

Now equation 1.18 can be rewritten as

$$\mathbf{F}_i \sum_{\alpha}^{N_{basis}} c_{\alpha i} \chi_{\alpha} = \epsilon_i \sum_{\alpha}^{N_{basis}} c_{\alpha i} \chi_{\alpha} \quad (1.23)$$

Multiplication from the left and integration leads to the so-called Roothan-Hall equations, which can be written in matrix notation

$$\mathbf{FC} = \mathbf{SC}\epsilon \quad (1.24)$$

where  $\mathbf{F}$  contains the Fock-matrix elements and  $\mathbf{S}$  are the overlap matrix elements.  $\mathbf{C}$  is the coefficient matrix for the basis functions and  $\epsilon$  is the matrix of orbital energies. To solve those equations, the Fock matrix has to be diagonalized. As this is only possible if all basis function coefficients are known, again a self-consistent field procedure has to be employed. The solution of the Roothan-Hall equation delivers  $N_{basis}$  orbitals, where  $N_{elec}$  orbitals are occupied and  $N_{basis} - N_{elec}$  are virtual orbitals, which have no direct physical interpretation.

Ideally the basis functions are constructed such, that an increase of basis set size systematically improves the accuracy of the orbitals. Usage of a complete basis set would therefore be equivalent to a direct mapping of orbitals, which is still only accurate up to the limits of the underlying method.

Basis functions should be designed hierarchical, to allow a systematic improvement towards this limit. Furthermore they should reflect the nature of the system and be available for all atom species of interest. For a universal applicability they should also be suitable for different methods and properties.

### Types of Orbitals

There are two types of analytic orbitals which are commonly used to model the atomistic electron distribution. The first are **Slater-type orbitals (STOs)**[3]

$$\chi_{\zeta,n,l,m}(r, \theta, \phi) = N Y_{l,m}(\theta, \phi) r^{n-1} e^{-\zeta r} \quad (1.25)$$

Here  $Y_{l,m}$  are spherical harmonics and  $N$  is a normalization constant.  $n, l$  and  $m$  resemble the principal, azimuthal and magnetic quantum numbers. The parameter  $\zeta$  determines the spread of the basis function and is important for the mixing of different basis functions. The mathematical form of Slater-type orbitals is very close to hydrogen-like orbitals, which ensures a rapid convergence with increasing number of functions, but analytic calculations of 2- and 4-center integrals are not possible due to the mathematical form.

Therefore, often **Gaussian-type orbitals (GTOs)**[4] are used.

$$\chi_{\zeta,n,l,m}(r, \theta, \phi) = N Y_{l,m}(\theta, \phi) r^{2n-2-l} e^{-\zeta r^2} \quad (1.26)$$

They do not resemble the analytic hydrogen orbitals that well, as the slope near the nucleus is too small and further away too large, which means that more GTOs (roughly 3-times as many) are needed to achieve the same level of accuracy as STOs. The main benefit however is, that calculations of

multi-center integrals become easy, as the Gaussians always transform into Gaussians. This more than compensates the increase in basis functions.

For computational applications, where the resulting integrals have to be solved numerically anyway, the basis functions can be numerical as well. One example for such **numerical atomic orbitals (NAOs)** are the ones used in the program package FHI-aims [5], that is also used in this work

$$\chi_i(r, \theta, \phi) = \frac{u_i(r)}{r} Y_{l,m}(\theta, \phi) \quad (1.27)$$

This is the formalism of the original publication, where the different quantum numbers (l,m) are coded into the index i.  $Y_{l,m}$  are again spherical harmonics. The radial function is numerically tabulated for each basis function, which has the advantage that the physical form of the atom orbital can be modeled very accurate. Because of their numerical form, the basis functions are only evaluated to a certain distance from the atom core.

## Size of the Basis Set

As soon as the type of basis functions is chosen, the size of the basis set has to be decided. For atom centered orbitals there exists a systematic classification for this task. This classification often refers to the exponent zeta used for STOs and GTOs. If each atom orbital is only modeled by a single set of basis functions with just one free parameter, it is called minimal or single zeta (SZ) basis set. A first improvement over this basis, is using a second set of basis functions for all occupied orbitals, which is commonly referred to as double zeta (DZ) basis and allows a better description of different bonding behavior ( $\sigma$  and  $\pi$  bonds) in molecules. Because normally only the valence electrons are involved in molecular bonds, often only the valence orbitals are doubled which is then called split valence basis. The next improvement is using three, four, five or more sets of basis functions for each occupied orbital, which is then called triple, quadruple, quintuple, . . . , -zeta basis. For the correct description of anions and dipole moments, basis functions with small exponents, which spread out far, are needed. They are called diffuse functions. Another significant improvement can be accomplished by adding functions for orbitals with higher angular momenta, which are denoted polarization functions. For Hartree-Fock the inclusion of a single set of polarization functions is often already sufficient. When using methods that calculate the electron correlation explicitly, many more polarization functions have to be used to account for the angular correlation.

When calculating hydrogen for example, the minimal basis consists of a single orbital for which a single parameter has to be optimized. The double zeta basis set adds a second orbital. Increasing the zeta by one always adds one new orbital. Using diffuse functions also adds one new orbital for each occupied orbital in the atom (1 for the case of hydrogen). Adding polarization functions means for hydrogen, that p-orbitals are included for the calculation, leading to three new orbitals.

## Plane Wave Basis Functions

In contrast to reconstructing the molecular orbitals of the system, here the full system is modeled with suitable functions. For periodic metals the valence electrons behave almost like free electrons, which suggests using solutions for the free electrons as basis functions for infinite systems. For such systems molecular levels coalesce into bands which can be written as

$$\chi_k(r) = e^{ikr} \quad (1.28)$$

with the corresponding energy

$$E = \frac{1}{2}k^2 \quad (1.29)$$

The allowed values for  $\mathbf{k}$  are determined by the unit cell translation vector  $\mathbf{t}$  via  $\mathbf{k}\mathbf{t} = 2\pi m$  with a positive integer  $m$ . Here one can see that the number of plane waves is determined by the highest  $\mathbf{k}$ -vector, which also defines the maximum kinetic energy via equation 1.29.

The quality (and size) of the basis set is often defined by a cutoff energy  $E_{cut}$  which is typically in the range of several hundred eV. One major benefit of plane waves is, that their basis set size only depends on the size of the unit cell and the cutoff energy, but not on the number of atoms within the system, which makes it attractive for larger systems. This benefit however is also a drawback when calculating surfaces and non-interacting molecules, which have to be separated by large distances to prevent interactions with periodic replicas. There a lot of basis functions are used solely to model the vacuum region, where no physical processes occur. Plane waves are very good at describing strongly delocalized electron densities with slow local variations but require a large number of heavily oscillating functions to model the core regions, leading to a basis with very large  $E_{max}$ . The core singularities of electron densities are also almost impossible to model, which is why plane waves are always used in combination with pseudopotentials. Those are analytic functions (e.g. Bessel Functions), which replace the full description of the electrons inside a specified core radius  $r_c$ . The functions have to be designed such, that at  $r_c$  the functions, as well as their first and second derivatives, are continuous.

### Basis Set Superposition Error

When using an atom centered basis set, such as GTOs, several hundred functions per atom would be needed to reach chemical accuracy (1 kcal/mol or around 40 meV) for the total energy. But usually one is only interested in energy differences, which can be achieved faster by choosing a balanced basis set with almost constant basis set error. When calculating a molecule with such an underconverged basis set, the electron distribution around the core can be improved with basis functions of another atom. This effect changes with different geometries, leading to a basis set error which is commonly called basis set superposition error (BSSE).

#### 1.1.4. Density Functional Theory

Hohenberg and Kohn have shown, that the ground state electron energy is fully determined by the electron density [6]. As this is the integral of the squared wave function over  $N-1$  electrons, the dimensionality of the problem can in principle be reduced from  $4N$  (3 spacial and one spin component for each electron) to 3 dimensions for spin up and 3 for spin down. The problem hereby is, that the functional that connects the electron density to the energy  $E = E\{\rho(\mathbf{r})\}$  is not known. From Hartree-Fock theory it seems reasonable to divide the energy functional into 4 terms: the kinetic energy  $T$ , the electron-nucleus attraction  $E_{ne}$ , Coulomb interaction  $J$  and electron-electron correlation  $K$

$$E[\rho] = T[\rho] + E_{ne}[\rho] + J[\rho] + K[\rho] \quad (1.30)$$

First theories, such as Thomas-Fermi and Thomas-Fermi-Dirac, tried to calculate the total energy directly from the electron density. Those so-called orbital-free DFT methods, which could make full use of the benefit of having only three variables independent of system size, lead to a significantly worse behavior compared to HF. The main flaw in those methods lies in the representation of the kinetic energy. Therefore Kohn and Sham [7] reintroduced a set of orthonormal orbitals and split the kinetic energy term into two parts, an exact solution for non-interacting particles (identical to HF) and a small correction for the kinetic electron correlation. This electron correlation is adsorbed into an exchange correlation functional  $E_{xc}$ . The total DFT energy functional can now be written as

$$E_{DFT}[\rho] = T_{HF}[\rho] + E_{ne}[\rho] + J[\rho] + E_{xc}[\rho] + \int V_{ext}(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} \quad (1.31)$$

$$E_{xc}[\rho] = (T[\rho] - T_{HF}[\rho]) + (E_{ee}[\rho] - J[\rho]) \quad (1.32)$$

In the second equation, the first bracket represents the kinetic correlation energy as the difference between the real kinetic energy and the kinetic energy in the Hartree-Fock formalism. The second bracket contains the potential correlation and exchange energy.

Here it becomes clear, that for Kohn-Sham (KS) DFT only functionals for the small correction  $E_{xc}$  have to be derived whereas for orbital free DFT, the kinetic and correlation energy functionals must be found. With KS-DFT methods the dimensionality of the problem is now again  $6N$ , but computationally much less expensive than post-HF because correlation effects are not calculated explicitly.

Once the exchange-correlation functional is chosen, a set of orthogonal orbitals that minimizes the energy needs to be found. Here the relevant equations are

$$\mathbf{h}_{KS}\phi_i = \sum_j^{N_{orb}} \lambda_{ij}\phi_j \quad (1.33)$$

$$\mathbf{h}_{KS} = \frac{1}{2}\nabla^2 + \mathbf{v}_{eff} \quad (1.34)$$

$$\mathbf{v}_{eff} = \mathbf{V}_{ne}(\mathbf{r}) + \underbrace{\int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}'}_{=J[\rho]} + \mathbf{V}_{xc}(\mathbf{r}) \quad (1.35)$$

$\mathbf{h}_{KS}$  now has the same role as the Fock Operator in HF. The set of orbitals can again be diagonalized

$$\mathbf{h}_{KS}\phi_i = \epsilon_i\phi_i \quad (1.36)$$

By expanding the KS orbitals in an atomic basis set, the problem formally becomes identical to equation 1.24

$$\mathbf{h}_{KS}\mathbf{C} = \mathbf{S}\mathbf{C}\boldsymbol{\epsilon} \quad (1.37)$$

The matrix elements of  $\mathbf{h}_{KS}$  for the kinetic, one-electron and Coulomb parts are identical to the corresponding elements in  $\mathbf{F}$ , but the exchange-correlation part  $\mathbf{V}_{xc}$  depends on the electron density and possibly also its derivatives. The evaluation of KS matrix elements for  $\mathbf{V}_{xc}$  is difficult, because of its implicit dependence on the integration variables via the electron density. For numerical evaluations of the integrals, the accuracy depends on the number of integration points used. This grid can lead to a superposition error similar to the BSSE and should therefore not be changed when calculating different properties of a quantum mechanical system.

### 1.1.5. Exchange-Correlation Functionals

It can be proven that there exists a unique exchange-correlation (XC) functional which is valid for all systems. If this functional was known explicitly, DFT would be exact. Sadly this is not the case, but a number of properties for the exact functional can be derived. All currently used functionals fulfill different properties, each having errors by neglecting others.

Usually the functional is split into an exchange and a correlation part. For the exchange part, in principle the HF exchange could be used. This has been shown to give poor results, because the definition of the exchange and correlation is not equivalent to the HF definition. In wave mechanics, both exchange and correlation energy have a short- and a long-range part. The long-range correlation effectively cancels the delocalized part of the exchange energy. For DFT those energies only depend on the local electron density and its vicinity and the cancellation at long range should be implicitly built into the functionals. Using exact exchange thus destroys this cancellation.

XC-functionals exhibit parameters which can either be fitted to experimental results (empirical) or be chosen in such a way that a maximum number of properties of the exact functional are fulfilled. The

quality of the functional has to be determined by comparing it to either a high level wave-mechanics calculation or to experiments.

Differently to wave function methods (where more Slater determinants are used), the quality of calculations can not be improved systematically, but J.P. Perdew has suggested a "Jacob's ladder" [8] where for each step up the ladder an improvement in accuracy can be expected. In the following some important steps on this ladder will be explained briefly.

## Local Density Approximation

The first approximation assumes, that the density can locally be treated as an uniform electron gas for which the exchange energy is known analytically and the correlation energy is also known to high accuracy. Here only the electron density is needed to calculate exchange and correlation energies. If the densities for different spins are not equal, this approach is generalized to two distinct spin densities which is then called *Local Spin Density Approximation*.

## Generalized Gradient Approximation

This class of functionals uses the first derivative of the electron density as additional parameter for the calculation of exchange and correlation energy. Directly adding the gradient to the energy would perform worse than LSDA, because some important properties would be destroyed. Therefore in *Generalized Gradient Approximation* (GGA) methods the gradient of the electron density is included as a variable which is tuned to fulfill certain requirements. One popular semi-empirical correlation functional is the one from Lee, Yang and Parr (LYP)[9] which has 4 parameters that are fitted to experimental data of the Helium atom. It is often used with the (also semi-empirical) B88 exchange functional from Becke [10]. A famous functional, that works without empirical data, was designed by Perdew, Burke and Enzerhof (PBE) [11]. Here the exchange and correlation functionals are designed as enhancement factors which are multiplied (exchange) and added (correlation) to the LSDA solution. The parameters for these functionals are optimized to fulfill the exact functional requirements, rather than fitted to experimental data.

## Meta-GGA

The next step on Jacob's ladder is allowing functionals to depend on higher order derivatives of the density. Inclusion of orbital kinetic energy dependence also falls into this category, as it gives essentially the same information as the second derivative of the electron density. One relatively new but very promising non-empirical representative of this class of functionals is called SCAN [12]. It is ought to replace the old-established functional from Tao, Perdew, Staroverov and Scuseria (TPSS)[13]. They can both be seen as a next improvement to PBE.

## Hybrid Functionals

For a system of non-interacting particles residing on the KS orbitals, Hartree-Fock would give the exact exchange energy, which is commonly referred to as exact exchange. Mixing this exact exchange with LSDA and GGA methods leads to a zoo of possible exchange-correlation functionals with even more parametrizations. One famous semi-empirical example is the combination of B88 with LYP called B3LYP [14]. On the non-empirical side there is also a hybrid version of PBE called PBE0 [15]. The improvement when mixing HF-exchange can be explained by the fact that most errors of LSDA (contained in most functionals) and HF display systematic errors in opposite directions and using a suitable linear combination of both cancels out most of those errors.

## Double Hybrid Functionals

Mixing in the full information of KS-orbitals leads to double hybrid methods, which are generally better than hybrid functionals but hindered by their computational cost.

## Range-Separated Methods

Current functionals have been found to describe the charge transfer excitation energies wrong. This is related to the overestimation of the delocalization of electrons and the self-interaction error stemming from the incorrect long-range behavior of exchange functionals which then fail to cancel the Coulomb term. To cope with this, range-separated functionals have been derived which use different exchange functionals for short and large distances to the nucleus. The parameter that usually controls the partitioning is commonly called  $\omega$ . For short range most commonly a density-exchange functional is used and for the long range part the exact HF-exchange is employed. Such range-separated versions [16] can be employed for all commonly used functionals but reparametrizations have to be done. Improvements to the previously mentioned functionals are LC-BLYP and LC-PBE. Livshits and Baer brought up the idea called *optimal tuning* [17] to choose  $\omega$  in such a way that Koopmans' theorem is fulfilled and the HOMO is equal to the ionization potential.

For use in periodic solid state systems sometimes HF is used for the short and density-exchange for the long-range part. Such functionals are then called screened-exchange. A famous example is from Heyd, Scuseria and Enzerhof (HSE) [18].

## Dispersion-Corrected Methods

A major disadvantage of standard DFT is, that most functionals are unable to describe long range dispersion forces (Van der Waals interactions). Therefore S. Grimme proposed additional empirical terms to account for vdW interactions [19]. First versions simply used the  $R^{-6}$  behavior together with a parameter  $C_6$  for every different species combination. To circumvent divergence at short distances, a damping function has been added. Further improvements were made by including higher order terms.

$$\Delta E_{disp} = \sum_{n=6(8,10)} s_n \sum_{AB}^{atoms} \frac{C_n^{AB}}{R_{AB}^n} f_{damp}(R_{AB}) \quad (1.38)$$

the damping function is typically of the form  $f_{damp}(R_{AB}) = \frac{s_n}{1+e^{-d(R_{AB}/s_n R_{AB0})^{-1}}}$  and in the simple approach by Grimme the  $C_6^{AB}$  coefficients are determined from single-species parameters via

$$C_6^{ij} = 2 \frac{C_6^i C_6^j}{C_6^i + C_6^j}$$

The scaling factor  $s_n$  is different for each XC functional to take account for the different amounts of short-range interaction included. More sophisticated methods calculate the dispersion coefficients  $C_n^{AB}$ , cutoff-radii  $R_{AB,0}$  and polarizability  $\alpha_i$  from their (often tabulated) free atomic values and the charge-density of the system taking into account the local chemical environment. For the method proposed by Tkatchenko and Scheffler [20] the effective volume  $v_i$  and the Hirshfeld weight  $w_i(\mathbf{r})$  are determined from the free atomic values via

$$v_i = \frac{\int r^3 w_i(\mathbf{r}) n(\mathbf{r}) d(\mathbf{r})}{\int r^3 n_i^{free}(\mathbf{r}) d(\mathbf{r})} \quad (1.39)$$

$$w_i(\mathbf{r}) = \frac{n_i^{free}(\mathbf{r})}{\sum_{j=1}^{N_{atom}} n_j^{free}(\mathbf{r})} \quad (1.40)$$

The dispersion interaction is then calculated with

$$C_{6ii} = v_i^2 C_{6ii}^{free} \quad (1.41)$$

$$\alpha_i = v_i \alpha_i^{free} \quad (1.42)$$

$$C_{6ij} = \frac{2C_{6ii}C_{6jj}}{\left[\frac{\alpha_j}{\alpha_i}C_{6ii} + \frac{\alpha_i}{\alpha_j}C_{6jj}\right]} \quad (1.43)$$

For this method there also exists a surface parametrization called  $TS^{surf}$  [21].

Another approach is to directly link the electron density to an appropriate dispersion kernel, leading to a six-dimensional integral with much higher cost. Those methods are called many-body dispersion correction [22]. Additionally the methods can be classified as self-consistent when employed in every SCF-cycle and a-posteriori if only added at the last step.

### 1.1.6. Errors Inherent to Common Density Functional Approximations

Here some of the most prominent errors that occur in density functional theory are summarized, a more detailed explanation is given in [23] and [24].

#### Self-Interaction Error

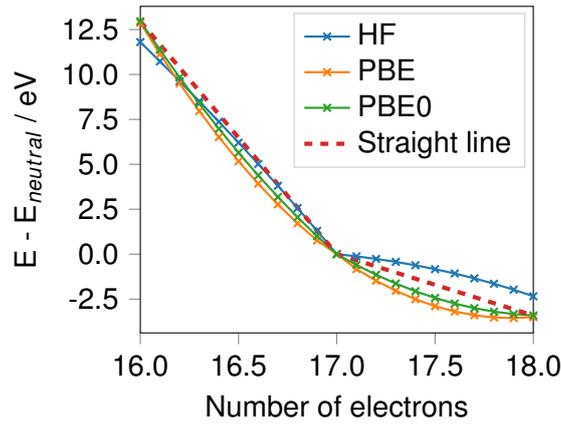
The one-electron self-interaction error stems from the fact, that for a quantum mechanical system the Coulomb self-energy should cancel out the exchange self-energy exactly (see eq. 1.15). Hartree-Fock takes this into account explicitly. For many DFT exchange functionals however this constraint does not hold exactly, which leads to an effective interaction energy of single electrons with themselves. As explained in section 1.1.5 this is a design criterion for functionals, but to date by far not all functionals in use take it into account.

#### Delocalization Error

This error is also called "deviation from straight line"-error or many-electron self-interaction error. When fractional charges are added or subtracted from an atom, the actual energy is a linear interpolation between the integer charge energies due to the discrete nature of electrons. For semi-local DFT this change is convex, for HF concave. This means that in DFT fractional charges are over-stabilized, leading to over-delocalization, underestimated bond lengths, underestimated band-gaps, too low transition state energies and unphysical charge transfer between molecules and metal surfaces. HF in comparison over-localizes electrons leading to the opposite behavior. Figure 1.1 shows this behavior for a single atom calculated with DFT and different XC functionals.

#### Static Correlation Error

For systems with strong electron-electron correlation, there can arise electronic states with fractional spins. For the exact XC-functional these states give the same energy as their comprising degenerate pure-spin states [25]. If the molecule  $H_2$  for example is dissociated, the spin states are degenerate which leads to fractional spins on both atoms. This energy should be equal to the one of pure spins located on both atoms, but currently used functionals deviate strongly from this behavior, which leads to an overestimation of covalent bonds. For the calculation of correlated systems however, such functionals will be needed.



**Figure 1.1.:** Deviation from straight line condition shown for DFT calculations of a single Cl atom with fractional charges and different XC functionals.

### 1.1.7. Periodic Quantum Mechanical Calculations

Calculating an extended system, such as a metal or semiconductor, would be extremely costly, as the computational expense is proportional to the number of electrons in the system. A periodic system can be described by an infinite repetition of a unit cell. The unit cell in three dimensions is described by three vectors ( $\mathbf{a}_1$ ,  $\mathbf{a}_2$  and  $\mathbf{a}_3$ ), defining the space and the shape of the cell. Any point  $\mathbf{r}'$  within the crystal can now be reached by defining a vector in the first unit cell  $\mathbf{r}$  and a linear combination of translation vectors  $\mathbf{a}_i$ .

$$\mathbf{r}' = \mathbf{r} + n_1 * \mathbf{a}_1 + n_2 * \mathbf{a}_2 + n_3 * \mathbf{a}_3 = \mathbf{r} + \mathbf{R} \quad (1.44)$$

The reciprocal cell is also defined by three orthonormal vectors obeying the condition  $\mathbf{a}_i * \mathbf{b}_j = 2\pi \delta_{ij}$ :

$$\mathbf{b}_1 = 2\pi \frac{\mathbf{a}_2 \times \mathbf{a}_3}{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)}, \quad \mathbf{b}_2 = 2\pi \frac{\mathbf{a}_3 \times \mathbf{a}_1}{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)}, \quad \mathbf{b}_3 = 2\pi \frac{\mathbf{a}_1 \times \mathbf{a}_2}{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)}$$

Those vectors span the unit cell in reciprocal space, also called Brillouin zone (BZ). Just as points in the real unit cell can be reached via  $\mathbf{r}'$ , points in reciprocal space can be reached via  $\mathbf{g}'$ , also called wave vector.

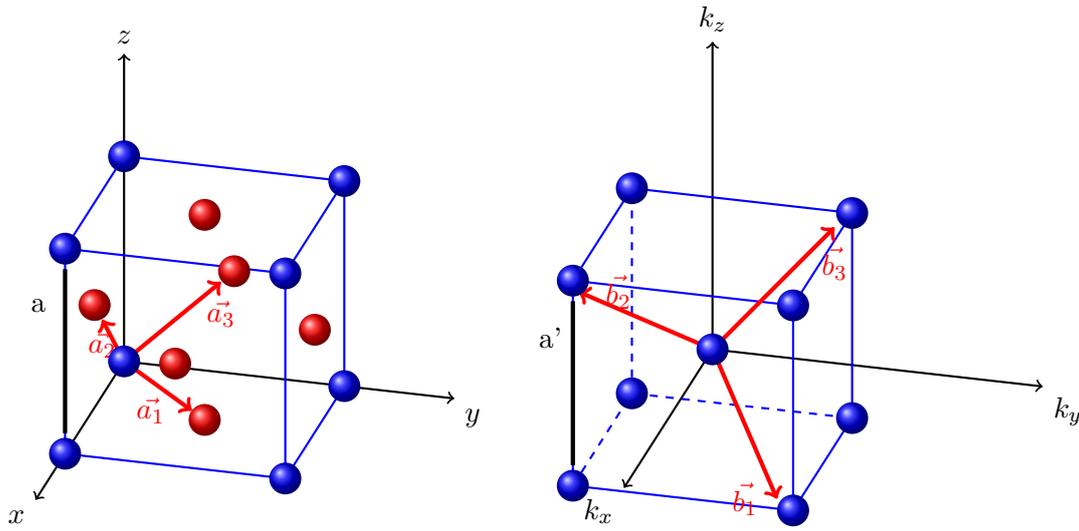
$$\mathbf{g}' = \mathbf{g} + k_1 * \mathbf{b}_1 + k_2 * \mathbf{b}_2 + k_3 * \mathbf{b}_3 = \mathbf{g} + \mathbf{k} \quad (1.45)$$

The periodicity of the system requires, that the square of the wave function at periodic replicas of the unit cell is the same, which means that the wave function can only differ by a complex wave factor. This requirement is fulfilled by Bloch waves [26]:

$$\phi(\mathbf{r} + \mathbf{R}) = \phi(\mathbf{r})e^{i\mathbf{k}\mathbf{R}} \quad (1.46)$$

with the unit cell vector  $\mathbf{R}$  introduced in equation 1.45 and the wave vector  $\mathbf{k}$ . Bloch orbitals can be created by combining spatially periodic functions with plane waves:

$$\phi_{n,\mathbf{k}}(\mathbf{r}) = \psi_n(\mathbf{r})e^{i\mathbf{k}\mathbf{r}}$$



**Figure 1.2.:** Left: fcc lattice in real space (balls on faces red), arrows are primitive lattice vectors. Right: Bravais lattice (blue balls) and reciprocal lattice vectors (red arrows) of the fcc lattice.

Those spatially periodic functions can either be expanded into plane waves (PW) or suitable atom-centered functions (ACO) in Bloch-form:

$$\phi_{n,\mathbf{k}}^{PW}(\mathbf{r}) = e^{i\mathbf{k}\mathbf{r}} \sum_i^{N_{basis}} c_{ni} \chi_i^{PW}(\mathbf{r})$$

$$\phi_{n,\mathbf{k}}^{ACO}(\mathbf{r}) = \sum_i^{N_{basis}} \sum_{\mathbf{R}} c_{ni} e^{i\mathbf{k}\mathbf{r}} \chi_i^{ACO}(\mathbf{r} + \mathbf{R})$$

for further discussion of basis sets see section 1.1.3.

This projects the problem from infinitely many electrons to those in the unit cell, but now the solutions are functions of the reciprocal space vector  $\mathbf{k}$ . The problem can now again be mapped onto a matrix equation like equation 1.24:

$$\mathbf{F}^{\mathbf{k}} \mathbf{C}^{\mathbf{k}} = \mathbf{S}^{\mathbf{k}} \mathbf{C}^{\mathbf{k}} \boldsymbol{\epsilon}^{\mathbf{k}}$$

The solutions to this equation are continuous functions of  $\mathbf{k}$  called bands. For every  $\mathbf{k}$ -vector many solutions are found, forming many energetically distinct bands.

To determine the number of states at specific energies, the density of states (DOS) is used. The electron density is then found by filling the bands up to the Fermi level

$$\rho(\mathbf{r}) = \int_{BZ} d\mathbf{k} \sum_{n=0}^{N_{bands}} |\psi_{i\mathbf{k}}(\mathbf{r})|^2 \phi(E_n(\mathbf{k}) - E_F)$$

$E_i(\mathbf{k})$  is the eigenvalue of the  $i$ -th crystalline orbital at the reciprocal point  $\mathbf{k}$ . The Heaviside function  $\theta$  stems from the Fermi-Dirac distribution

$$f(E) = \frac{1}{\exp\left(\frac{E-E_F}{k_B T}\right) + 1}$$

evaluated at 0 Kelvin. The Fermi level  $E_F$  determines the occupation of each orbital. It can be defined via the occupation number of the system:

$$n = \int dE DOS(E) f(E)$$

$$DOS(E) = \int_{BZ} \frac{d^3k}{(2\pi)^3} \delta(E - E(\mathbf{k}))$$

With this the Fermi level is the energy where a state (if there is one) is half filled, independent of temperature. If this energy is in the band gap of a material,  $E_F$  at 0 K (which is the case for QM calculations) can be at any arbitrary energy between the highest occupied and lowest unoccupied state and therefore can not be used as a physically meaningful parameter.

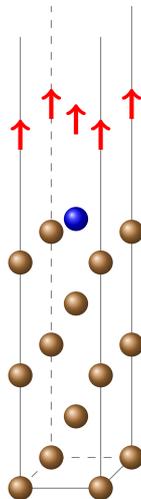
Due to the relatively smooth form of the bands, often only a small amount of k-points needs to be calculated for a good result, but the choice of k-points is critical for the calculation performance.

One of the most used methods to distribute points in k-space is the Monkhorst-Pack scheme [27], which spaces the points equally in the first Brillouin-zone but with an offset from the Gamma point (origin in k-space).

### Repeated Slab Approach

For the modeling of bulk-like properties considering a large number of atoms is essential, which can only be realized with periodic calculations. For the calculation of surfaces however, 3D-periodic calculations are not suitable, as there is no periodicity perpendicular to the surface. While there exist methods to calculate 2D systems [28], the modeling of surfaces is still a niche in the field of ab-initio modeling, so there is no reliable code for such systems. To calculate 2D systems with 3D code, the unit cell is constructed to have a large vacuum area above the surface slab. As there is no coupling between the periodic surface-replicas in z-direction, only one k-point needs to be used in this direction.

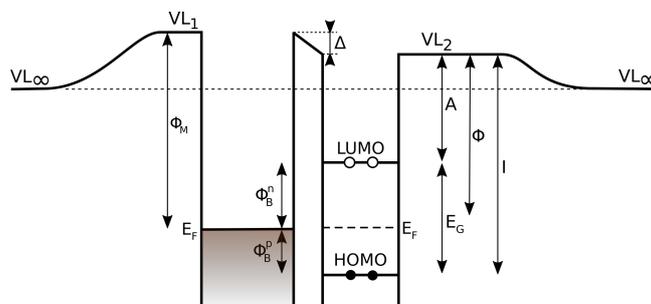
Systems that are not symmetric in z-direction can have different electrostatic properties at the top and at the bottom. This would lead to an electrostatic potential gradient in the slab and thus to physically unreasonable charge rearrangements. To overcome this problem, an artificial dipole layer is introduced in the slab, creating a step in the electrostatic potential. An example of a surface slab with 1 surface atoms in the unit cell and a single atom adsorbed on top is shown in figure 1.3. The red arrows indicate the dipole layer.



**Figure 1.3.:** Illustration of the repeated slab approach. The red arrows indicate the artificial dipole layer introduced.

## 1.2. Interactions of Organic Molecules and Surfaces

Here a brief introduction to some specific interactions between molecules and surfaces is given. Figure 1.4 shows the energy level diagram of a molecule layer adsorbed on a metal surface. The interaction of the molecule with the surface introduces a surface dipole, which shifts the surface vacuum level  $VL_1$  of the metal surface by  $\Delta$ , whereas the vacuum level at infinite distance  $VL_\infty$  stays constant. The surface work function is changed from the metal work function  $\psi_M$  to  $\psi$ .  $\psi_B^n$  and  $\psi_B^p$  are the electron and hole injection barriers. The energy of the molecules' fundamental gap  $E_G$  plus the electron affinity  $A$  gives the ionization energy  $I$  of the molecule on the surface.



**Figure 1.4.:** Explanation of the different energies used when explaining the interactions of molecules on surfaces. Possible origins of the interface dipole are explained below.

Here some of the interactions creating the surface dipole are explained briefly, for a detailed explanation see the reviews from Ishii et al. [29] and Braun et al. [30].

**Electrostatic Interaction** Molecules with intrinsic dipoles can generate a dipole layer at the interface when aligned properly, generating shift in the vacuum level. Charged molecules can create mirror charges in the substrate, leading to attractive forces.

**Chemical Bonds** The formation of chemical bonds also leads to an effective surface dipole which influences the work function of the surface.

**Push back** For metals, the core charges are usually at fixed positions, but the electrons can move freely in the solid and also 'leak' out of the surface. When a molecule physically adsorbs on such a surface (via van der Waals forces), the molecules' electron density pushes the metal electrons away from the molecule and partly back into the slab. This generates a surface dipole that shifts the vacuum level. The whole process is then called push back.

**Charge Transfer** The substrate and molecule layer can have different electronic chemical potentials at the interface. This will lead to an exchange of charges until an energetic equilibrium is reached. If the number of mobile charge carriers is high enough, this will lead to a redistribution of electrons until the chemical potential is aligned. Some specific characteristics of charge transfer are

**Fermi Level Pinning:** As soon as the Fermi level of the metal is below the HOMO or above the LUMO of the molecule, electrons can be transferred from the HOMO or to the LUMO which leads to a pinning of the HOMO or LUMO at the Fermi level, independent of the metals' work function.

**Backdonation:** When charge is transferred to the LUMO of the molecule, some of the charge from energetically lower lying orbitals, which are closer to the metal surface, can be partially depleted, leading to a backdonation of electrons into the metal.

**Integer Charge Transfer** The charge transferred from an inorganic substrate to a full organic layer (or back) is typically lower than one electron per molecule. It is not completely clear, if those charges are distributed homogeneously over the whole layer or whether some molecules have integer charged whereas others remain neutral. There is evidence for such integer charge transfer situations for systems with a nonconducting buffer between a metal substrate and the molecule layer, but the topic is still highly debated.

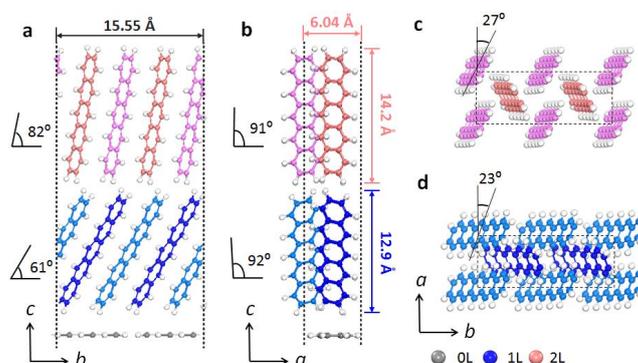
### 1.3. Structure Search

In thermodynamic equilibrium, the crystal structure that is lowest in Gibbs free energy will form. In many cases, many competing low enthalpy structures exist and the experimentally observed polymorph is governed by external parameters (like temperature or pressure). If only one polymorph is present from 0 K until melting point, it is called monotropic. For enantiotropic systems on the other hand the relative stability of polymorphs changes with the temperature. If the condition of thermodynamic equilibrium is not fulfilled, also kinetically trapped polymorphs can be observed. Thus for many systems a range of possible polymorphs can be found. Already in 1965 Walter McCrone stated, that the amount of polymorphs found corresponds to the time and money spent searching for it [31].

The search for crystal structures of organic molecules has been a task over decades [32]. First it was primarily driven by drug research as different polymorphs can change the biological activity and legal regulations require the knowledge of all possible polymorphs present in the drug. Since the uprising of organic electronics, the influence of the relative molecule arrangements has also been recognized as a dominant factor in the performance of organic devices.

A very prominent example therefore is the mobility of rubrene, which does not exceed  $2 \text{ cm}^2/Vs$  in bulk phases [33] but can increase up to  $40 \text{ cm}^2/Vs$  when grown in the right polymorph in a thin film transistor [34]. Here it needs to be highlighted, that this high mobility can only be achieved by one of four possible polymorphs and that this is not a surface induced phase.

In recent years, the importance of the interface between inorganic and organic materials for charge injection and other device properties was also realized. One crucial insight was the role of surface-induced phases for the device performance, which offers new design strategies for organic electronics [35]. Surface polymorphism with properties varying from the bulk systems have been found for various molecules. The role of such surface polymorphs for crystal growth on inorganic and metal substrates has been reviewed by Andrew Jones and colleagues [36]. They state that the formation of such polymorphs can directly be governed by the surface or that a wetting layer can form and act as a mask for the following layers. The phases of subsequent molecular layers usually perform a transition into bulk phases with increasing thickness. A sketch of such a transition from a wetting layer over a surface phase to the bulk structure for pentacene is shown in figure 1.5.



**Figure 1.5.:** Schematic representation of the first three layers of pentacene on boron nitride. Taken from [37].

## 1.4. Computational Structure Search

Following the importance of organic polymorphism, much effort has been put into computational prediction of polymorphism for single molecules and molecular crystals. Application of structure search to organic-inorganic interfaces is still at the very beginning. One reason therefore is the configurational explosion. It states that the number of possible surface structures grows exponentially with the number of molecules and coverage considered. Thus the configurational space can become intractably large already for very small systems. Another prohibitive factor is the computational cost of *ab-initio* calculations of periodic surface-molecule interfaces, which is orders of magnitude larger than for bulk crystals or single molecules. This means that already established molecular dynamics methods for crystal structure prediction, such as simulated annealing [38, 39], basin hopping [40, 41], minima hopping [42, 43], and also evolutionary approaches such as genetic algorithms [44, 45] or particle swarm optimization [46] are not applicable to organic/inorganic interfaces due to the large data sets needed for successful prediction. For this reason, most structure search methods for such interfaces focus on the adsorption of single molecules on the surface [41, 47, 48]. Nevertheless there are very promising approaches for the *ab-initio* prediction of full molecular monolayers using kernel ridge regression [49].

### 1.4.1. Linear Regression Methods

The machine learning model used in this thesis is based on linear regression, therefore first the basic concepts of this method will be explained.

Suppose a linear model with a target variable  $y$  dependent on variables  $x$  and model parameters  $w$ .

$$y_i = w_1x_{i1} + \dots + w_px_{ip} \quad (1.47)$$

A given set of known data points  $\mathbf{y}'$  should be used to estimate the best parameters to predict further points  $\mathbf{y}$ . The errors of the data points are supposed to be normally distributed  $\epsilon_i \propto \mathcal{N}(0, \sigma_M)$ . For a single data point  $y'_i$  this can be written as

$$y'_i = w_1x_{i1} + \dots + w_px_{ip} + \epsilon = \mathbf{w}\mathbf{x}_i + \epsilon_i \quad (1.48)$$

and for a vector of  $N$  data points  $\mathbf{y}'$  in matrix form

$$\mathbf{y}' = X\mathbf{w} + \boldsymbol{\epsilon} \quad (1.49)$$

$$X = \begin{pmatrix} x_{11} & \dots & x_{1p} \\ \vdots & \vdots & \vdots \\ x_{i1} & \dots & x_{ip} \\ \vdots & \vdots & \vdots \\ x_{N1} & \dots & x_{Np} \end{pmatrix} \quad (1.50)$$

The conditional probability of the  $y'_i$  given  $x_i$  now reads

$$p(\mathbf{y} = \mathbf{y}'|X) = p(X\mathbf{w} + \boldsymbol{\epsilon} = \mathbf{y}'|X) = p(\boldsymbol{\epsilon} = \mathbf{y}' - X\mathbf{w}|X)$$

$$p(\boldsymbol{\epsilon} = \mathbf{y}' - X\mathbf{w}|X) = \mathcal{N}(\mathbf{y}' - X\mathbf{w}, \sigma_M) \propto e^{-\frac{(\mathbf{y}' - X\mathbf{w})^2}{2\sigma_M^2}}$$

Maximizing this probability with respect to  $w$  is widely known as maximum likelihood estimation. It is equal to the problem

$$\min_w \left( \|X\mathbf{w} - \mathbf{y}'\|_2^2 \right) \quad (1.51)$$

which is also known as least squares approach.

The solution is given by  $(X^T X)^{-1} X^T \mathbf{y}$ , but if the number of data points is smaller than the number of parameters, this is an ill-conditioned problem and some more information needs to be added to the model.

This can be done using Bayes' theorem with the already known likelihood  $p(\mathbf{y}'|\mathbf{w}, X)$

$$p(\mathbf{w}|\mathbf{y}', X) = \frac{p(\mathbf{y}'|\mathbf{w}, X)p(\mathbf{w}|X)}{p(\mathbf{y}'|X)} \quad (1.52)$$

To maximize  $p(\mathbf{w}|\mathbf{y}, X)$  some prior information about the model can be added. Using a normal distributed Gaussian prior

$$p(\mathbf{w}|X) \propto \mathcal{N}(\hat{\mathbf{w}}, \Gamma) \quad (1.53)$$

leads to a minimization problem which is known as maximum a posteriori optimization. It is equivalent to least squares with Tikhonov regression [50] using a Tikhonov matrix  $\Gamma$

$$\min_{\mathbf{w}} \left( \|X\mathbf{w} - \mathbf{y}'\|_2^2 + \|\Gamma\mathbf{w}\|_2^2 \right) \quad (1.54)$$

Choosing the matrix as a multiple of the identity matrix is also called L2 regularization or ridge regression

$$\min_{\mathbf{w}} \left( \|X\mathbf{w} - \mathbf{y}'\|_2^2 + \alpha\|\mathbf{w}\|_2^2 \right) \quad (1.55)$$

and leads to a preference of very many non-zero but small parameters.

If the parameter space should ideally be very sparse, a Laplacian prior can be used

$$\min_{\mathbf{w}} \left( \frac{1}{2n_{samples}} \|X\mathbf{w} - \mathbf{y}'\|_2^2 + \alpha\|\mathbf{w}\|_1 \right) \quad (1.56)$$

also called LASSO regularization. This prior favors sparse parameter vectors where the nonzero elements are still small.

An analytic solution for those minimization problems exists only, if conjugate priors are used, which means that the likelihood and prior are from the same functional family.

### 1.4.2. Bayesian Linear Regression

One such method with a conjugate prior is Bayesian linear regression. Here the prior is a multivariate Gaussian distribution. The posterior probability for the parameters  $\mathbf{w}$  given some data  $\mathbf{y}'$  is

$$p(\mathbf{w}|\mathbf{y}') = \frac{1}{Z} p(\mathbf{y}'|\mathbf{w}) p(\mathbf{w}) \quad (1.57)$$

with the likelihood  $p(\mathbf{y}|\mathbf{w})$  and the prior  $p(\mathbf{w})$

$$p(\mathbf{y}|\mathbf{w}) \propto e^{-\frac{1}{2\sigma^2} \|X\mathbf{w} - \mathbf{y}'\|^2} \quad (1.58)$$

$$p(\mathbf{w}) \propto e^{-\frac{1}{2}(\mathbf{w} - \bar{\mathbf{w}})^T C^{-1}(\mathbf{w} - \bar{\mathbf{w}})} \quad (1.59)$$

combining those two leads to the full posterior probability

$$p(\mathbf{w}|\mathbf{y}) \propto e^{-\frac{1}{2\sigma^2} (X\mathbf{w} - \mathbf{y}')^T (X\mathbf{w} - \mathbf{y}')} e^{-\frac{1}{2}(\mathbf{w} - \bar{\mathbf{w}})^T C^{-1}(\mathbf{w} - \bar{\mathbf{w}})} \quad (1.60)$$

To maximize the probability to get the correct parameters, only the exponent needs to be minimized

$$\min_{\mathbf{w}} \left( \|X\mathbf{w} - \mathbf{y}'\|^2 + (\mathbf{w} - \bar{\mathbf{w}})^T C^{-1} (\mathbf{w} - \bar{\mathbf{w}}) \right) \quad (1.61)$$

The right part of the equation can again be seen as regularization.

This probability can now be rewritten to a standard Gaussian. For better readability this is done with the logarithmic probability.

$$\log p(\mathbf{w}|\mathbf{y}') \propto -\frac{1}{2} \left[ \frac{1}{\sigma^2} (X\mathbf{w} - \mathbf{y}')^T (X\mathbf{w} - \mathbf{y}') + (\mathbf{w} - \bar{\mathbf{w}})^T C^{-1} (\mathbf{w} - \bar{\mathbf{w}}) \right] \quad (1.62)$$

$$\propto -\frac{1}{2} \left[ \mathbf{w}^T \left( \frac{X^T X}{\sigma^2} + C^{-1} \right) \mathbf{w} - 2 \left( \bar{\mathbf{w}}^T C^{-1} + \frac{\mathbf{y}'^T X}{\sigma^2} \right) \mathbf{w} + \frac{1}{\sigma^2} \mathbf{y}'^T \mathbf{y}' + \bar{\mathbf{w}}^T C^{-1} \bar{\mathbf{w}} \right] \quad (1.63)$$

Equating the coefficients with

$$\log p(\mathbf{w}|\mathbf{y}') \propto -\frac{1}{2} \left[ (\mathbf{w} - \mathbf{w}')^T A^{-1} (\mathbf{w} - \mathbf{w}') \right] + R \quad (1.64)$$

leads to the new covariance  $A$  and mean  $\mathbf{w}'$

$$A^{-1} = \frac{X^T X}{\sigma^2} + C^{-1} \quad (1.65)$$

$$\mathbf{w}' = A \left( C^{-1} \bar{\mathbf{w}} + \frac{X^T \mathbf{y}'}{\sigma^2} \right) \quad (1.66)$$

which leads to the posterior probability

$$p(\mathbf{w}|\mathbf{y}') = e^{-\frac{1}{2}(\mathbf{w}-\mathbf{w}')^T A^{-1}(\mathbf{w}-\mathbf{w}')} / \mathcal{Z} \quad (1.67)$$

This is now a multivariate probability distribution. The mean of the distribution are the best parameters  $\mathbf{w}'$  and their variance and also covariance is encoded in  $A$ . Those parameters can then be used to predict new values  $y$ , given the coordinates  $x$ .

## 1.5. The SAMPLE Approach

The structure search method used in this thesis is called **Surface Adsorbate polyMorph Prediction with Little Effort** or short **SAMPLE**. It is an approach to efficiently predict all possible surface polymorphs that molecules can form on any given substrate. At the moment the approach is limited to monolayers of molecules with significant molecule-substrate interactions as it is limited to commensurate structures.

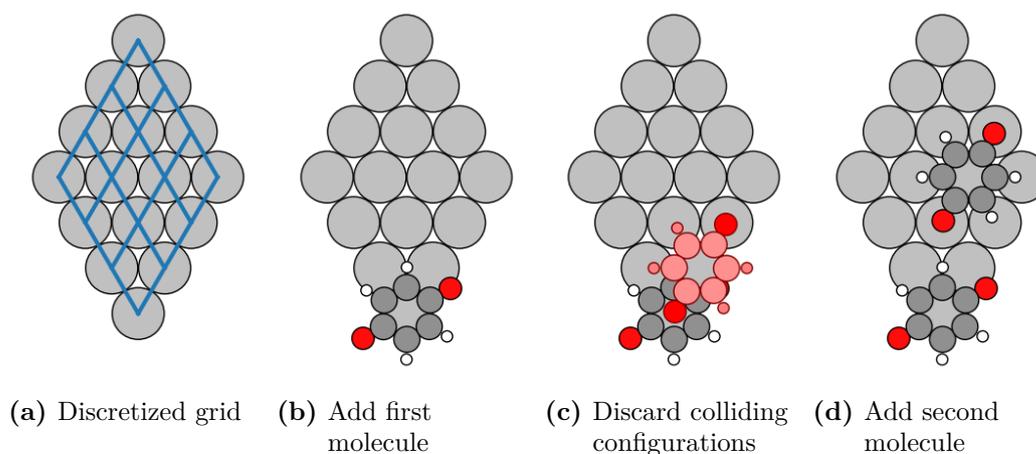
The polymorph prediction is done in four distinct steps from a discretization of the potential energy surface to the prediction of all possible polymorphs on a surface. In this section a short overview of these four steps will be given, for an extensive explanation see [51]. The approach has already been shown to be able to explain the formation of triangular TCNE structures on Au(111) [52] and could identify the defect energies of kink formation for TCNE on Ag(100) [53].

### 1.5.1. Step 1: Generation of Local Geometries

In this step the interaction of single molecules with the surface is explored. The basically infinite potential energy surface for the position of a single molecule on the surface is reduced to the local energy minima. Those geometric local minima are then used as discrete building blocks with a common anchor point, which is the surface atom closest to the geometrical center of the molecule. This approach combined with basin hopping was successfully applied for the structure prediction of TCNE on Au(111)[52]. A detailed explanation of the determination of the local geometries can be found in section 2.3.

### 1.5.2. Step 2: Combination to Configurations

Once the building blocks are found, they need to be combined in every possible way for a full exploration of the discretized configurational space. In principle this is easy: First one needs to take all possible inequivalent surface unit cells. Then, for each unit cell, all combinations of symmetry inequivalent local geometries that fit into the cell have to be built. This is done by representing the unit cell as a discretized grid with every surface atom in the unit cell acting as an anchor point. The first local geometry is then put onto the first anchor point. The next molecule is positioned on another anchor point and then the structure is checked for collisions. If the molecules collide, in the cell or in any periodic replica, the second molecule is moved to the next position, otherwise another molecule gets added with the same procedure. This procedure is repeated for all possible local geometries on all possible positions with specified cell and number of molecules. A representative sketch is shown in figure 1.6. The combination of a specific unit cell with local adsorption geometries will further be called configuration. Periodically replicated configurations will be called structures.



**Figure 1.6.:** Creation of a configuration with 2 molecules in a 4x4 unit cell. First a discretized grid is specified, then the first molecule is added at the first position and the second molecule is shifted until the molecules do not collide with each other (adapted from [54]).

The main problem of this method is, that the number of configurations increases exponentially with the number of molecules and the size of the unit cell. Lukas Hörmann was able to tackle this problem in his master's thesis [54]. He developed methods to remove all symmetry equivalent configurations and parallelized the code to a degree where it is possible to create millions of configurations within a day on a standard personal computer.

### 1.5.3. Step 3: Learning of Interactions

Now that the millions of possible inequivalent discretized configurations are found, a formalism to estimate their energies is needed. The main problem here is, that in contrast to many machine learning applications, each data point is very expensive to evaluate ( $\approx 1000$  CPUh). Thus an energy model, which enables the description of the system, and a method to efficiently train the model are needed. A very general model would be an expansion of all interaction energies in a configuration into n-body contributions

$$E_{conf} = \sum_i^{N_{1body}} n_i E_i + \sum_j^{N_{2body}} n_j E_j + \sum_k^{N_{3body}} n_k E_k + \dots \quad (1.68)$$

To train this model, all n-body interaction energies  $E_\alpha$  for all n-body contributions need to be fitted sufficiently. This model is still too complicated and would lead to a highly underdetermined system given the affordable number of calculations. A reasonable simplification of the model is a truncation at the 2-body contributions as 3-body and higher interactions are not expected to have a significant influence in the formation of surface polymorphs.

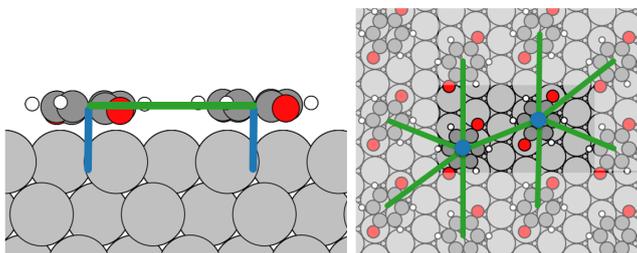
The energy model used then reads

$$E_{conf} = \sum_{geometries} n_g U_g + \sum_{pairs} n_p V_p \quad (1.69)$$

$$= N * \mathbf{E} \quad (1.70)$$

$$\mathbf{E} = (U_1, \dots, U_{N_g}, V_1, \dots, V_{N_p})^T \quad (1.71)$$

and is visualized in figure 1.7.



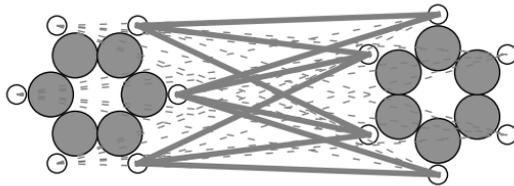
**Figure 1.7.:** Visualization of the interactions used for the energy model. Molecule-slab interactions are shown in blue and molecule-molecule interactions in green (adapted from [51]).

Here the 1-body contributions represent the interaction of a single molecule with the surface and the 2-body contributions represent interactions of pairs of molecules with each other.

The number of 1-body contributions is limited to the number of local adsorption geometries. For the 2-body interactions the situation is different. Here, in principle, there are still infinitely many contributions, but many of them are very similar. To encode this similarity, pairs of molecules are represented with feature vectors (FV). A common feature vector used for the representation of molecules is the Coulomb matrix, where the molecule is represented as a matrix of all interatomic distances  $r_{ij}$  scaled with the atomic numbers. As this FV operates in the internal coordinate system of the molecule, it is insensitive with respect to symmetry operations. A reasonable adaption of this FV for molecule pairs is the replacement of interatomic distances with distances of atoms on different molecules. To reduce the amount of redundant information, only a subset of the atoms in the molecule (usually those at the rim of the molecule) are considered. The distances are then scaled with a minimal distance  $d_{min}$ , specifying the closest possible arrangement of two molecules. The minimal distance needs to be chosen such that attractive interactions are fully represented, but Pauli repulsion is excluded. How it is chosen in this thesis is explained in section 2.4.2.

One major drawback of the Coulomb matrix feature vector is, that it is discontinuous with respect to a permutation of atom indices. Thus, for the feature vector in SAMPLE the distances are sorted, and the feature vector is populated with the  $N$  smallest distances. In figure 1.8 all distances between hydrogen atoms surrounding two benzene rings are shown. The ten smallest distances, which could be used for a FV are highlighted by a thicker line width. To increase the separation in feature space for smaller distances, the renormalized distances are potentiated by  $n$ , where  $n$  is typically between 1 and 5. The full mathematical form of the feature vector is displayed in equation 1.75.

$$\mathbf{f}(g1, g2) = \begin{pmatrix} \left( \frac{d_{min}}{|r_i - r_j|} \right)^n \\ \left( \frac{d_{min}}{|r_i - r_j|} \right)^n \\ \vdots \end{pmatrix} \quad (1.72)$$



**Figure 1.8.:** Visualization of all distances between hydrogen atoms of two benzene molecules. The thick lines mark the 10 smallest distances used for the feature vector.

For an estimation of the interaction energies  $\mathbf{E}$ , Bayesian linear regression (equation 1.61) is used. The regression problem here takes on the form

$$\min_{\mathbf{E}} \left( \|N\mathbf{E} - \mathbf{E}_{DFT}\| + (\mathbf{E} - \hat{\mathbf{E}})^T C^{-1} (\mathbf{E} - \hat{\mathbf{E}}) \right) \quad (1.73)$$

The parameter vector now consists of the interaction energies  $\mathbf{E}$  with the corresponding mean values  $\hat{\mathbf{E}}$ . The model matrix  $N$  consists of as much rows with model vectors  $\mathbf{n}$  as there are training configurations  $\mathbf{E}_{DFT}$ . The number of columns is equivalent to the length of the interaction energy vector  $\mathbf{E}$ . Each element of the matrix represents the number of occurrences of this specific interaction in the configuration, normalized by the number of molecules in the configuration. As an example: a configuration with 2 of 4 possible adsorption geometries in the unit cell and a distinction between six different pairwise interactions, where one is present four times and the other twice in the configuration, the model vector  $\mathbf{n}$  will look as follows:

$$\begin{aligned} n_i &= \underbrace{(1, 0, 1, 0)}_{\text{1body terms}} \underbrace{(0, 2, 0, 4, 0, 0)}_{\text{2body terms}} / 2 \\ &= \left( \frac{1}{2}, 0, \frac{1}{2}, 0, 0, 1, 0, 2, 0, 0 \right) \end{aligned}$$

Those model vectors for all data points are then combined to the model matrix  $N$ .

Now only values for the mean  $\hat{\mathbf{E}}$  and covariance  $C$  for the prior  $p(\mathbf{E})$  are missing. As prior energies for the molecule-slab interactions  $U_g$  the adsorption energies from the local geometry search are used. All interaction energies are chosen to be zero and the physical knowledge is encoded in the covariance matrix. The single-body submatrix of  $C$  is given by a constant value  $\sigma_{ads}^2$  times a unity matrix. For the covariance of the pair energies a twofold kernel is used.

The first part  $\sigma_i$  consists of the mean pairwise prior deviation  $\sigma_{int}$  and a factor that scales this correlation with the minimal distance between two geometries renormalized by a typical decay length  $\tau$ . This kernel represents the physical intuition, that pair interactions will decay to zero for large distances.

$$\sigma_i = \sigma_{int} \cdot e^{-\frac{d_{min} - \min(|\mathbf{r}\mathbf{g}_i - \mathbf{r}\mathbf{g}_j|)}{\tau}} \quad (1.74)$$

For the covariance of a pair of feature vectors, the  $\sigma_i$  are multiplied with an additional kernel that encodes the similarity of two molecule pairs in feature space. This encodes the assumption, that similar pairs will have similar interaction energies. The elements of the pair-energy part of the covariance matrix now read

$$C_{ij} = \sigma_i * \sigma_j * e^{-\frac{|f_i - f_j|}{\gamma}} \quad (1.75)$$

When training the model, a small D-optimally selected subset of configurations is calculated with DFT and the covariance and mean of the parameter posterior  $p(\mathbf{E}|\mathbf{y})$  are updated according to equations 1.65 and 1.66.

D-optimal selection is an established method in experimental design theory. It chooses a set of configurations that maximizes the determinant of the information matrix  $N^T N$ .

**The hyper parameters used in SAMPLE** are

$\sigma_{DFT}$	DFT uncertainty
$\sigma_{ads}$	Estimated deviation of adsorption energies
$\sigma_{int}$	Estimated energy range of interaction energies
$\tau$	Characteristic decay length for interactions
$d_{min}$	Minimal distance between atoms
$\gamma$	Characteristic decay length of feature vector

They are preoptimized on artificial gasphase monolayers (full configuration with removed metal substrate). This enables a reasonable hyper parameter optimization with little computational effort. The exact procedure is explained in section 2.4.

#### 1.5.4. Step 4: Prediction

Once training of the model is finished, the energies for all remaining configurations can be predicted. Therefore only the model vectors  $\mathbf{n}$  need to be determined. Then the energies can be calculated by evaluation of a simple matrix multiplication

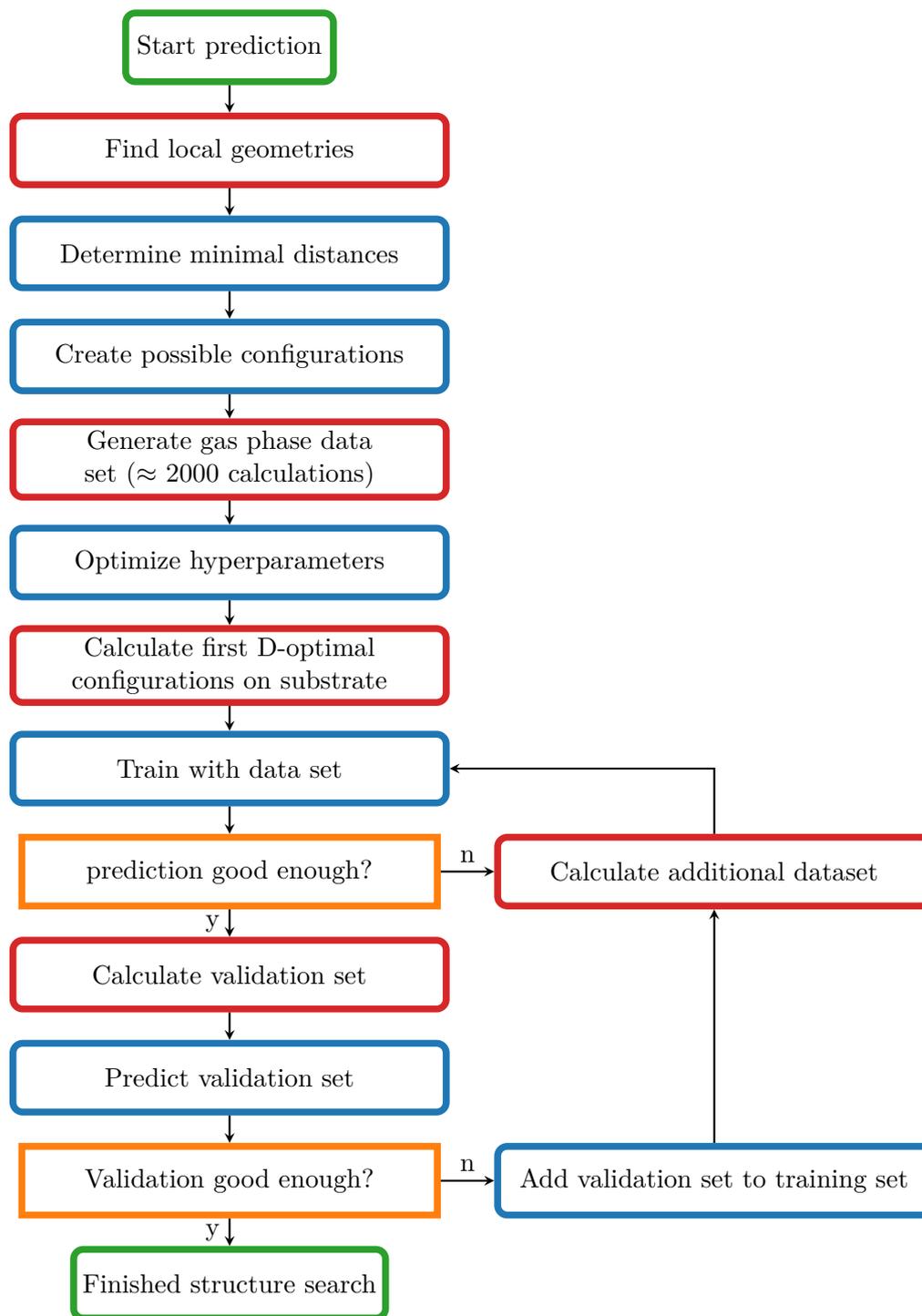
$$\mathbf{E}_{pred} = N * \mathbf{E} \quad (1.76)$$

Thus the configuration energy of millions of structures can be evaluated within a few CPUh, enabling the prediction of large data sets.

The quality of the prediction is evaluated by validation on a test set consisting of DFT calculated configurations that were not part of the training set. The root mean square error (RMSE, see equation 1.77) between prediction and DFT energy is representative for the uncertainty of all predictions.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (E_{pred,i} - E_{exact,i})^2} \quad (1.77)$$

A full workflow of the SAMPLE method is shown in figure 1.9.



**Figure 1.9.:** Schematic workflow of the sample method from the calculation of local geometries to the prediction of surface polymorphs. Red boxes involve costly DFT calculations. Blue boxes are done by the SAMPLE package and orange boxes are user decisions.

## 2. Making Tractable

In the first part of this chapter, the necessary adaptations to the computational method are explained. As an increasing molecule size inevitably increases the computational expense, the largest molecule used here would have needed resources beyond our capacity. Thus, the first part will introduce the computational methodology and an approach to reduce the cost of every single DFT calculation significantly, which enabled the calculations of the results presented in chapter 3.

While being able to calculate the properties of interest quantum chemically for a few configurations is nice, it is still not sufficient for a reliable structure prediction. Therefore the second part of this chapter will focus on adaptations to the the SAMPLE approach (sec. 1.5). Some of these were necessary to enable the treatment of molecules with functional groups, while others massively improved the predictive power.

### 2.1. Quantum Mechanical Method Used

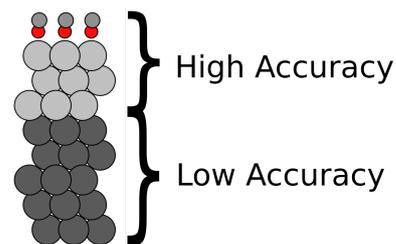
The computational effort of periodic surface calculations scales formally  $\mathcal{O}(N^3)$  with respect to the number of basis functions  $N$ . For periodic systems, the use of plane waves would in principle be beneficial (see section 1.1.7). In this thesis mainly surface calculations employing the repeated slab approach and large areas of vacuum were performed. The description of vacuum with plane waves in the periodic slab approach is very costly, thus for this thesis the DFT code FHI-aims [5] was mainly used. Here, by using numeric atom-centered basis functions and employing suitable numerical integration schemes, the scaling can be brought down to roughly  $\mathcal{O}(N)$ . As exchange correlation functional PBE was used. For the dispersion correction the method by Tkatchenko and Scheffler (TS) in its surface parametrization  $TS^{surf}$  was taken. The k-point distribution was carried out with the Monkhorst-Pack scheme, but no gamma offset was used. For the species basis sets, tight settings with modifications due to the convergence tests were used.

### 2.2. Mixed Basis Set Method

Despite the good scaling of numeric atomic orbitals, larger systems ( $>5000$  electrons) are still hard to handle. For metal-organic interfaces however, such large systems occur, as due to the strong delocalization of electrons in metals, surfaces have to comprise several atom-layers and large surface areas need to be considered for molecule-surface interactions.

One little recognized detail is, that high accuracy is only needed in the vicinity of the surface, where the interaction with the molecules takes place. The slab underneath only has to act as an electron bath with reasonable bulk properties. Therefore a reasonable approach for cost-reduction is to adapt the number of basis functions used with respect to the position in the slab.

For the slab used in this thesis this means: Use a well converged basis for the uppermost slab layers and as simple and cheap basis functions as possible for the lower lying layers. This procedure is here often called mixed basis set method.



**Figure 2.1.:** Accuracy requirements for the metal slab layers.

### 2.2.1. Application to Ag(111)

The structure search in this thesis is performed on the Ag(111) surface, thus it is the system of choice for a test of the basis set optimization. The first step when starting with a new system in DFT is performing convergence tests for all relevant parameters. Those tests for the standard parameters can be found in section A. Within these tests one problem arose: With the standard methods it was not possible to gain reasonable accuracy (below 20 meV) for the convergence of the number of substrate layers within the available computational time. This problem was the driving force in the development of the mixed basis set method.

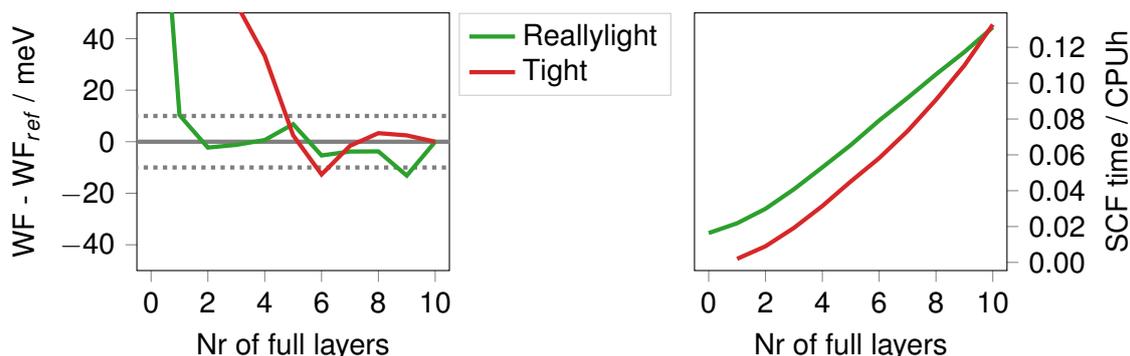
#### Setup

The assumption of the approach was, that the lower lying layers only act as an electron bath and thus only the minimal basis is needed. As there was no reference available, 3 different test species with decreasing accuracy were used. This was achieved by starting with the standard *tight* Ag basis set and subsequent removal of basis functions and sparsening of the integration grid up very light settings which are further on called *reallylight*. The explicit settings and full calculations are shown in section C.

#### Converging the Work Function

The resulting slab should still represent the correct physical properties of a silver surface. A reasonable slab thickness is assumed to be reached when the surface-work function is reasonably converged (around 10 meV error). The suitability of the different basis-sets was checked in the following way: A single-pillar slab of *tight* Ag with 10 atoms was created and the species was replaced with the corresponding lighter settings atom by atom.

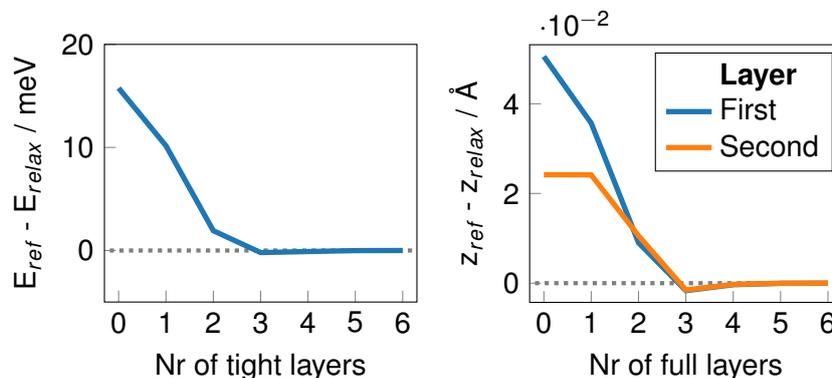
A check of the work function (figure 2.2) shows, that a suitable convergence of the work function can already be achieved when only 2 of the atoms are calculated with the better species definitions, taking only *tight* Ag, 5 layers of silver are needed for the same convergence. The computational offset of this method is comparable to 2-4 layers of tight silver because there are always 8-10 atoms present. However this is outweighed by the better convergence behavior. Differences of the electron densities were also checked, but already for 3 tight layers the fluctuations in the surface region are negligible.



**Figure 2.2.:** Work function evolution and mean SCF-cycle time of an Ag pillar of 10 atoms when gradually replacing the basis set.

## Slab Relaxation

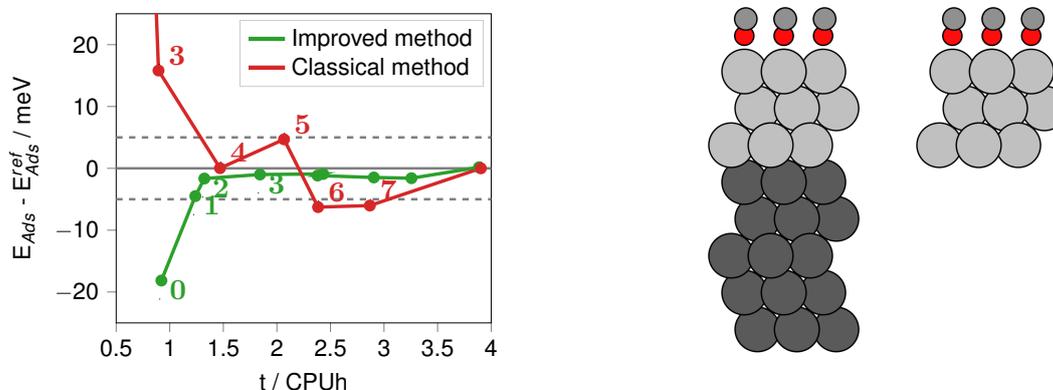
To allow for the geometric interaction of the surface with adsorbed molecules, usually the uppermost two layers are relaxed during a geometry optimization and the lower lying layers are held fixed. Therefore the number of *tight* Ag layers on the surface needed for correct reconstruction has to be checked. This was done by creating a slab with 8 layers, fixing the 6 lowermost and successively replacing the *reallylight* Ag species with *tight* Ag from the top. The results in figure 2.3 suggest to use three layers of *tight* Ag which leads to an error in energy below 0.5 meV and a geometric deviation below 0.001 Å .



**Figure 2.3.:** Evolution of surface relaxation with increasing number of *tight* Ag layers starting from an 8-layer slab with *reallylight* Ag. Left: Energy difference to a relaxed *tight* Ag only reference slab. Right: deviation of the z coordinate for the top, and second layer from a reference slab.

## Convergence of Adsorption Energy

To further check the method, the adsorption energy of carbon monoxide on a silver slab with 8 atoms and subsequently altered species from *tight* to *reallylight* was investigated. Convergence tests for the adsorption energy with respect to the number of *tight* layers and k-points showed that 3 layers of *tight* Ag and 36 k-points show reasonable convergence accuracy. The full tests are shown in section C. Figure 2.4 shows the dependence of the adsorption energy accuracy on the computational time invested for successive replacement of *tight* Ag with *reallylight* Ag (green) and increasing number of *tight* Ag layers (red) at 36 k-points. It is again visible that the improved method reaches convergence with 2 *tight* Ag layers whereas otherwise 7 to 8 layers would have been needed.

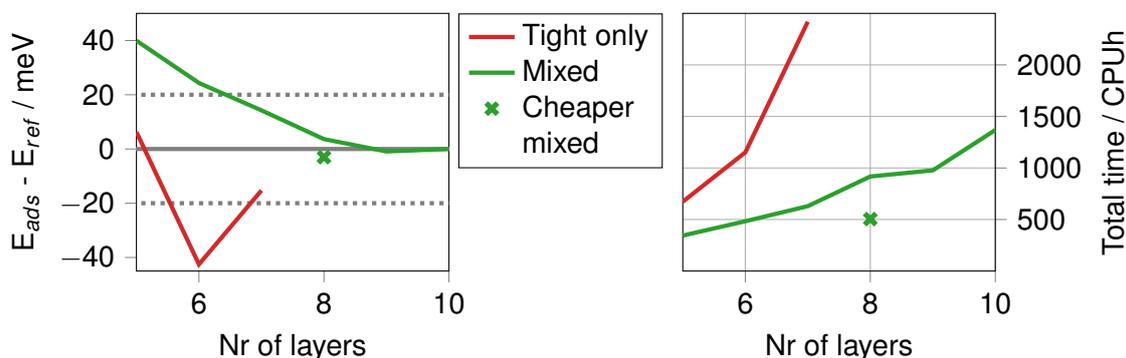


**Figure 2.4.:** Left: Adsorption energy accuracy dependent on the calculation time for a layer convergence with *reallylight* Ag and 36 k-points, The number of *tight* layers is written next to the calculation points. Right: Adsorption geometry of CO on an Ag-slab with three *tight* Ag layers for both methods.

## Number of Layers Needed

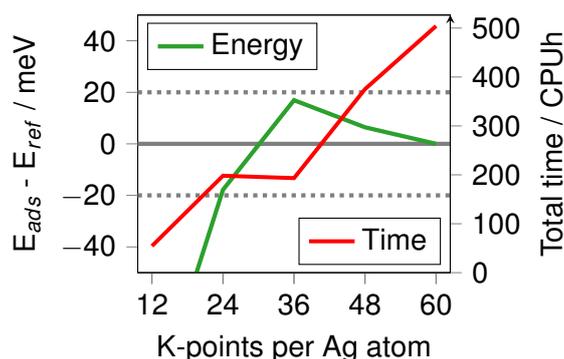
All tests up to now were for a small toy system and with a fixed total amount of layers. For a reliable estimation of the error, the adsorption energy for pentacenequinone, which will be used later, with respect to the total number of slab layers was investigated. Therefore a slab with 3 layers of *tight* Ag was taken and layers of *tight* Ag or *reallylight* Ag were added. To minimize the k-point error, those calculations were performed with 60 k-points. The evolution of the adsorption energy error and computational expense are shown in figure 2.5. Due to computational cost and convergence issues the calculations with solely *tight* Ag were carried out only up to 7 layers.

For 8 layers the adsorption energy with the mixed method is converged to 3 meV but with the cost of 1000 CPUh for a single calculation. Because this is still too expensive for structure search, 8 layers were recalculated with looser settings for the *tight* Ag (tabulated in listing 2). The resulting adsorption energy is displayed as green cross in figure 2.5. This minor adjustment changes the adsorption energy by only 6 meV but cuts the computational cost by nearly one half.



**Figure 2.5.:** Layer convergence for pentacenequinone. The calculations were performed with 3 layers of *tight* Ag and N-3 layers of *reallylight* or *tight* Ag below. The point marked with a cross is a single calculation with cheaper settings described in the text.

With the cheaper settings also a k-point convergence test for a 8-layer slab was carried out, which is shown in figure 2.6. It indicates, that with an adsorption energy error of around 20 meV, another halving of computational expenses can be achieved.



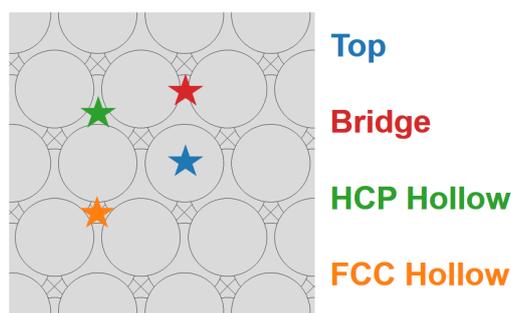
**Figure 2.6.:** K-point convergence for PQ on Ag(111) with mixed Ag species.

## Final Settings

Taking all convergence tests into account, the k-point density was chosen to 36 per substrate atom. The slab used for further calculations consists of 3 layers of *tight* Ag and 5 layers *reallylight* Ag underneath. The species definitions and full settings for the further calculations are tabulated in section C.

## 2.3. Local Adsorption Geometry Search

One important part of the SAMPLE approach is the search for local adsorption geometries. For such a search for local energetic minima, the number of degrees of freedom (DOF) is  $3N-6$ , where  $N$  is the number of atoms allowed to move. For the surface slabs used here, the two uppermost layers of silver and the molecule have to be optimized, which leads to several hundred degrees of freedom. But luckily it can be assumed that the slab and molecule will not change their general structure strongly, so for rigid molecules, the most important degrees of freedom are the rotations and movement of the molecule with respect to the surface. As large molecule-substrate interactions are expected, this six-dimensional subspace (3 rotations and 3 translations for the molecule) is further reduced to only 3 important dimensions, namely the movement along the surface plane and the rotation of the whole molecule on the plane. The vertical distance between the lying molecule and the substrate influences the adsorption energy strongly, but is also fixed as it is similar for all x-y positions.



**Figure 2.7.:** High-symmetry points on the fcc(111) surface.

Now a rough presampling of the surface can be performed within this three-dimensional subspace. Full geometry optimizations are then performed from selected points on this potential energy surface (PES). Up to now, the starting points were selected by positioning the molecules center of mass above all high symmetry points of the substrate (for fcc see figure 2.7) and rotating it in discrete angle-steps. From all those points full optimizations were performed. This approach led to slow optimization convergence and possibly missed minima, because the starting points were often far from local minima of the reduced PES.

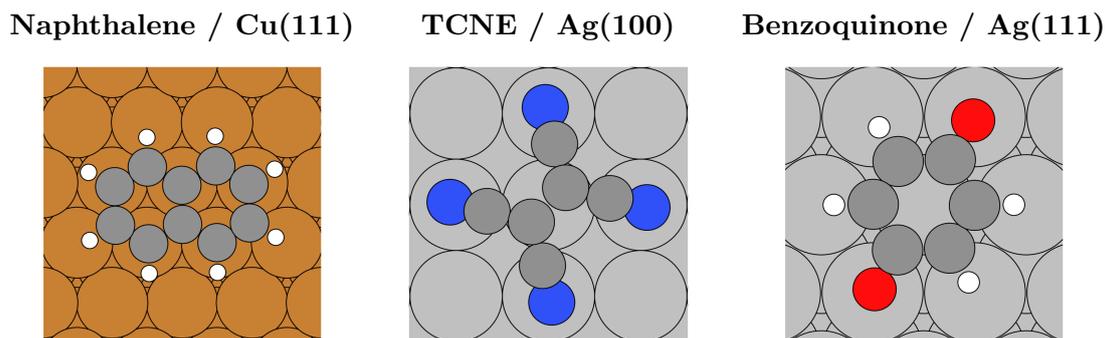
To eliminate this problem, a reasonable presampling of this reduced PES needs to be performed. One possible method is presented here:

First the approximate adsorption height is determined by putting a single molecule on a point on the surface, that is supposed to be high in adsorption height (e.g. a top position). Then a geometry relaxation restricted to the z-coordinates of the molecule is performed. With this adsorption height, the molecules are placed on all high symmetry points and then rotated to distinct points on the surface. In contrast to the old method explained above, instead of full, CPU time consuming, geometry optimizations, now only single points are calculated. The energies for the full rotations are then interpolated with cubic piecewise polynomial splines using the torque on the molecule as the first derivative of the energy with respect to the rotation. The full mathematical derivation is depicted in section F. From the minima on those splined curves, full geometry optimizations are performed. Some examples for this method are shown in section 2.3.

This approach still uses only the PES around the high symmetry points. An improvement can be achieved with the help of Gaussian process regression, which allows a continuous interpolation of the full subspace.

## 2.4. Improvements to the Learner

The only molecules the SAMPLE approach was already successfully applied to are naphthalene and TCNE [51, 54] (figure 2.8). Naphthalene is solely surrounded by hydrogen atoms, which dominates the interactions with other naphthalenes. For TCNE, the description of pairs using the nitrogen atoms is unique and as nitrogen and carbon have similar van der Waals radii, there is also no problem when defining minimal distances between molecules. To describe the pairwise interactions of molecules with multiple surrounding species correctly, some improvements to the method were implemented. In the following section these improvements will be shown on the example of benzoquinone on Ag(111) (see figure 2.8). It is surrounded by hydrogen and oxygen atoms which have very different size and interaction range, which makes a distinction in the characterization necessary.



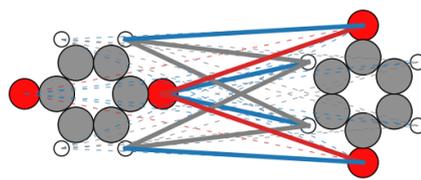
**Figure 2.8.:** Left and middle: molecules and surfaces already used for structure searches. Right: system used for learner improvements.

### 2.4.1. New Feature Vector

The feature vector (FV) introduced in section 1.5.3, which defines the similarity between pairs, does not distinguish between different species. As shown in table 2.1, the interaction range is very different for different species combinations. To encode this physical requirement, the feature vector was split into sub-vectors, which all consider only one specific species pair. The entries of the sub-vectors are then sorted and cropped at a specific length. In principle this approach is a stacked combination of the standard FV (see section 1.5) for each species pair.

In equation 2.1, the structure of such a feature vector for two geometries  $g1$  and  $g2$  and a distinction between species H and O is shown. The minimal distances  $d_{min}^{XY}$  are hyper parameters that need to be chosen reasonably.  $r_i$  and  $r_j$  are positions of atoms on different molecules.

$$\mathbf{f}(g1, g2) = \left( \begin{array}{c} \left( \frac{d_{min}^{OO}}{|r_i^A - r_j^O|} \right)^n \\ \vdots \\ \left( \frac{d_{min}^{OH}}{|r_i^A - r_j^H|} \right)^n \\ \vdots \\ \left( \frac{d_{min}^{HH}}{|r_i^B - r_j^H|} \right)^n \\ \vdots \end{array} \right) \left. \begin{array}{l} \left. \begin{array}{l} \left. \left( \frac{d_{min}^{OO}}{|r_i^A - r_j^O|} \right)^n \right\} \mathbf{f}^{OO}(g1, g2) \\ \left. \left( \frac{d_{min}^{OH}}{|r_i^A - r_j^H|} \right)^n \right\} \mathbf{f}^{OH}(g1, g2) \\ \left. \left( \frac{d_{min}^{HH}}{|r_i^B - r_j^H|} \right)^n \right\} \mathbf{f}^{HH}(g1, g2) \end{array} \right\} \end{array} \right. \quad (2.1)$$



**Figure 2.9.:** Representation of the feature vector for species O and H.

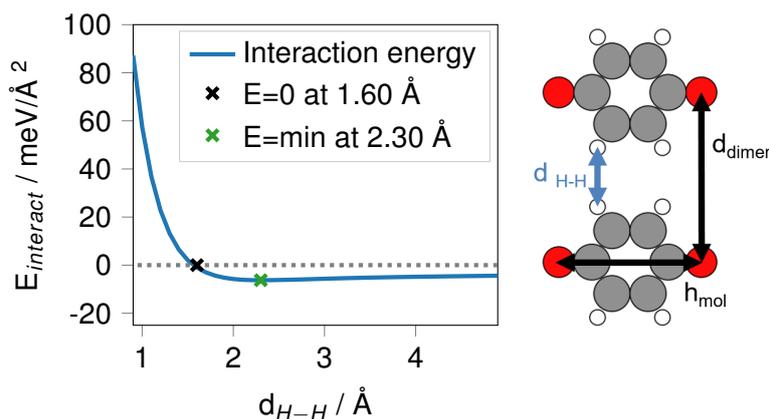
This new feature vector allows to distinguish between different species and thereby interaction ranges. A test of the new FV and all other improvements introduced in this section is shown in section 2.4.5.

## 2.4.2. Choice of Minimal Distances

For the generation of all configurations and also the learning of the interactions, a physically reasonable minimal distance between all different species occurring in the molecule has to be provided, but there is no straightforward way of choosing those minimal distances. To find reasonable values, the interaction energy of selected BQ pairs dependent on the dimer distance was probed. The pairs were selected such, that the dominating interactions are solely between a single species on either molecule. The interaction energies are then corrected by the adsorption energy of single molecules on the surface:

$$E_{interact} = \frac{E_{interact}^{gasphase} - E_{ads}^{substrate}}{d_{dimer} * h_{mol}} \quad (2.2)$$

where the dimer distance  $d_{dimer}$  (measured at the molecule centers) times  $h_{mol}$ , the length of the molecule orthogonal to the movement, represents the area a single molecule would occupy on the surface (see figure 2.10). This allows for an estimation of the interaction energy on the surface. A value lower than zero means, that if the molecule-surface interaction does not change the molecule-molecule interaction, at this distance it would still be energetically favorable for a line of molecules on the surface to stay on the surface. For values higher than zero, the line would gain energy by desorption. As the remaining of the molecules on the surface is a crucial requirement for surface structures, the distances where the interaction energy becomes zero were chosen as the distance thresholds. From equation 2.2 it is also evident, that this point does not depend on the length of the molecule, but only on the distance and interaction energy, which makes it a very robust parameter.



**Figure 2.10.:** Interaction energy evolution of a BQ dimer used to probe the minimal distance between hydrogen atoms.

The results for all distance thresholds are shown in table 2.1 and the exemplary pair energy curve of the H-H distance of benzoquinone on Ag(111) can be seen in figure 2.10. All potential curves are shown in section E. From the results it is evident, that the different species pairs have distinct scaling behavior, making a diverse treatment necessary. It is also noteworthy, that the minimal distance between two hydrogen atoms, and the one between hydrogen and oxygen is the same, whereas the distance between hydrogen and carbon is much larger. This indicates that the system is likely to form hydrogen bonds which could stabilize certain configurations.

**Table 2.1.:** Final distance thresholds for benzoquinone mapped from gas phase pairs

species pair	min distance / Å
C-H	2.30
H-H	1.60
C-O	2.50
O-H	1.60
O-O	2.40

### 2.4.3. New Covariance Kernels

Because of the new feature vector, the old definition of the real-space kernel (see equation 1.74) is no more applicable, as the definition of the minimal distance is no more unique. Therefore four new kernels were tried. The first takes only one minimal distance parameter and uses the minimum of all species pair distances for the uncertainty.

$$\text{ignore: } \sigma_{int} \cdot \exp \left[ -\frac{d_{min} - d_{pair,min}}{\tau} \right] \quad (2.3)$$

This causes the minimal distances of the smallest atoms (most often hydrogen) to be the dominating factor for the uncertainty. A slight improvement is achieved by the *minimal* kernel, which takes a different minimal distance for all pairwise interactions and then takes only the maximal variance for the prior.

$$\text{minimal: } \sigma_{int} \cdot \max \left( \exp \left[ -\frac{d_{min}^{AB} - d_{pair,min}^{AB}}{\tau^{AB}} \right] \right) \quad (2.4)$$

If several species pairs have similar values, this could influence the effective uncertainty. Another option is to build a mean of all variances by summing up all contributions and then dividing by the number of species combinations.

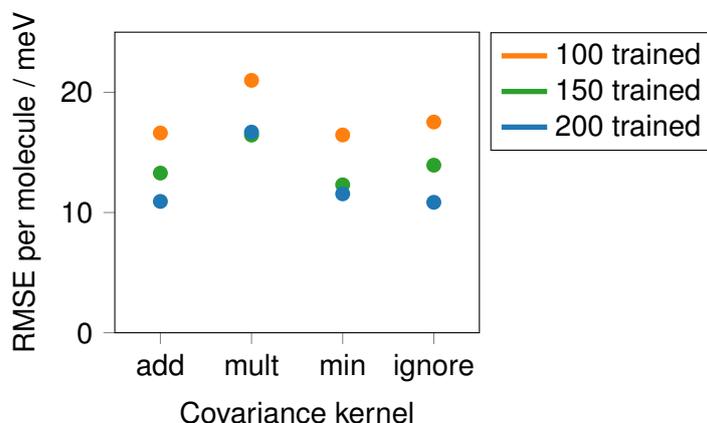
$$\text{additive: } \sigma_{int} \cdot \frac{1}{N_{AB}} \sum_{AB} \exp \left[ -\frac{d_{min}^{AB} - d_{pair,min}^{AB}}{\tau^{AB}} \right] \quad (2.5)$$

This leads to a kernel that considers not only the largest, but also smaller variances and should be well suitable for systems with only a few different species. If the number of species becomes too large, the averaging will lead to an underestimation of the uncertainty. One more possibility is to multiply all variances.

$$\text{multiplicative: } \sigma_{int} \cdot \prod_{AB} \exp \left[ -\frac{d_{min}^{AB} - d_{pair,min}^{AB}}{\tau^{AB}} \right] \quad (2.6)$$

This approach is frequently used in Gaussian process regression, but leads to a significant, and in this application unwanted, reduction of the correlation between pairs.

The performance of the different kernels for the training of BQ in gas phase is shown in figure 2.11. Their applicability is due to be tested with more systems. For this system there is no significant difference between *additive*, *minimal* and *ignore*, but the *multiplicative* kernel performs noticeably worse.



**Figure 2.11.:** Performance of the different kernels for a BQ gas phase monolayer. From a test set of 2370 calculations, 100, 150 and 200 were chosen D-optimal for training and all other configurations were predicted.

### 2.4.4. Improvements of the Prior

For the reliable energy prediction of all constructed configurations, a certain number of training calculations is necessary. With Bayesian linear regression, this number can be reduced by introduction of physical knowledge into the model via the mean and covariance of the prior. For the interaction of the molecule with the substrate, the adsorption energy can be used as a suitable prior mean. On the other hand, the molecule-molecule interactions on the surface are manifold. Some examples are electrostatic interactions, van der Waals forces, formations of chemical bonds, Pauli repulsion or exchange and correlation energies. An approximate estimation without quantum mechanical calculations is possible only for the electrostatic interactions and for van der Waals interactions. As the electrostatic prior is still quite expensive to calculate and has not shown significant improvement for our systems, the formalism needed is only shown in section D.

**Van der Waals Forces** can easily be computed with the correction suggested by Stefan Grimme[19] which was already introduced in section 1.1.5. In this thesis the interaction energy between two molecules is estimated by equation 1.38 in sixth order. The free molecule coefficients  $C_6^{AA}$  and  $R_0^{AA}$  were taken from Grimme’s publication[19]:

Species	H	C	N	O	F
$C_6^A / \frac{eV}{\text{\AA}^6}$	1.66	17.10	11.50	7.25	5.90
$R_0^A / \text{\AA}$	1.11	1.61	1.55	1.49	1.43

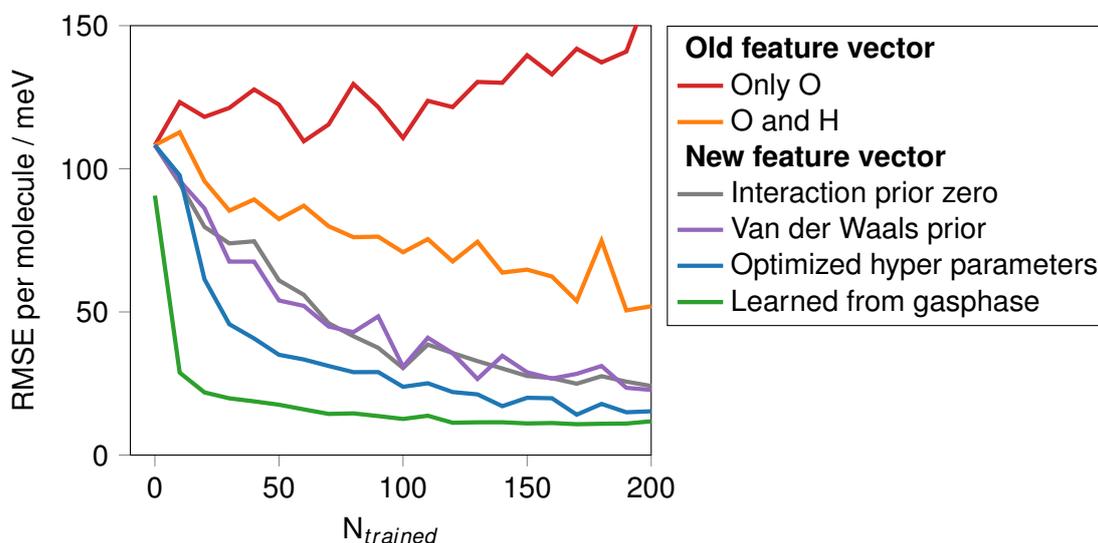
For PBE the  $s_6$  coefficient was chosen to  $s_6 = 0.94$  and the damping constant in the damping function was selected to  $\alpha = 23$ . Those are the same values as used inside the FHI-aims code.

**Learning from Gas Phase** on the other hand is a useful tool if the dominating interactions in the gas phase are similar to those on the surface, which is true for hydrogen bonds and van der Waals interactions, but completely wrong for charge transfer driven interactions. If applicable, pairwise interaction energies from gas phase calculations can be used as a prior for the pairwise interactions on the substrate.

An application of those two prior improvements to benzoquinone on silver is shown in section 2.4.5.

### 2.4.5. Comparison of the Improvements

To test the different improvements introduced in this section, a RMSE evolution for the training of benzoquinone on Ag(111) was calculated. Therefore all substrate calculations of BQ were used (220 with 1 or 2 molecules per UC, 27 with 3 or 4 molecules per UC). For the feature vector the atom distances between all hydrogen and oxygen atoms were used, if not noted differently. Then a training set with size  $N$  was chosen D-optimal from the calculations with 1 or 2 molecules per UC and all other configurations were predicted. All results can be seen in figure 2.12.



**Figure 2.12.:** Comparison of the different feature vectors for BQ on Ag(111). The RMSE per molecule is calculated by D-optimal training from a set of surface calculations and prediction on the remaining data points.

For the first test, only the oxygen atoms of the BQ molecules were used to define the FVs. The error increases because some configurations in the test set can not be learned and as the test set becomes smaller, their influence increases. The old feature vector with oxygen and hydrogen atoms is able to decrease the error by learning, but not to a satisfactory level.

The new feature vector allows for a significantly increased prediction result, but the use of a van der Waals prior leads to no real improvement. An optimization of the hyper parameters allows for the RMSE to decrease even further. To get even better results, first a gas phase system with pairs equivalent to the substrate was trained with 2850 data points. The pair energies of this system were then used as interaction prior for the substrate system. What is remarkable about this approach, is that the RMSE is very low right from the beginning. The best value of the optimized learner is here already reached after 60 data points.

## 3. Application

To date the SAMPLE approach has only been tested for small and highly symmetric molecules (TCNE and naphthalene). Molecules used in organic electronics are often much larger. The aim of this thesis was to extend the applicability of the approach toward larger systems.

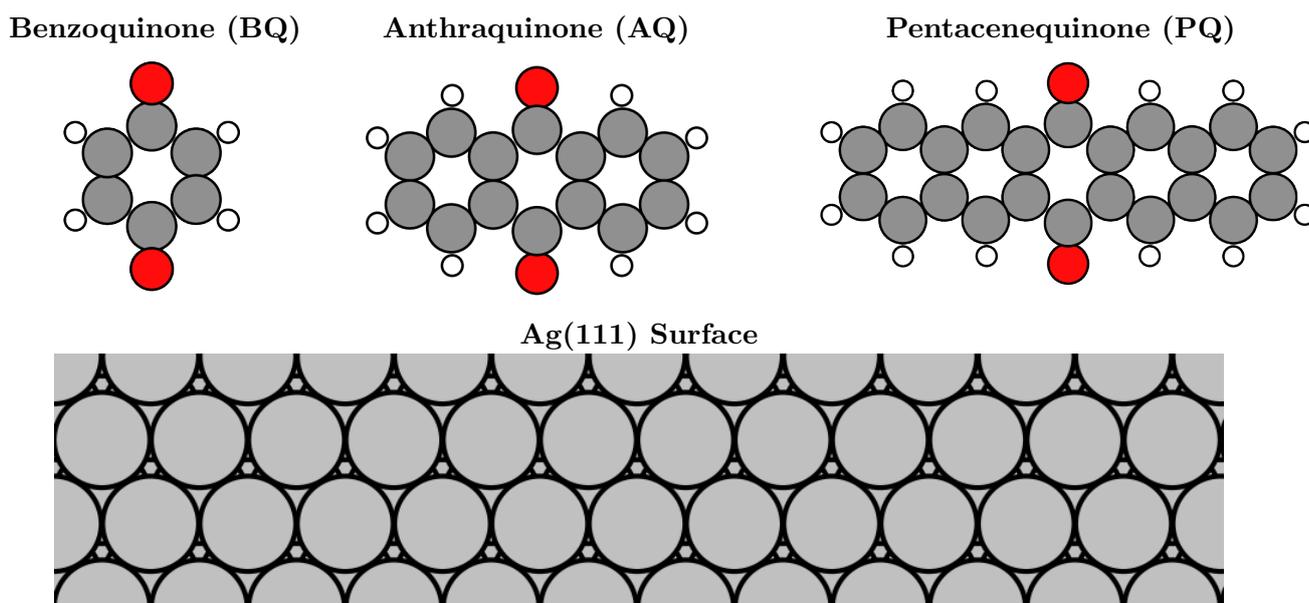
### 3.1. Systems of Interest

For this extension of system size, a system with specific characteristics was needed. The main functional groups of the molecule should not change with molecule size, as otherwise the results would not be directly comparable. Also the smallest unit, which is used for methodology development, should have a size, where a larger set of test calculations is within the available computational expense.

With these requirements in mind, acenequinones were chosen as the ideal playground for the advance towards larger molecules. Their smallest representative, benzoquinone, allows reasonably fast and cheap calculations which are needed for method development. Adding a benzene ring on either side leads to a significantly larger molecule, which still has similar interactions with the surface. With two more benzene rings, the even larger molecule pentacenequinone is obtained. The *ab-initio* structure search for this molecule is at the computational limits realizable with the available resources (2 high-performance clusters).

The Ag(111) substrate as a surface was chosen, because the interaction with the molecules was expected to be between physisorption and chemisorption (at least for pentacenequinone [55]), which was expected to lead to commensurate structures, but also rich surface polymorphism.

The molecules and substrate surface used in this thesis are visualized in figure 3.1.



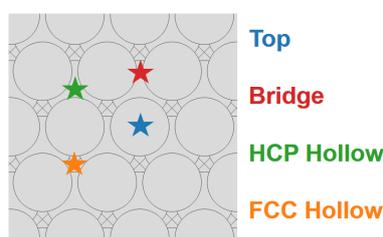
**Figure 3.1.:** Molecules used for structure search on Ag(111).

## 3.2. Local Geometry Search

For the SAMPLE approach, the first step is to find local adsorption geometries. Therefore the search method introduced in section 2.3 was used. The procedure is here explained in detail on the example of benzoquinone. For anthraquinone and pentacenequinone the same practice was used.

### 3.2.1. Search Procedure

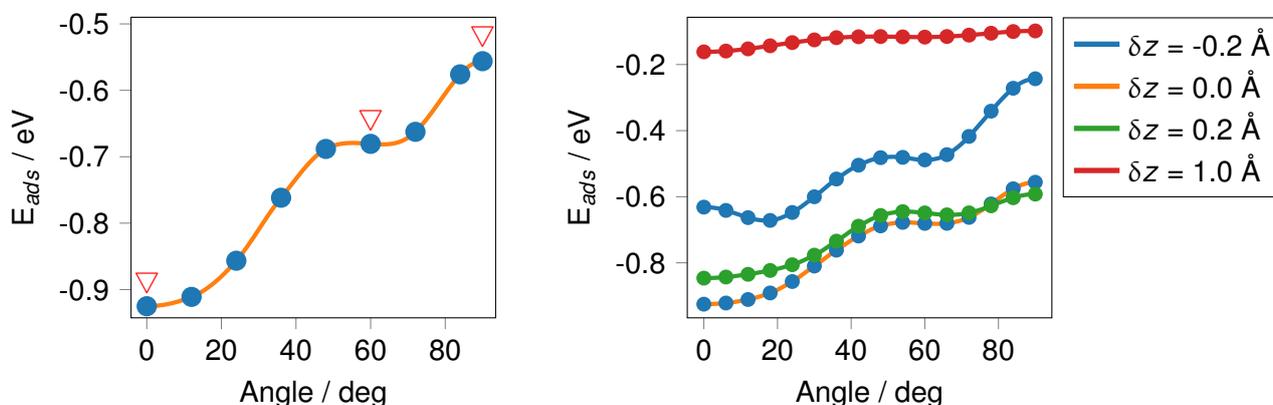
First a geometry relaxation of the BQ molecule in the gas phase was performed. Then a single molecule was relaxed on a position at the surface which was suspected to have a large adsorption height. With this resulting adsorption geometry lifted by  $0.1 \text{ \AA}$ , the rotational potential energy surface was mapped around the high-symmetry points depicted in figure 3.2. For all high symmetry points single-point calculations of 6 to 8 geometries equally distributed in the symmetrically inequivalent rotation area were performed. As all three molecules share the same symmetries, this means that for the bridge position, 90 degrees, and for the other high symmetry points 30 degrees had to be screened.



**Figure 3.2.:** High symmetry points on the fcc(111) surface.

An example of the adsorption energy dependence of BQ for rotation at the bridge position is plotted on the left side of figure 3.3. Here the starting position of 0 degrees is always the one, where the oxygen atoms are aligned vertically. For BQ at bridge the energy range is about 300 meV and a local minimum is present at around 60 degree.

Before this curve can be used for further calculations, its stability with respect to the chosen adsorption height needs to be tested. The main findings of this test (right side of figure 3.3) are, that a reduction of the height by  $0.2 \text{ \AA}$  shifts the minimum at 0 degree due to Pauli push back effects. For an increase of  $0.2 \text{ \AA}$  the curve stays qualitatively the same and increasing it by  $1 \text{ \AA}$  results in a very flat curve which still comprises the same minima. Thus it can be assumed that the chosen height is suitable and that for the rotational sampling the use of a geometry, which is slightly higher than the initial relaxed geometry is advisable.

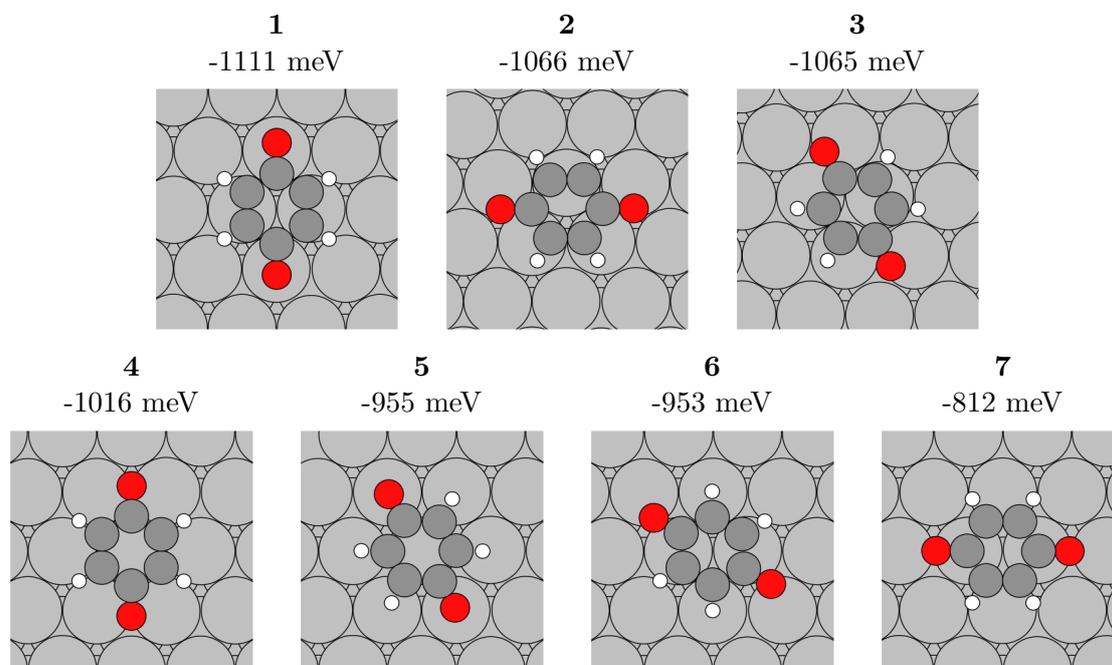


**Figure 3.3.:** Left: Rotational adsorption curve for BQ in the bridge position. Starting points for geometry optimization are marked with red arrows. Right: Adsorption curve in different heights. The reference height is  $3.53 \text{ \AA}$ .

As the exact form of the rotational PES-curves for all three molecules is not really relevant, they can be found in section G. The main findings of the rotational presampling are, that the energy spread is largest for the bridge position ( $\approx 300$  meV), intermediate for top ( $\approx 100$  meV) and smallest for the hollow positions ( $\approx 50$  meV) independent of molecule size. A significant change of the shape with changing molecule size is only observable for the bridge rotations, where more local minima are developed. From all curves, full geometry optimizations were performed from all local minima and saddle points. In the following section the resulting geometries for the three molecules are discussed.

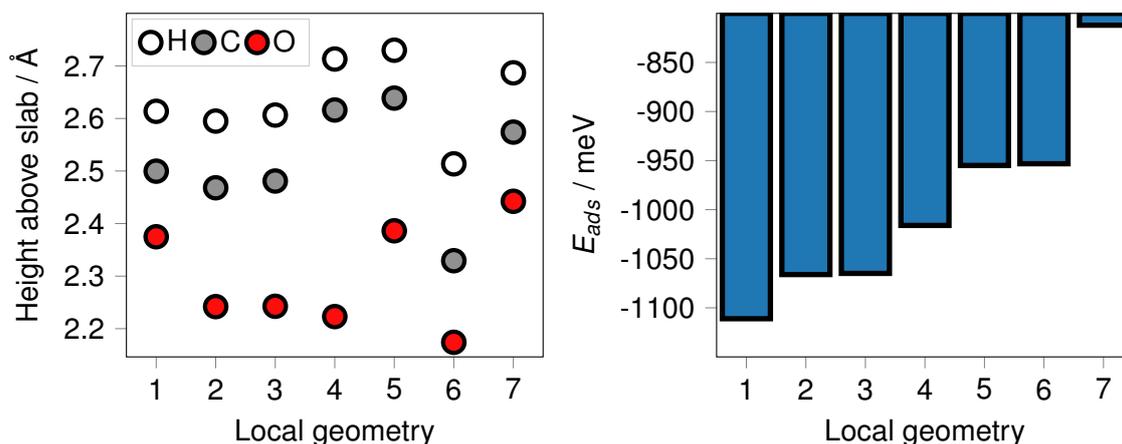
### 3.2.2. Results for Benzoquinone

For this molecule, 9 starting points for the full geometry optimization resulted in 7 distinct geometries, because three geometries relaxed into the global minimum (local geometry 1). The geometries together with their adsorption energies are shown in figure 3.4. The energetically most favorable site is  $\approx 50$  meV lower in energy than geometries 2 and 3. Those are symmetry equivalent with respect to the first Ag layer, but the different hollow sites make them geometrically distinguishable. From 3 to 4 there is again an energetic penalty of  $\approx 50$  meV and then geometries 5 and 6 are another  $\approx 60$  meV less favorable. Local geometry 7 is 300 meV worse in adsorption energy than the energetically most favorable site, thus it was neglected for the structure search due to its low adsorption energy. Nonetheless it is shown here, because it also appears in the local geometries for AQ and PQ.



**Figure 3.4.:** Local adsorption geometries of BQ with their adsorption energies. The uncertainty of the adsorption energies is  $\approx 10$  meV.

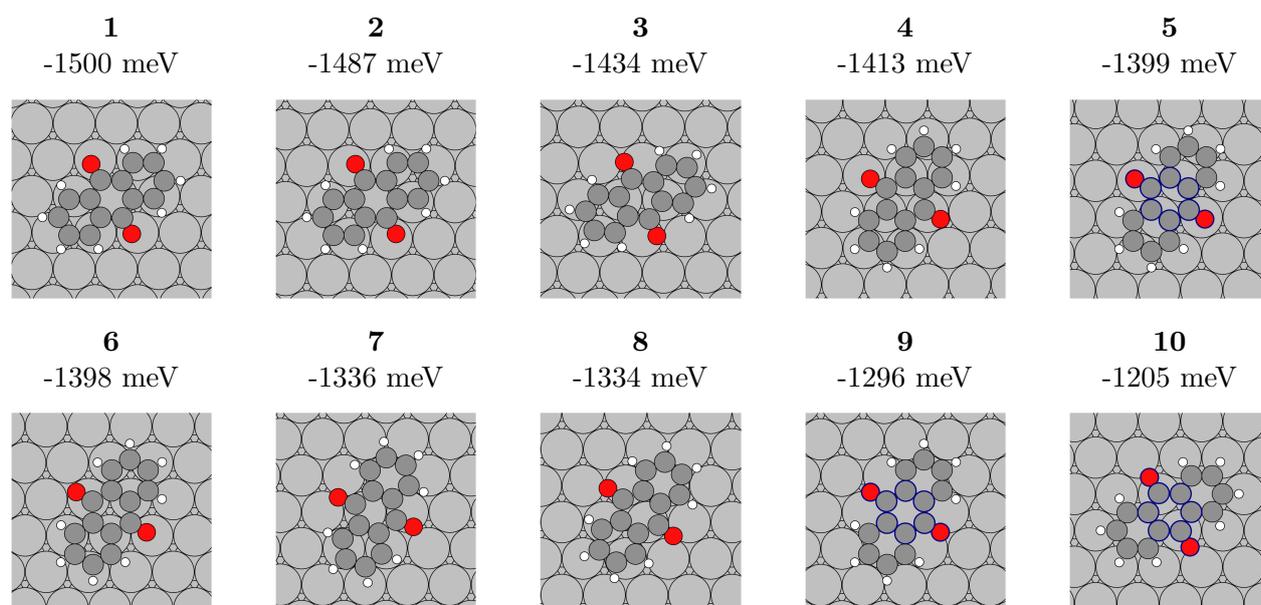
Figure 3.5 shows the mean adsorption heights for all local geometries divided into the different species (carbon, oxygen and hydrogen). The variation of adsorption height is in the range of  $0.2 \text{ \AA}$ . As expected, it mainly depends on the positions of the atoms on the surface. This means, that atoms right above metal atoms are higher in position compared to those on bridge and hollow sites. The only exception to this is the adsorption height of geometry 7, which is the only geometry where carbon atoms are directly on top of silver atoms.



**Figure 3.5.:** Adsorption heights and energies of the local geometries of BQ.

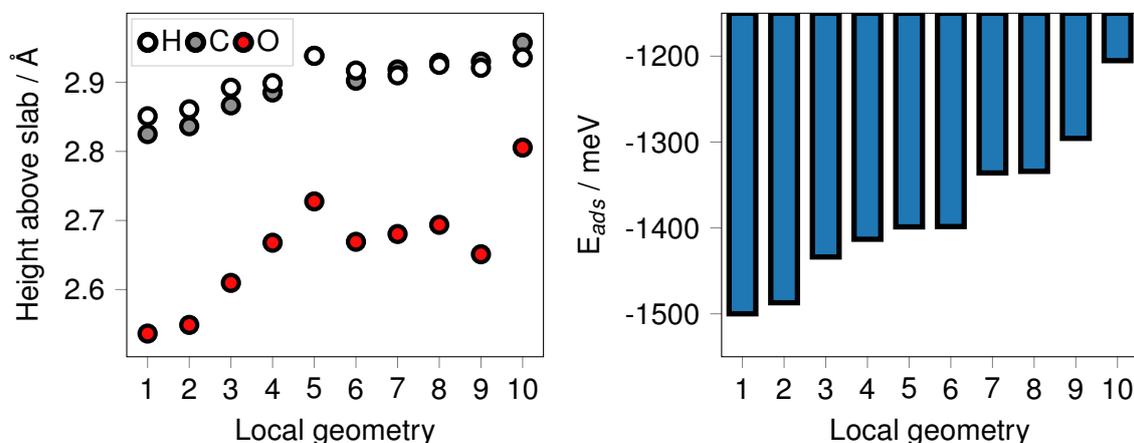
### 3.2.3. Results for Anthraquinone

Adding a benzene ring on both ends of benzoquinone leads to anthraquinone. Here geometry optimizations were not solely started from the minima and saddle points of the rotational PES sweep, but additionally a slightly more sophisticated PES-prediction method involving Gaussian process regression was used. This resulted in slightly different starting points. From the form of the resulting adsorption sites it can be assumed, that they would have also been found with the method explained in this thesis and the improved approach only led to faster results with fewer calculations. The resulting local geometries for AQ are displayed in figure 3.6. Geometries 1 and 2 are again almost symmetry equivalent with respect to the first Ag layer and their energy difference stems from the different adsorption heights, which could be a geometry convergence artifact. Around  $\approx 50$ -80 meV less favorable is a range of 4 adsorption geometries, now also with a hollow position as central adsorption point. Another  $\approx 60$  meV off are geometries 7 and 8 which are also hollow-ignoring symmetry equivalents. As already for BQ, the last structure is  $\approx 300$  meV lower in adsorption energy than the best.



**Figure 3.6.:** Local adsorption geometries of AQ with their adsorption energies. The related geometries of BQ are indicated with blue circles. Due to the convergence tests a maximum uncertainty of adsorption energy of  $\approx 15$  meV can be assumed.

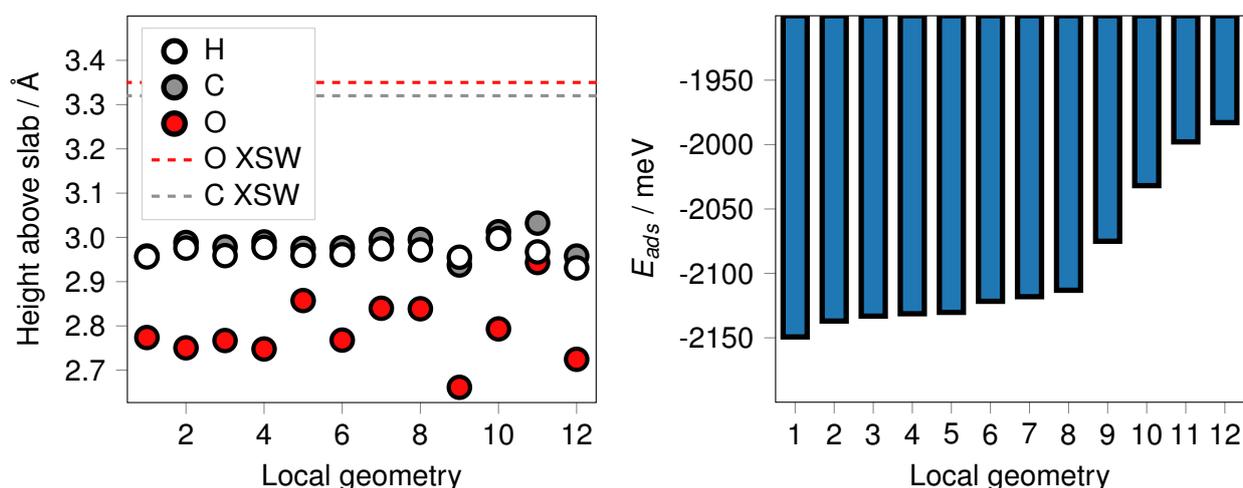
Interesting about these structures is, that the adsorption geometries of BQ can be found as direct part of the AQ geometries for geometries 5, 9 and 10. The backbone of geometries 1 and 2 is almost equivalent to BQ geometry 5, and the one from geometry 3 is very similar to the 6th of BQ, but a little bit shifted due to the additional benzene rings. The worst geometry of BQ is here also part of the worst adsorption geometry. A general observation for all those structures is, that it is energetically beneficial for the benzene-ring carbons to be at positions far from the Ag atoms.



**Figure 3.7.:** Adsorption heights and energies of the local geometries of AQ.

### 3.2.4. Results for Pentacenequinone

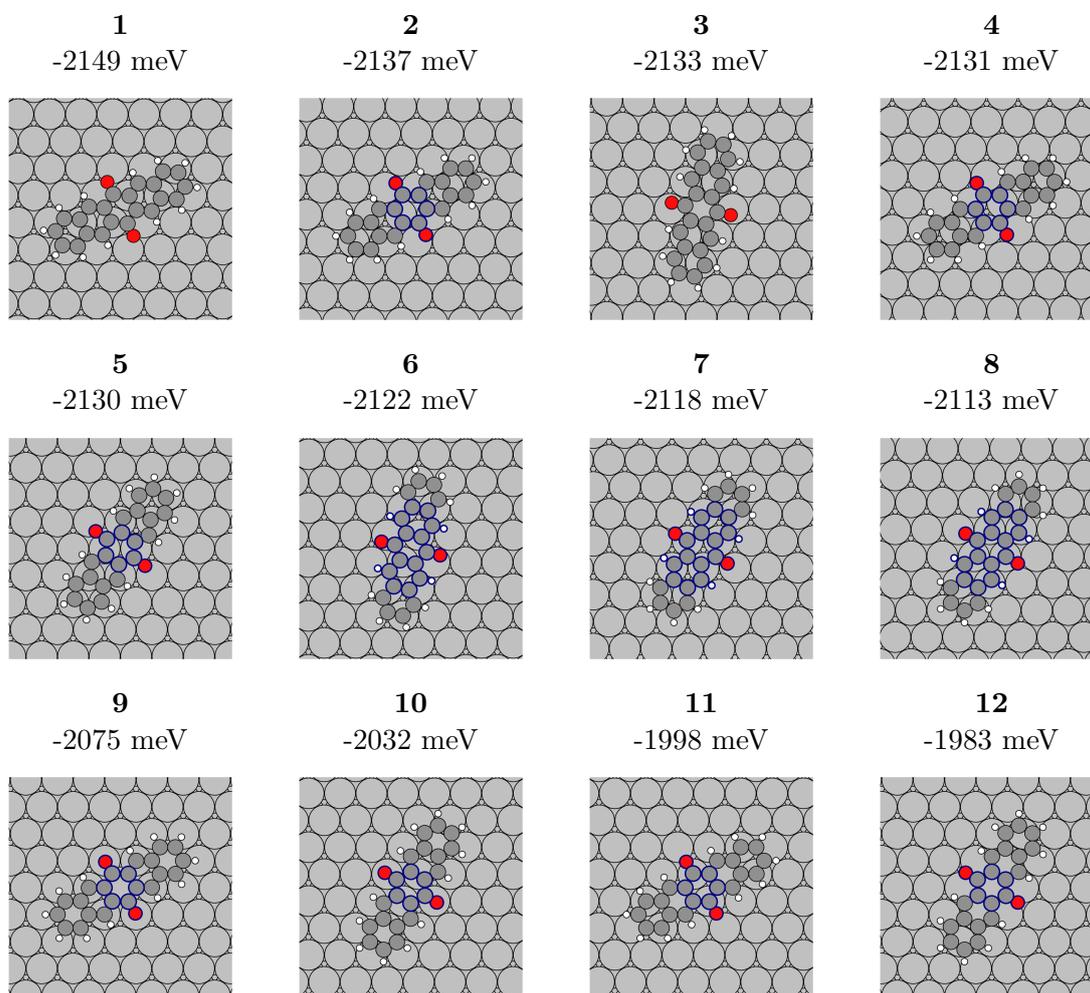
The largest molecule investigated exhibits a range of 12 distinct local geometries, which are all visualized in figure 3.9. Unfortunately, the first 8 are within an energy range of only 40 meV. Here again geometries 2 & 4 and 7 & 8 are hollow-ignoring symmetry equivalents. In contrast to the two other molecules, the total energy spread is now only  $\approx 170$  meV. Geometry 9 of AQ, which had a reasonable adsorption energy, exhibits the worst local minimum for PQ, closely followed by the geometry that was already the least favorable for BQ and AQ.



**Figure 3.8.:** Left: Adsorption heights of pentacenequinone. The dotted lines indicate experimental results obtained by Georg Heimel and colleagues [55]. Right: energies of all local geometries of pentacenequinone.

In figure 3.8 the species resolved adsorption heights are plotted and compared to experimental results obtained by XSW [55]. To date no clear indications for the large difference between simulation and experiment were found, but new experiments are already being performed.

It is remarkable, that for this molecule all 7 local geometries of BQ can be identified as parts of the local geometries, and that three more can be explained with geometries of AQ. Those results indicate, that the backbone (BQ) is influencing the preferred adsorption sites strongly. The remaining distinctions arise due to preferred positions of the benzene rings, which leads to an increase in the number of adsorption sites.



**Figure 3.9.:** Local adsorption geometries of PQ with their adsorption energies. Related geometries of BQ and AQ are indicated with blue circles. The uncertainty of adsorption energies is  $\approx 20$  meV (see section 2.2).

### 3.3. Structure Search

Now that the local geometries are found, the actual task of polymorph search can be tackled. In this section the results for the structure search of benzoquinone and anthraquinone are presented.

#### 3.3.1. Preparations

Before the real structure search on the surface could start, some preparatory tasks had to be accomplished. After finding the adsorption geometries, the first challenge was to find out, how close such geometries should be allowed to come. Those minimal distances for the different species combinations have already been found for BQ in section 2.4.2 and due to the structural similarity of BQ and AQ, they were also used for the latter. With those minimal distances, all configurations in a reasonable coverage range could be constructed. In the following the word coverage will always be used in the sense of surface silver atoms per molecule. In total, for benzoquinone 279000 and for anthraquinone 3 million structures were constructed. Detailed numbers for the configurations created and the maximum coverage used for the corresponding number of molecules are shown in table 3.1. Configurations larger (and less dense) than the upper limits would not lead to energetically more favorable structures, but increase the number of configurations exponentially, due to the configurational explosion.

**Table 3.1.:** Total number of configurations constructed for benzoquinone and anthraquinone with the maximal unit cell size used for structure creation. Additionally the total number of DFT calculations used for the training in gas phase and on the substrate is depicted for both molecules.

molecules per unit cell	Structure creation				DFT calculations			
	Configurations		Max UC area		Gas phase		Ag(111)	
	BQ	AQ	BQ	AQ	BQ	AQ	BQ	AQ
1	132	2179	12	30	125	300	39	46
2	4592	17827	18	30	1068	300	181	189
3	60956	1790444	24	45	913	300	7	8
4	213904	1210220	29	51	743	0	20	0

After all configurations were constructed, there was one last step before the actual learning could start. In the SAMPLE approach there are in total 8 hyper parameters (see section 1.5), for all of which reasonable values need to be found. The physically motivated parameters, like the energy uncertainties and decay lengths can readily be estimated by human intuition. For the choice of the feature space parameters, this task is much harder. Thus for the choice of hyper parameters, sweeps with different combinations of parameters were performed for gas phase monolayers. This is a configuration of molecules on the surface, but calculated without the surface, which implies, that molecule-molecule interactions, such as hydrogen bonds, van der Waals interactions and Pauli push-back are represented well, whereas all effects that arise due to the surface, like charge transfer and other interactions with the surface, are not considered.

For all the hyper parameter tests, RMSE evaluations were performed with learners trained on configurations selected D-optimally from the available 1 and 2 molecule configurations and validated on all remaining calculations (also including 3 and 4 molecules). The hyper parameters chosen after the tests (shown in section H) are depicted in table 3.2. The number of gas phase calculations available is shown in table 3.1.

A little remark on one delicate detail about the optimization: This procedure was previously not part of the standard workflow of SAMPLE, but rather settings that seemed reasonable were chosen. This led to the belief, that for the feature decay parameter  $\gamma$ , only values between 0 and 1 should be chosen. A sweep of this parameter revealed, that the RMSE starts diverging for values lower than 0.4 and the optimum parameter for the molecules used here lies at around 10. So it was possible to cut the RMSE of several systems, used in the group, in half within a single day of convergence tests.

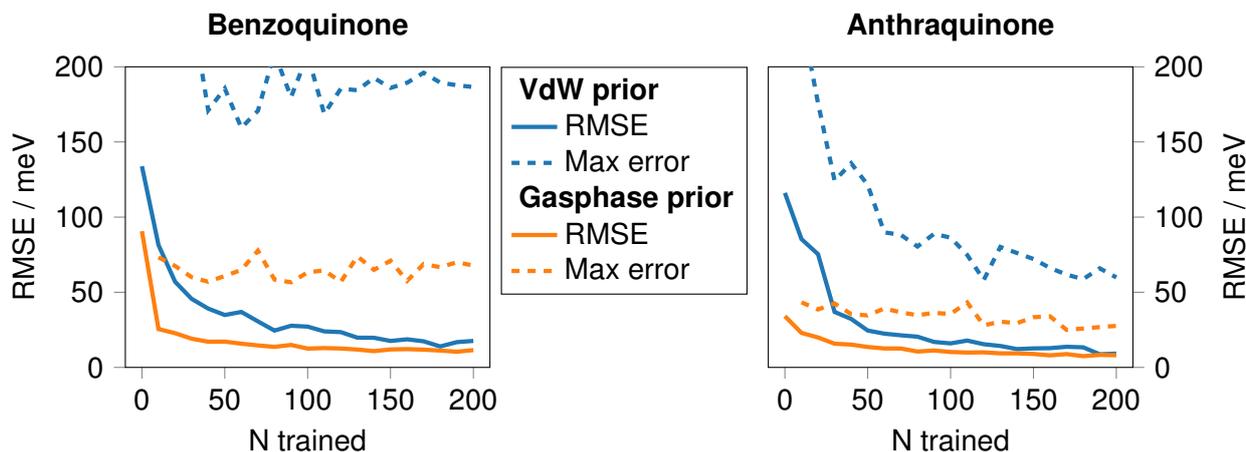
**Table 3.2.:** Hyper parameters used for the structure prediction of BQ and AQ. The full hyper parameter optimization is shown in section H.

Hyper parameter	Symbol	Value BQ	Value AQ
uncertainty adsorption energy	$\sigma_{ads}$	100 meV	100 meV
uncertainty interaction energy	$\sigma_{int}$	300 meV	300 meV
uncertainty DFT data	$\sigma_{DFT}$	10 meV	5 meV
decay length	$\tau$	5 Å	5 Å
decay power	n	3	3
decay length feature space	$\gamma$	12	9
feature threshold		0.0075	0.0075
covariance kernel		additive	additive

### 3.3.2. Inspection of the Learning Behavior

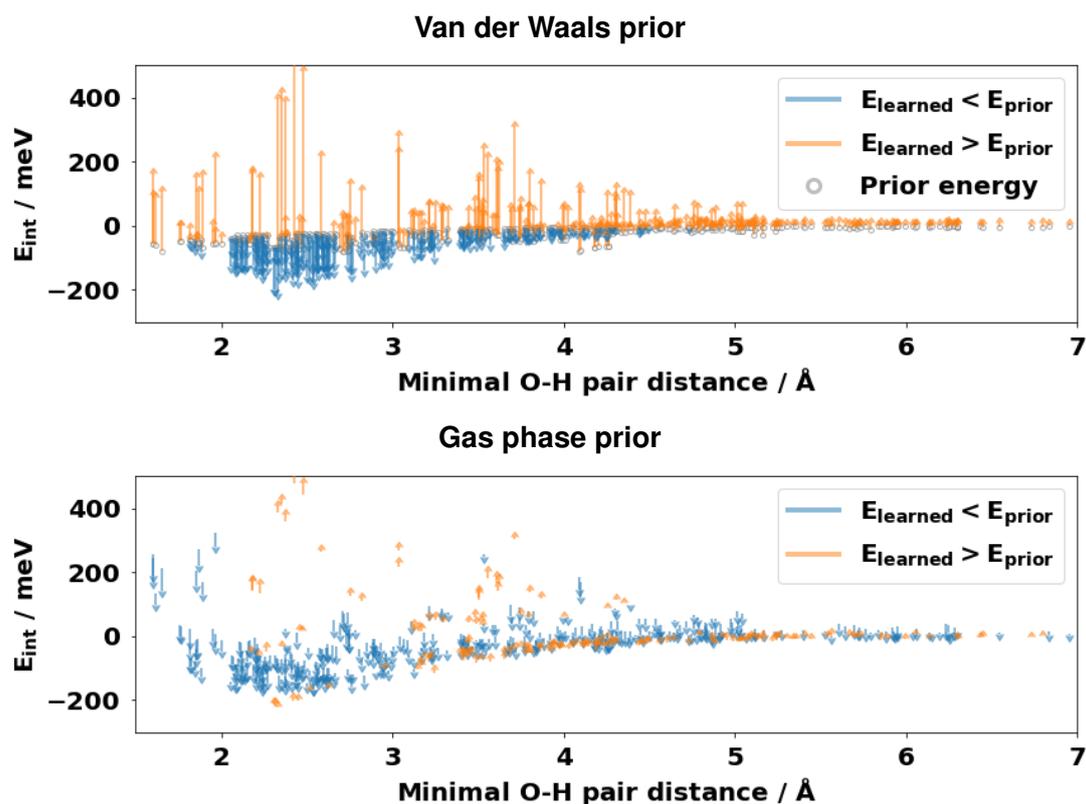
Once all configurations were created and the optimal hyper parameters were found, the structure search on the surface could be tackled. Before predictions can be presented, the reliability of the learner needs to be evaluated. Therefore, in the following a comparison of the learning behavior with the van der Waals and with the gas phase prior will be given. For the second prior, all gas phase calculations were used to train a gas phase model. From this trained model, the pair energies were taken as prior values for the substrate learner. The prior for the single-body energies is in both cases the adsorption energy of the local adsorption geometries.

One small side remark: due to the small amount of calculations, the standard RMSE, which only considers data points that have not been learned, could no longer be used reliably. Therefore the leave-one-out error was utilized. There the deviation of a single calculation from the prediction is evaluated by training with all but this configuration and subsequent prediction of this single value. The RMSE is then calculated with all those single-value errors. The maximum error in this case is simply the largest single-value error. This means that the biggest error calculated with this method is the maximum obtainable error, given the calculated configurations.

**Figure 3.10.:** Comparison of vdW and gas phase prior for benzoquinone and anthraquinone. The training/test set size was 247 for BQ and 243 for AQ. The error shown here is obtained with leave-one-out validation.

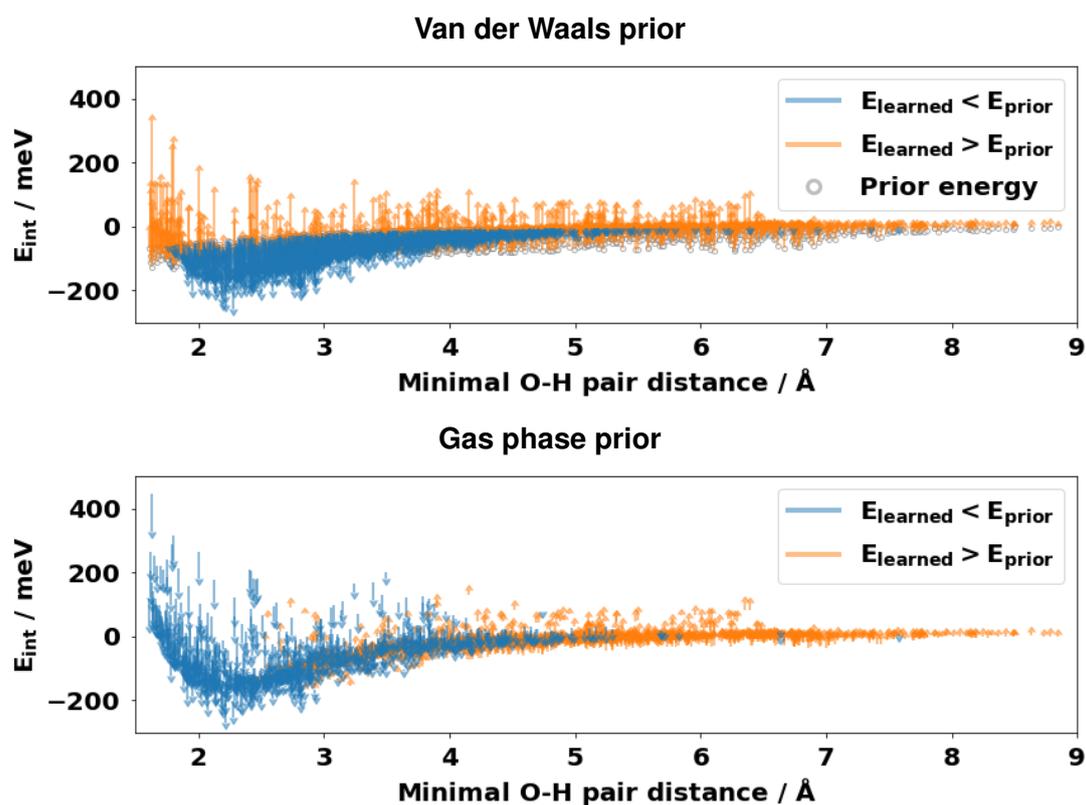
The first property of interest was the learning behavior with increasing training set size. For this, the learners were trained with  $N$  D-optimally selected calculations from the 1 and 2 molecule configurations, with  $N$  ranging from 10 to 200, and subsequent calculation of the RMSE on all other calculations. The results in figure 3.10 show, that for both systems, the gas phase prior led to a faster convergence of the RMSE and also the final RMSE was slightly lower. More important is the fact, that the maximum error was also much lower, which is beneficial for the confidence in the predictions further on.

The question now was, why the gas phase prior is so much better than the van der Waals prior. To answer this question, the evolution of the interaction energies while learning was probed. To picture this, all unique pairs in the learner were sorted by minimal distance between oxygen and hydrogen. Then, for each pair, the evolution of the energy from the prior value to the learned value was indicated with an arrow. An arrow going upwards means, that this pair interaction energy was corrected towards higher energy while training and the tip represents the final value. For the downwards arrows the same holds, but in the other direction. A comparison of the two learners for BQ with this method is shown in figure 3.11. Here it is evident, that with the vdW prior large changes in interaction energy (up to 400 meV) need to be learned, while for the gas phase prior only small adjustments need to be made. For anthraquinone (figure 3.12) the situation is a little less obvious. Here the changes of the gas phase prior are also in the range of 100 meV, but due to the high correlation of the pairs, only a few substrate calculations are needed for a good representation of the whole system.



**Figure 3.11.:** Comparison of the interaction energy change from the van der Waals and the gas phase prior for BQ.

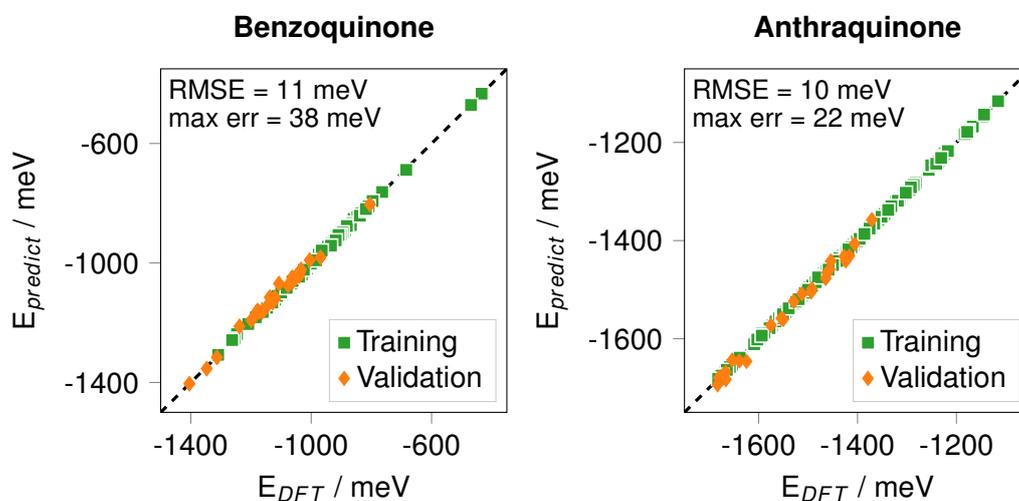
As a general conclusion, the vdW prior is not able to model the highly attractive O-H interactions in the range of 2-3 Å correctly and Pauli repulsion, which becomes important below 2 Å is also not considered. In all those plots, the general structure of a Lennard-Jones potential is recognizable. The differences to this structure are caused by other species pairs, which can be closer than the shown O-H value and have an other orientation and specific interactions.



**Figure 3.12.:** Comparison of the interaction energy change from the van der Waals and the gas phase prior for AQ.

### 3.3.3. Predicted Configurations

For the final predictions, D-optimal sets with a fixed number of configurations (usually 50) were drawn, and the learner with gas phase prior was trained until convergence of the RMSE in the range of the DFT error (around 10 meV) was reached. Thus in total 220 configurations with 1 or 2 molecules per UC were calculated and used for the training of AQ and BQ. With the trained model, then all generated configurations were predicted. To test the prediction accuracy, the best predicted structures and some more random structures with 3 and 4 molecules were calculated. The total number of structures calculated with DFT is shown in table 3.1.

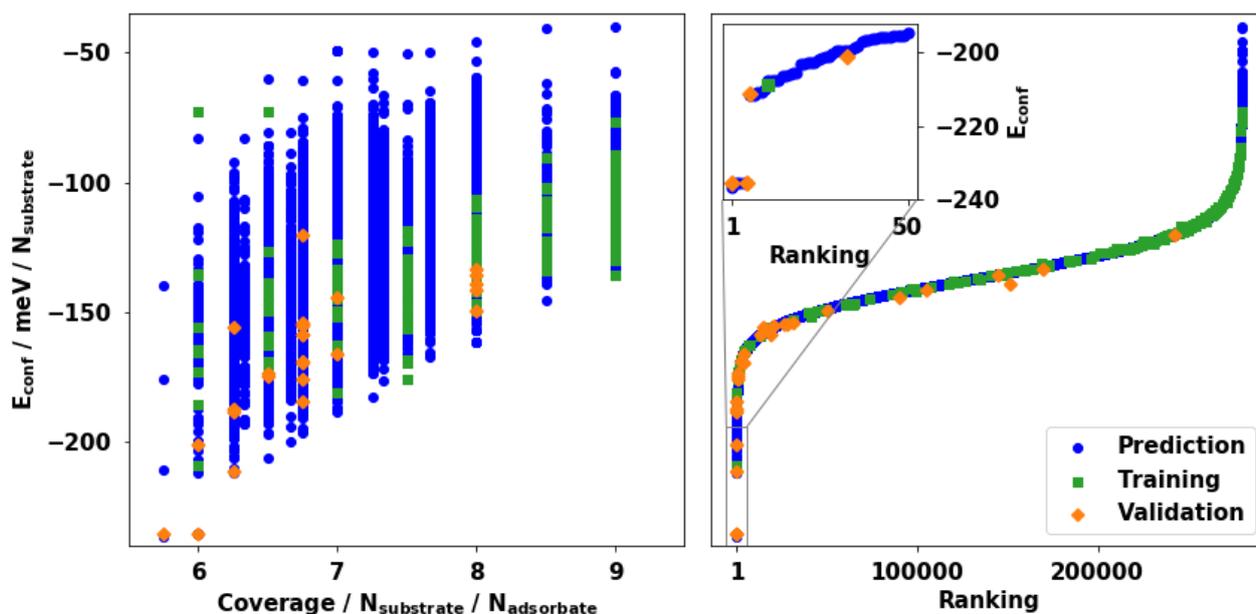


**Figure 3.13.:** Prediction versus DFT calculations for BQ and AQ on Ag(111). The errors are only with respect to the validation set.

In figure 3.13 the match of the prediction with the calculation results for final training with the gas phase prior and the full training set is shown. There the validation set for BQ consists of all structures with 3 and 4 molecules, which were never seen by the learner. For AQ the validation set consists of 16 configurations with 2 molecules and the 8 best 3-molecule configurations, because the computational cost for single configurations is much higher. As visible there, the validation RMSE could be cut to the desired 10 meV, but the maximum error is still higher.

Because a thermodynamic equilibrium with a surface molecule monolayer and an infinite molecule reservoir is assumed as experimental condition, the configuration energy per area is the property of interest. The case without infinite molecule reservoir, where the configuration energy per molecule would become the relevant property, could also be considered easily.

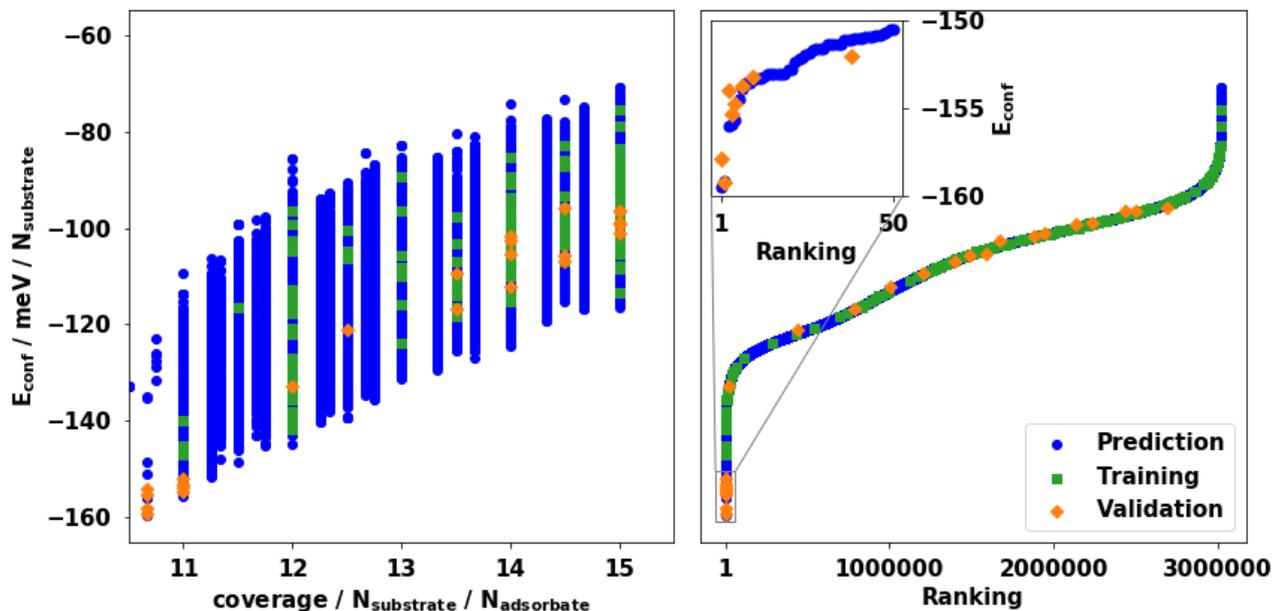
With these conditions in mind, the final prediction is plotted, ordered by coverage, in figure 3.14 for BQ and in figure 3.15 for AQ. Both systems exhibit a large spread in energy because all possible configurations were predicted, which means that also such, that are energetically unfavorable are shown in the plots. The non-integer lines arise from the cells with more molecules, for which the coverage in number of adsorbate atoms per molecule can become a rational number. As can also be seen, most training calculations were not performed at the best coverages and still the prediction of the closer packed structures is very good. This implies, that the chosen energy model, which cuts the interactions at two-body terms, is indeed able to reproduce the configuration interactions with reasonable accuracy.



**Figure 3.14.:** Prediction of configuration energies for benzoquinone on Ag(111) dependent on coverage (left) and ranked by energy (right). The validation and training points represent energies stemming from DFT calculations. The maximal uncertainty based on the maximum error of the training is  $10 \text{ meV}/N_{ads}$ .

For both molecules, the best configurations are found at the highest coverage (which is the smallest value of substrate atoms per molecule) for which configurations were created. Smaller configurations were not possible due to the choice of minimal distance. For BQ, in the beginning calculations with even higher coverage were also used. There the high Pauli repulsion energy contributions caused the predictive power of the method to deteriorate rapidly, because the assumption of a single decay length for the interactions is no more valid. This might be solved in the future by implementing a suitable kernel for the interaction prior. With the chosen minimal distances it would in principle be possible, that structures with even higher coverage are energetically more favorable, but it is not very likely as the interaction energies (figure 3.11) are already very high for the distances covered.

It is also possible, that the best structure has more than 4 molecules in the unit cell. Such structures were not considered, because it would increase the number of configurations dramatically and the predicted configurations would not be validateable because DFT calculations would be too costly.



**Figure 3.15.:** Prediction of configuration energies for anthraquinone on Ag(111) dependent on coverage (left) and ranked by energy (right). The validation and training points represent energies stemming from DFT calculations. The maximal uncertainty based on the maximum error of the training is  $2.5 \text{ meV}/N_{\text{ads}}$ .

To get a better feeling for the energy distribution of the configurations, a different method of visualization was used. In figures 3.14 and 3.15 the energies of all calculations, ranked by energy, are therefore shown. The general form of this curve is similar for both molecules. A small subset is very good in energy and another subset is very bad, while the bulk of configurations is within a smaller energy range. For the training and validation set, the blue points at a ranking position are the predicted results while the orange and green points are energies from the DFT calculations. With this visualization, the fit of the validation set to the predicted data points can be seen nicely. What is also common for both systems is, that although the training set was only chosen at specific coverages, the energies are distributed over the whole range. At first sight it seems that the BQ validation fits much better, but this is partly caused by the different energy scale used for the two systems. Both exhibit two structures, which are lowest in energy and energetically separated from all the other structures by more than the RMSE. The difference is, that for BQ this energy difference is over 20 meV while it is only around 2 meV for AQ. One has to keep in mind that the configuration energy per molecule is rather similar for both molecules whereas the area needed for AQ is almost the double of BQ.

Before the best structures are examined in detail, the energetic composition for the lowest structures is discussed. Thus in figure 3.16 the energy contributions from the adsorption and interaction energy are shown for the 50 best structures for BQ and AQ. One can see, that the variation in adsorption and interaction energy is much larger for BQ. There is one specific combination of local geometries, that is especially beneficial and a combination of this with strong interaction energies leads to the best configuration. The second best configuration has a slightly worse adsorption energy, but an equivalently good interaction contribution.

The spread in adsorption energy for AQ is by far not so high and the best structures are mainly distinguished by their good interaction energy. It will be shown in the next section, that for AQ a specific combination of local geometries is the cause of those good interaction energies. This is also the reason for the smaller energy spread and the larger number of similar structures.

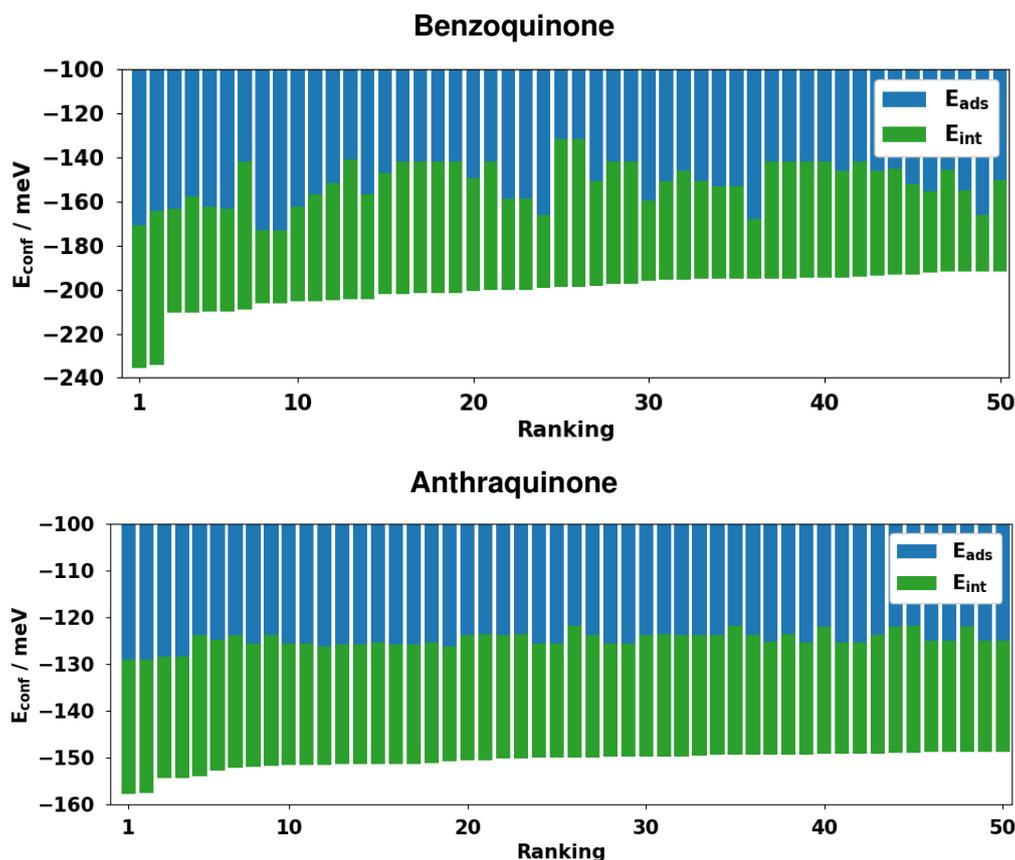


Figure 3.16.: Comparison of the energy contributions for the 50 best structures of BQ and AQ.

### 3.3.4. Explanation of the Best Structures

It is now time to take a closer look at the best predicted structures. For BQ, the two structures that are significantly better in energy compared to all other structures are shown in figure 3.17. The first has 2 molecules per unit cell, which are local geometries 1 and 4, and the second has a slightly tighter packing and additionally the second and third best local geometry.

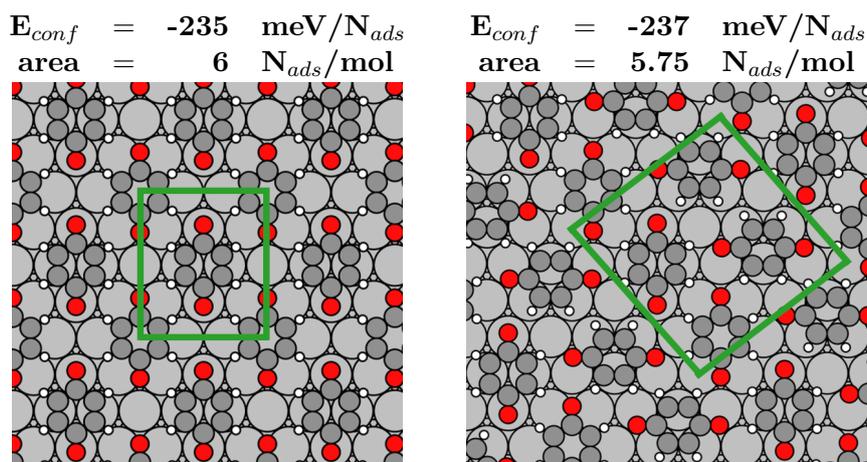
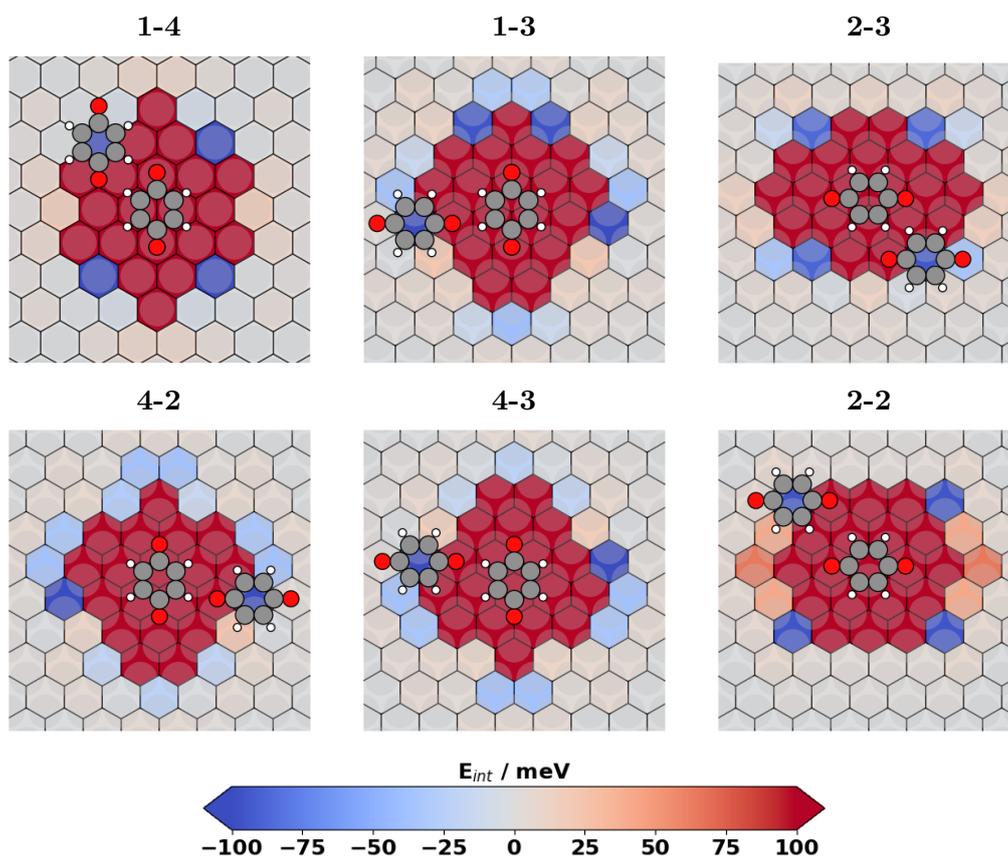


Figure 3.17.: Best configurations for BQ on Ag(111). The unit cell is indicated in green.

This is already an important factor for the favorability of those structures, but the interaction energies also need to be taken into account. Due to the energy model, the pairwise interaction energies can be visualized directly. The pair potential for a selection of geometry combinations is shown in figure 3.17.

In all those pair potential plots, the gray circles indicate the atoms of the silver surface. Each hexagon represents one pair and the color the energy of this pair. All hexagons are aligned such, that the center of the surrounding molecule is equal to the center of the hexagon. Remarkable about the interactions of those two molecules is, that for the combination of 1 and 4, there is one specific position, which is highly attractive. This position is the one that the best 2-molecule structure of BQ mainly consists of, which explains the high configuration energy. For the interactions of geometry 2 and 3 with 1 and 4 there are also very specific and highly attractive positions, from which almost all are present in the configuration. The interaction energies for hollow-ignoring symmetry equivalent geometries are almost equivalent but mirrored, which is shown on the example of the geometry combinations 4-2 and 4-3.



**Figure 3.18.:** Pairwise interaction plots for all pairs of the two best configuration for BQ.

The situation for anthraquinone is slightly different. Here two structures are equal in energy and their only structural difference is a shift of one molecule row by a single silver atom. Many of the energetically favorable structures are composed of local geometries 4/6, 7/8 and 5 and most of them have such stripe-like patterns. The energetically most favorable local geometry is only rarely present. In some configurations, local geometry 9 is also present, which then leads to kinks in the structures. The fifteen best structures are visualized in section I. The two best structures and the best with kinks are shown in figure 3.19. To understand them, the pair interaction maps in figure 3.20 are helpful. Local geometry 5 has very attractive pairs with geometries 7 and 4 at positions, that favor such stripe-like structures. For geometry 4 and 6 there are also a few pairs that allow these stripes to form. Such highly attractive pairs can thereby only form, if both hollow-ignoring symmetry equivalent geometries are combined. Using only one of them would not result in such preferential pairs. Configurations with geometry 9 are only low in energy, because one specific pair of the geometry with itself is very attractive, while interactions with other local geometries are not so beneficial.

The best structures for benzoquinone however look very promising, and experimental results are curiously awaited.

$$E_{conf} = -159 \text{ meV}/N_{ads}$$

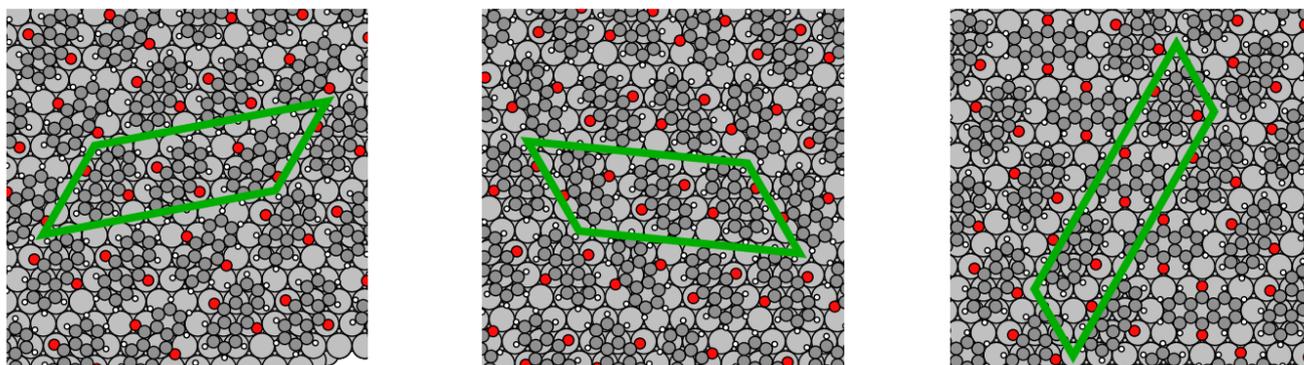
$$\text{area} = 10.66 \text{ N}_{ads}/\text{mol}$$

$$E_{conf} = -159 \text{ meV}/N_{ads}$$

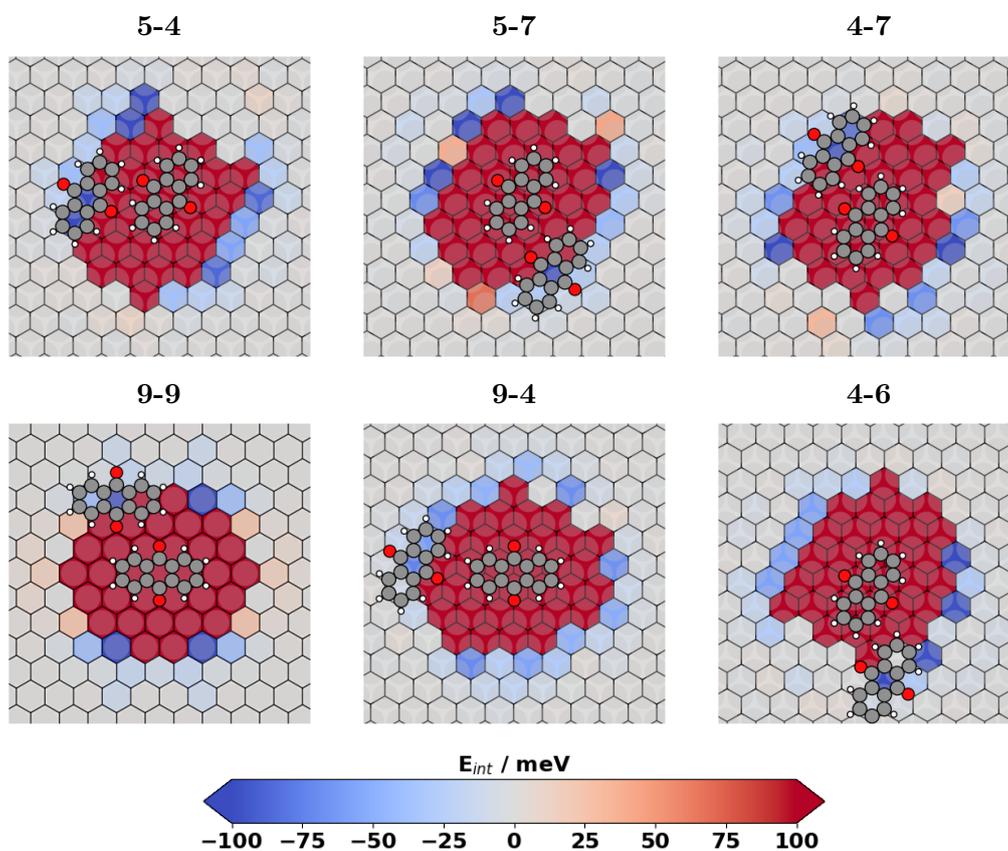
$$\text{area} = 10.66 \text{ N}_{ads}/\text{mol}$$

$$E_{conf} = -156 \text{ meV}/N_{ads}$$

$$\text{area} = 11 \text{ N}_{ads}/\text{mol}$$



**Figure 3.19.:** The two best configurations for AQ on Ag(111) and the fifth best structure, which contains an adsorption geometry that causes a kink to form. The unit cell is indicated in green.



**Figure 3.20.:** Pairwise interaction plots for the relevant pairs of the best configurations for AQ.

## Conclusion

In this thesis, the SAMPLE approach was improved to a degree that enables surface polymorph prediction of large molecules with functional groups. It utilizes coarse graining of the potential energy surface, by discretization of the possible adsorption positions for single molecules to adsorption sites, combined with exhaustive configuration creation and prediction using Bayesian linear regression.

The first step was the development of a mixed-basis set approach for the numerical atomic basis set of the FHI-aims package, that cut the computational cost to an affordable level. As a second step, the method for the generation of local geometries was improved by making use of the torque of a single molecule on the surface for a smart interpolation of the rotational degree of freedom for molecules on the surface. Then the Bayesian learner, which was before only applicable to small, highly symmetric molecules, was generalized by consistent introduction of species-distinguishing methods. One more improvement was achieved by application of knowledge from gas phase calculations to the surface structures.

With this improved method the local geometries for benzoquinone, anthraquinone and pentacenequinone were found. Thereby, a clear connection between the different adsorption geometries came to light, as the number of geometries increased with increasing molecule size. However, most positions for even the largest molecule could be explained by the adsorption behavior of the smaller molecules.

The real structure prediction was then performed with benzoquinone and anthraquinone. For benzoquinone two energetically significantly favorable structures could be found and explained by closer examination of the interaction energies on the surface. Anthraquinone showed a smaller energy spread within the best structures, but the emergence of the manifold of favorable structures could also be illuminated with physical insight.

Those two molecules already provide insight into the influence of molecule size on surface polymorphism. For a deeper understanding the structure search will also be conducted for the larger molecule pentacenequinone and as an intermediate step between benzoquinone and anthraquinone, naphtaquinone will be investigated as well.

# Appendices

## A. Convergence of DFT Settings for an Ag Bulk

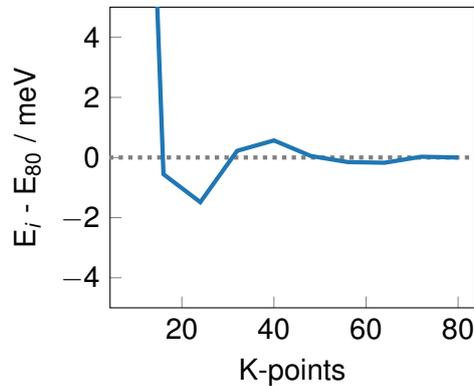
For all following calculations some standard aims settings were used:

Exchange Correlation Functional	PBE
Spin	Restricted
Charge	None
Relativistic Effects	Atomic_zora scalar
Occupation Smearing	Gaussian 0.1 or 0.01
Multipole Error Compensation	True
vdW correction	TS <sup>surf</sup>
RI method	lvl_fast

For all slab calculations dipole correction was used.

### Convergence of the Ag Bulk

The first step was to get a correct lattice constant for the silver bulk. Therefore first a primitive unit cell for a silver fcc bulk with the experimental lattice constant was constructed. Then a sweep of k-points from 8 to 80 in steps of 8 was performed in each direction. VdW corrections with Tkatchenko-Scheffler surface parametrization were used also for the bulk calculations to ensure compatibility with the surface calculations later on.



**Figure A1.:** Convergence behavior of the total energy of an Ag-bulk with respect to the number of k-points in each direction.

To find the equilibrium lattice constant  $a$  of the silver bulk, it had to be chosen such that the cohesive energy  $E_c$  is minimized.

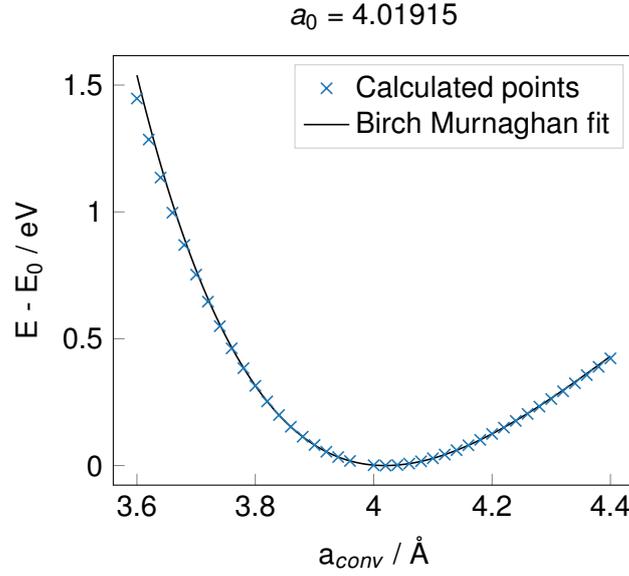
$$E_c(a) = \left( \frac{E(a)}{N} - E_{isol} \right)$$

- E(a) ... total energy of the atoms in a unit cell
- N ..... number of atoms in the cell
- $E_{isol}$  ... total energy of an isolated atom

Therefore single point(SP) calculations for a fcc-bulk with a lattice constant varying from 3.6 Å to 4.4 Å with 64 k-points (clearly converged as visualized in figure A1) were performed. A Birch Murnaghan fit was then used to determine the equilibrium lattice constant. The Birch-Murnaghan equation of state [56] is a relationship between the volume of a body and the pressure to which it is subjected. The internal energy of the system is given by

$$E(V) = E_0 + \frac{9V_0B_0}{16} \left\{ \left[ \left( \frac{V_0}{V} \right)^{\frac{2}{3}} - 1 \right]^3 B'_0 + \left[ \left( \frac{V_0}{V} \right)^{\frac{2}{3}} - 1 \right]^2 \left[ 6 - 4 \left( \frac{V_0}{V} \right)^{\frac{2}{3}} \right] \right\}$$

For mono-atomic unit cell bases, the minimization of the internal energy is equivalent to the minimization of the cohesive energy. The results are shown in figure A2.

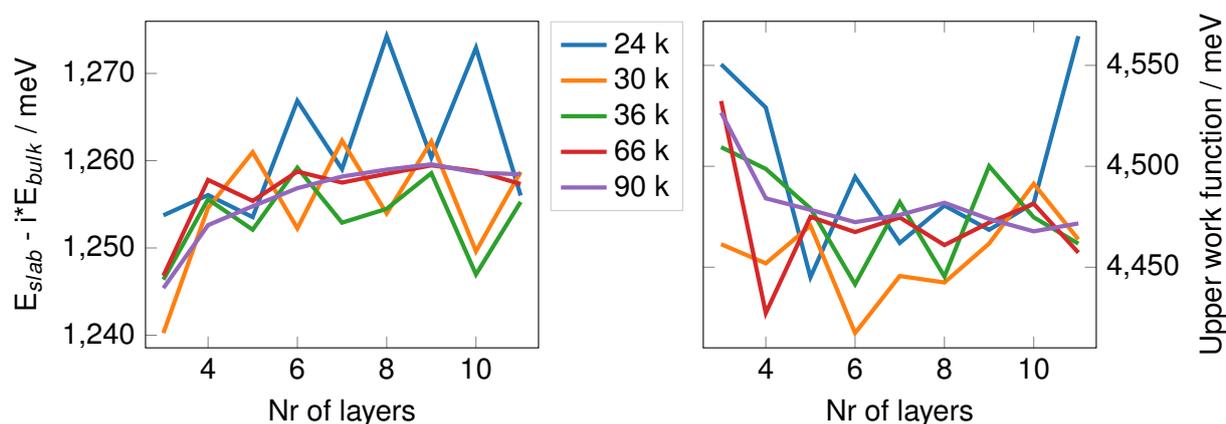


**Figure A2.:** Evolution of total energy of a fcc Ag-bulk with variation of the lattice constant.

With this method the lattice constant of the conventional unit cell was determined to  $a_0 = 4.01915$  Å , which is close to the experimental value of 4.079 Å [57].

### Convergence of the Ag(111) Slab

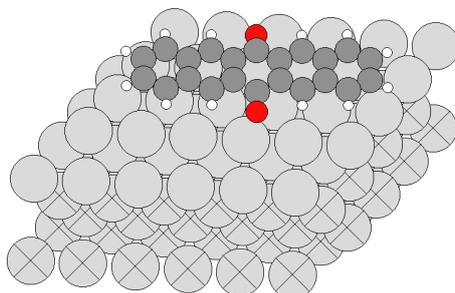
Another value to consider was the surface energy, which was obtained by taking the total energy of the slab minus the bulk energy times number of layers. Here convergence could be enforced by taking many k-points and a medium amount of layers (5-6). Convergence of the work function is a harder task, here many k-points have to be taken and the work function is still not sufficiently converged.



**Figure A3.:** Evolution of the surface energy and work function of an Ag(111) slab in dependence of the number of layers and k-points.

## B. Convergence Tests for the Adsorption Energy of Pentacenequinone

As for the work in this thesis only proper convergence of the adsorption energy is necessary, further tests of several parameters were conducted. For all parameter tests a preoptimized geometry from Elisabeth Wruß was taken (see figure A4). It consists of a single pentacenequinone molecule on a silver slab with 5 layers and 4 by 6 Ag atoms per layer in the unit cell.



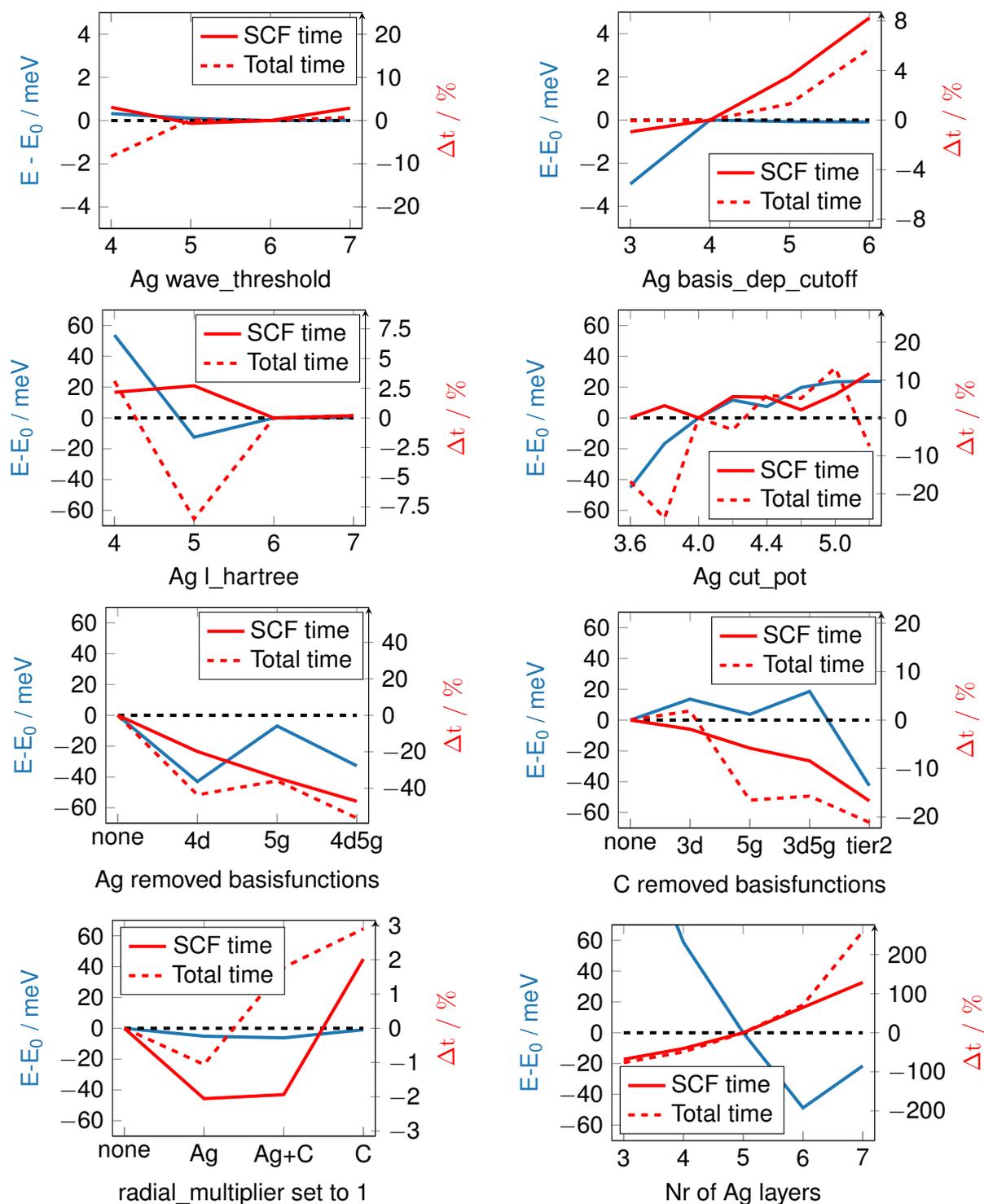
**Figure A4.:** Geometry used for parameter convergence tests.

The varied parameters and their effects on adsorption energy and computation time are shown in figure A5. All values are differences of adsorption energies between different settings. For all calculations the affected parts were recalculated with the respective parameters (e.g. for radial\_multiplier the full system and Ag slab were recalculated whereas the molecule was not).

Based on these results, the following settings were chosen:

species	setting	value	reason	error / meV	save per SCF / %
Ag	cut_pot	5.0	convergence	<b>5</b>	-8
Ag	radial_multiplier	1	runtime	<b>5</b>	2
Ag	basisfunctions	tier1 -5g	runtime	<b>6</b>	<b>34</b>
Ag	wave_threshold	1E-5	runtime	0.1	0.6
Ag	l_hartree	6	convergence	0	0
Ag	basis_dep_cutoff	1E-4	convergence / runtime	0	0
C	basisfunctions	tier2 -5g	runtime	3.7	5.8

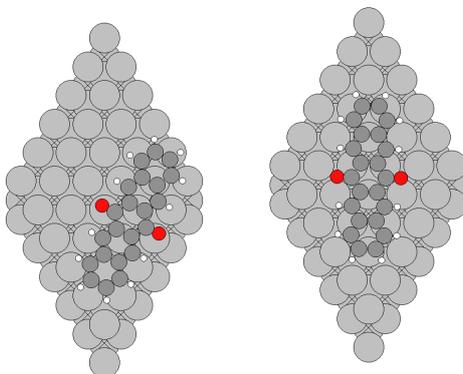
Another calculation with all chosen settings gives a total error of **16 meV** and a computation time saving of **40 %**.



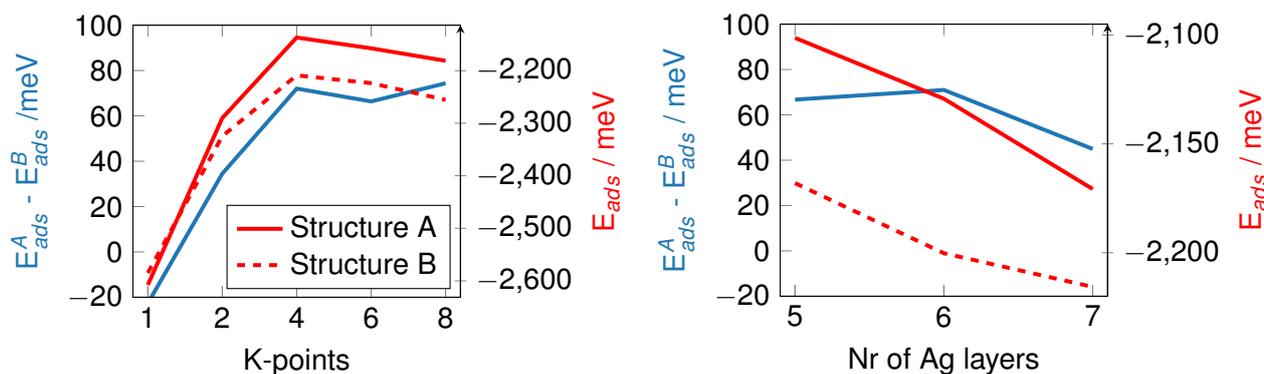
**Figure A5.:** Convergence calculations for several parameters of the tight Ag and tight C species.  $E_0$  is always the reference of the standard tight basis set. The basis functions were removed from the highest tier (tier1 for Ag and tier2 for C). For all calculations only one parameter was swept and the other were kept at the default settings. The differences between total time and SCF time arise from the different numbers of SCF cycles needed.

### Convergence of Differences of Adsorption energies

Due to the horrible layer convergence, a second adsorption geometry was created and the convergence with respect to the difference of adsorption energies between the two systems was checked.



**Figure A6.:** Previously used configuration (left) and second adsorption geometry (right) on a 6x6 Ag(111) slab.



**Figure A7.:** Dependence of the adsorption energies (red) and difference of adsorption energies (blue) on the number of k-points (left) and number of Ag layers (right).

First a check of k-points was performed and then with a nearly converged value of 4 k-points in x and y direction the layer convergence was carried out (see figure A7). The layer convergence exhibits a strong dependence of the adsorption energy on the number of layers, which was just the opposite of what was expected to happen.

A solution to this behavior was then found with the mixed-basis set approach. The results for the layer convergence are shown in figure 2.5.

## C. Computational Details of the Mixed Basis Set Method

### Trial Basis sets Used

**Table A.1.:** Test species for the slab basis-set modification

species	name of the designed basis set					
l_hartree	maximum angular momentum of charge-density expansion					
cut_pot	outer cutoff radius of numeric species definition					
rad_mult	multiplicity for the number of integration points given by radial_multiplier					
last grid	outermost radial shell with explicit integration grid density					
outer grid	integration grid density from last grid to cut_pot					
basis functions	basis functions additional to the minimal basis					
species	l_hartree	cut_pot	rad_mult	last grid	outer grid	basis functions
<i>tight</i> Ag	6	4.0	2	1.59	434	5p; 4f; 3s; 5g; 4d
<i>5p</i> Ag	6	4.0	1	1.59	434	5p
<i>light</i> Ag	4	3.5	1	1.56	302	5p; 4f; 3s; 4d
<i>reallylight</i> Ag	4	3.5	1	1.27	194	5p

The basis set *tight* Ag is the standard tight basis set used in aims, from there three different bases were created. *5p* Ag stems from the *tight* Ag basis except the 5p orbitals were removed. The *light* Ag consists of all tight-Ag basis functions except the 5g functions, but the integration grid was reduced significantly. The *reallylight* Ag has as little basis functions as *5p* Ag and an even worse integration grid than *light* Ag. The *tight* basis set was also modified to a cutoff potential of 5 Å, due to convergence tests.

### Convergence issues

When trying to calculate long slabs with different basis sets, convergence issues were encountered, which manifested with a large amount of SCF cycles. To overcome this problem, a check of the manual revealed the following recommendations for metal slabs: The FHI-aims manual suggests that the mixing parameter for the Pulay mixer should have a value of 0.05 or lower. For the preconditioner Kerker can be used but with a value between 1 and 2.

A test of the number of SCFs for a slab with 6 *tight* Ag and 4 *reallylight* Ag atoms confirmed the manual suggestions by reducing the cycles to around 20 (see table A.2).

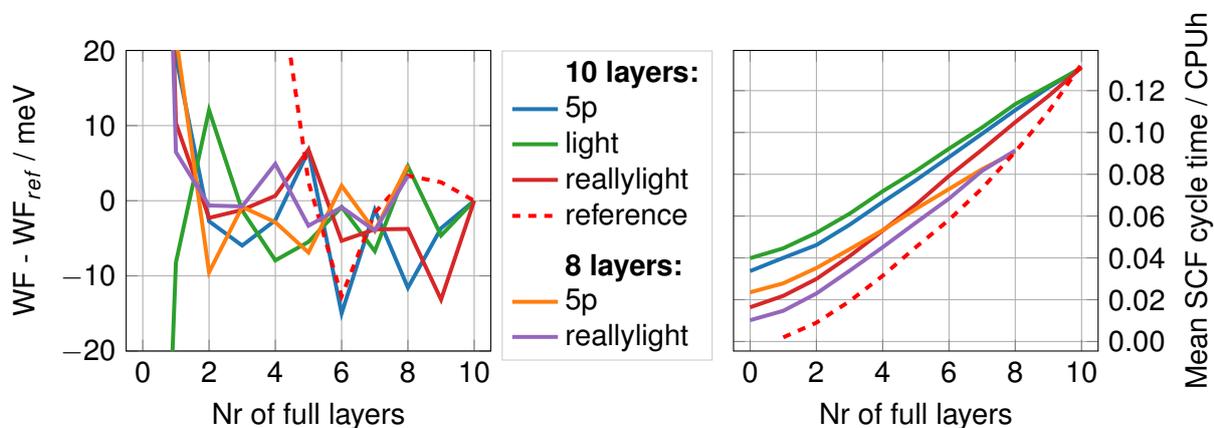
**Table A.2.:** Convergence behavior of an Ag slab with 10 atoms. Gray atoms are *tight* Ag, red atoms *reallylight* Ag.

mixer specifications	Nr SCFs
None	329 → convergence error
pulay 0.05	27
kerker 1.5	38
mixer + pulay	20



## Convergence Behavior of the Work Function

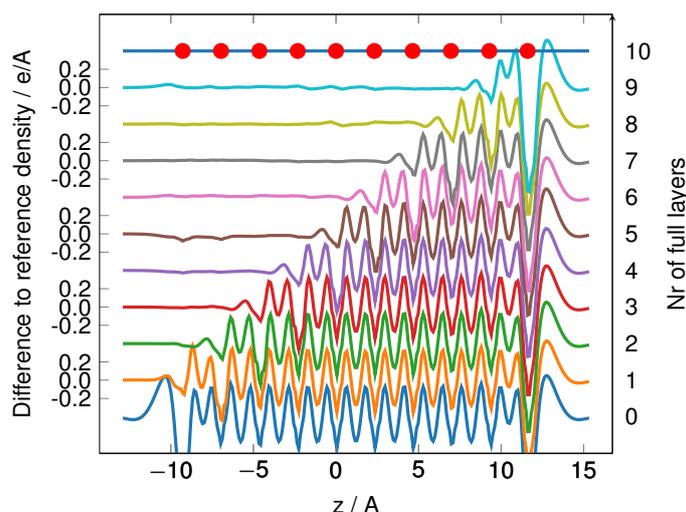
Here the convergence behavior of the work function for the three basis sets introduced above are shown. A single-pillar slab of *tight* Ag with 8 and 10 atoms was created and the species was replaced with the corresponding lighter settings atom by atom.



**Figure A8.:** Work function evolution and mean SCF-cycle time of an Ag pillar of 8 / 10 atoms when gradually replacing the basis set.

Checking the work function (figure A8) shows, that for a slab with 10 but also 8 atoms a suitable convergence of the work function for all basis sets can already be achieved when only 2 of the atoms are calculated with the *tight* Ag species, whereas 7 layers of silver are needed for the same convergence taking only *tight* Ag. The computational offset of this method is comparable to 2-4 layers of *tight* silver, because there are always 8-10 atoms present, which is easily outweighed by the better convergence behavior.

The differences of the electron densities projected onto the x-axis, when replacing the *reallylight* basis with *tight* silver, are shown in figure A9. One can see that the main differences occur where different basis functions are used. The fluctuations of electron density in the desired region are very small even for only 3 layers of *tight* silver.

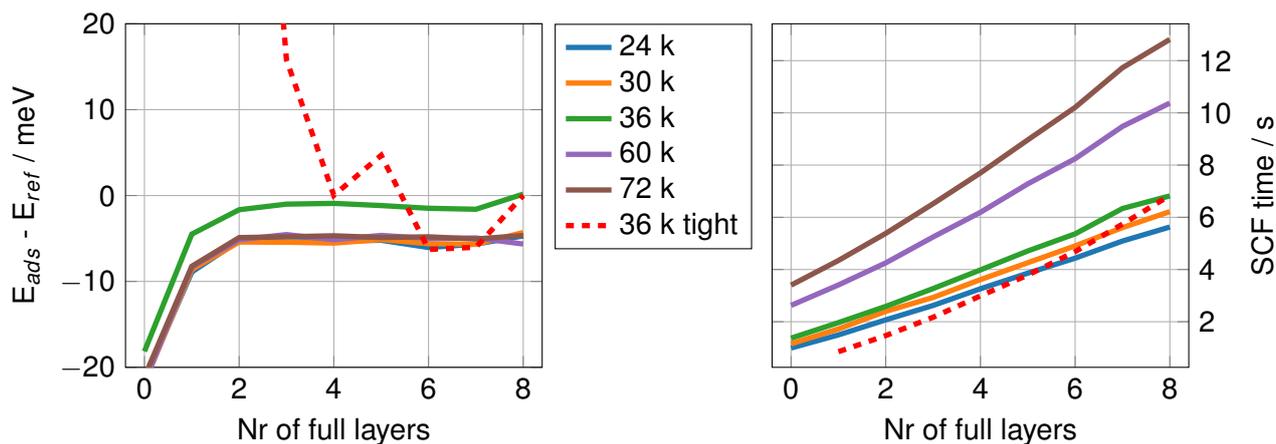


**Figure A9.:** Difference of electron densities to the reference density for substitution with *reallylight* Ag species. The lowermost line represents a slab with *reallylight* species only. *Tight* Ag were added from the bottom (left in the plot).

## CO Adsorption

Here the adsorption energy of carbon monoxide on a silver slab with 8 atoms and subsequently altered species from *tight* to *reallylight* is shown. In figure A10 the computational cost and the energy convergence of an 8 layer slab with consecutive replaced atoms for different k-point densities is depicted.

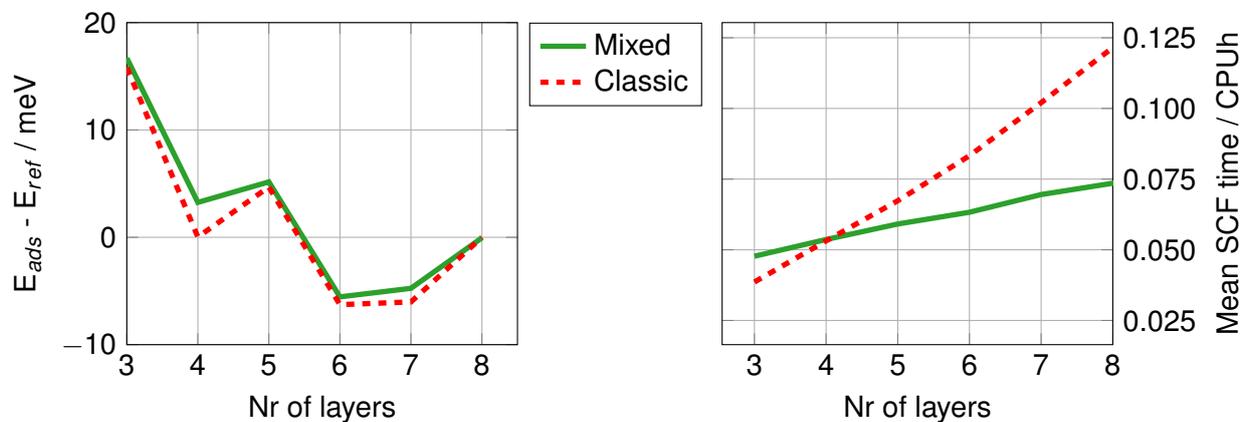
With the mixed-basis set slab the adsorption energy is again converged within 3 layers of *tight* Ag, compared to 6-8 layers for solely *tight* Ag. The reason for different convergence energy for 36 k-points might stem from a second very similar solution of the SCF-cycle that is reached when specific k-points are calculated. As this does not affect the future calculations no further investigations were made.



**Figure A10.:** Convergence behavior of CO adsorption energy and time per SCF cycle with different k point densities.

## Number of Layers needed

Before the layer convergence tests for pentacenequinone were performed, they were conducted on CO. Therefore a slab with 3 layers of *tight* Ag was taken and layers of *tight* Ag or *reallylight* Ag were added. The results in figure A11 show, that the layer convergence is very similar, but the computational effort is much lower. As the computational time scales with the number of basis functions in the system, the savings should become even larger. Those calculations were performed with 36 k-points and an influence of the k-point density on the convergence behavior can not be ruled out, which is why the convergence of pentacenequinone was initially performed with 60 k-points.



**Figure A11.:** Dependence of of the adsorption energy on the total number of layers for the mixed basis and the standard tight-only Ag.

## Final Settings

The structure search calculations performed in this thesis worked with a k-point density of 36 per substrate atom and a slab with 3 layers of *tight* Ag and 5 layers of *reallylight* Ag. The species definitions are tabulated in listing 1 and 2.

The settings for the control file used in the convergence tests and also later are:

### General Settings:

```
xc                pbe
spin              none
charge            0
relativistic      atomic_zora scalar
occupation_type   gaussian 0.1 / 0.01
k_grid            36 36 1
use_dipole_correction .true.
compensate_multipole_errors .true.
```

### Accuracy:

```
sc_accuracy_rho   1E-2
sc_accuracy_etot  1E-5
```

### Charge mixer:

```
mixer              pulay
charge_mix_param   0.05
preconditioner     kerker 1.5
```

**Listing 1:** Settings for *reallylight* Ag

```
species          Ag_reallylight
# global species definitions
  hirshfeld_param 122 15.4 2.57 #vdW
  surf
  nucleus          47
  mass             107.8682
  l_hartree        4
  cut_pot          3.5 1.5 1.0
  basis_dep_cutoff 1e-4
# integration grid
  radial_base      62 5.0
  radial_multiplier 1
  angular_grids specified
  division         0.5617 50
  division         0.9788 110
  division         1.2700 194
  outer_grid      194
# valence basis states
  valence         5 s 1.
  valence         4 p 6.
  valence         4 d 10.
# ion occupancy
  ion_occ         5 s 0.
  ion_occ         4 p 6.
  ion_occ         4 d 9.
# Additional basis functions
  ionic 5 p auto
```

**Listing 2:** Settings for *tight* Ag

```
species          Ag
# global species definitions
  hirshfeld_param 122 15.4 2.57 #vdW
  surf
  nucleus          47
  mass             107.8682
  l_hartree        6
  cut_pot          5 2.0 1.0
  basis_dep_cutoff 1e-4
# integration grid
  radial_base      62 7.0
  radial_multiplier 1
  angular_grids specified
  division         0.3947 50
  division         0.7739 110
  division         1.1156 194
  division         1.3117 302
  division         1.5936 434
  outer_grid      434
# valence basis states
  valence         5 s 1.
  valence         4 p 6.
  valence         4 d 10.
# ion occupancy
  ion_occ         5 s 0.
  ion_occ         4 p 6.
  ion_occ         4 d 9.
# Additional basis functions
  ionic 5 p auto
  hydro 4 f 7.6
  hydro 3 s 2.6
  hydro 4 d 8.4
```

## D. Electrostatic Prior

For systems with much charge transfer and reasonable distance this is assumed to be a major part of the interaction energy. The first assumption is often given, but the second, namely reasonable distance between molecules, is only rarely fulfilled at the search for dense packed structures. Nevertheless, in the following the interaction energy dependent on the interacting charges and potentials will be derived.

The energy stored inside a charge arrangement is equal to the energy needed to generate the arrangement. The potential of two point charges at a specific distance is:

$$U = \int_{\infty}^{r_{12}} \frac{1}{4\pi\epsilon_0} \frac{q_1 q_2}{|r_1 - r_2|^2} dr_1 = \frac{1}{4\pi\epsilon_0} \frac{q_1 q_2}{r_{12}} \quad (\text{A1})$$

And because of the superposition principle, the change in potential energy when adding a single charge into an arrangement of  $q_i$  is:

$$\Delta U = \frac{1}{4\pi\epsilon_0} \sum_i \frac{q_1 q_i}{r_{1i}} = \frac{1}{2} \int q_1 \frac{1}{4\pi\epsilon_0} \sum_i \frac{q_i}{r_{1i}} dr = \frac{1}{2} \int q_1 V(r) dr \quad (\text{A2})$$

which leads to the total potential energy of a point charge distribution:

$$U = \frac{1}{2} \frac{1}{4\pi\epsilon_0} \sum_{i,j,i \neq j} \frac{q_i q_j}{|r_i - r_j|} \quad (\text{A3})$$

where the 1/2 comes from double counting and the self energy of charges is excluded from the sum.

Generalizing the last part of equation A2 to a charge distribution  $\rho(r)$  leads to the potential energy of an electron density inside a potential:

$$U = \frac{1}{2} \int \rho(r) V(r) dr \quad (\text{A4})$$

Next, the interacting charge parts need to be extracted. The electron density of a single local adsorption geometry on the substrate can be written as:

$$\rho_{tot}^A = \rho^0 + \rho^A + \delta^A$$

Where  $\rho^0$  and  $\rho^A$  are the undisturbed densities of substrate and molecule A.  $\delta^A$  is the difference to the full system which represents the charge rearrangement due to interactions.

The total electrostatic energy of this configuration is given by the interaction of all parts with themselves and with each other:

$$E_{tot}^A = (\rho^0, \rho^0) + (\delta^A, \delta^A) + (\rho^A, \rho^A) + (\rho^0, \delta^A) + (\rho^0, \rho^A) + (\delta^A, \rho^A)$$

The interaction terms  $(\rho^i, \rho^j)$  can either be calculated only via charge densities, leading to a 6-dimensional integral, or alternatively by considering the potential generated by one of the distributions, which reduces the mathematical effort to a 3-dimensional integration:

$$\begin{aligned} (\rho^A(r), \rho^B(r')) &= \frac{1}{2} \frac{1}{4\pi\epsilon_0} \int dr dr' \frac{\rho^A(r) \rho^B(r')}{|r - r'|} \\ &= \frac{1}{2} \int dr \rho^A(r) V^B(r) \end{aligned}$$

The electrostatic contribution to the adsorption energy is:

$$\begin{aligned} E_{ads}^A &= E_{tot}^A - E_{slab} - E_{mol}^A \\ &= (\delta^A, \delta^A) + (\delta^A, \rho^0) + (\delta^A, \rho^A) + (\rho^0, \rho^A) \end{aligned}$$

with  $E_{slab} = (\rho^0, \rho^0)$  and  $E_{mol} = (\rho^A, \rho^A)$ .

When two molecules are adsorbed on a common substrate, the total charge density is given by:

$$\rho_{tot}^{A+B} = \rho^0 + \rho^A + \rho^B + \delta^A + \delta^B$$

The total electrostatic interaction energy between the two molecules, neglecting electronic interactions leading to charge rearrangements, can then be evaluated by first calculating the total electrostatic energy of the whole system

$$\begin{aligned} E_{full}^{A+B} &= (\rho^0, \rho^0) + (\rho^A, \rho^A) + (\rho^B, \rho^B) + (\delta^A, \delta^A) + (\delta^B, \delta^B) \\ &+ (\rho^0, \delta^A) + (\rho^0, \delta^B) + (\rho^0, \rho^A) + (\rho^0, \rho^B) + (\rho^A, \delta^A) + (\rho^B, \delta^B) \\ &+ (\delta^A, \delta^B) + (\rho^A, \rho^B) + (\delta^A, \rho^B) + (\delta^B, \rho^A) \end{aligned}$$

And then subtracting the self-energies of the molecules and the substrate, and the adsorption energies of the single molecules

$$E_{int}^{A+B} = E_{full}^{A+B} - E_{slab} - E_{mol}^A - E_{mol}^B - E_{ads}^A - E_{ads}^B$$

The different colors represent the corresponding energy terms.

Now interaction energy between two adsorbed molecules becomes

$$E_{int}^{A+B} = (\delta^A, \delta^B) + (\delta^A, \rho^B) + (\rho^A, \rho^B) + (\rho^A, \delta^B)$$

Due to the additivity of the charge distributions and the electrostatic potentials, the total charge of the two non interacting systems with one molecule each can be reintroduced:

$$\begin{aligned} &= (\delta^A, \delta^B + \rho^B) + (\rho^A, \delta^B + \rho^B) \\ &= (\delta^A + \rho^A, \delta^B + \rho^B) \\ &= (\rho_{tot}^A - \rho^0, \rho_{tot}^B - \rho^0) \end{aligned}$$

Or in another form, which is more stable with respect to numerical treatment

$$E_{int}^{A+B} = (\rho^A, \rho_{tot}^B - \rho^0) + (\delta^A, \rho_{tot}^B - \rho^0) \quad (A5)$$

This last form is beneficial, because singularities close to the cores of the atoms can be treated. For molecule B, the total electrostatic potential minus the substrate potential can be used, as the singularities of the potential are far from the electron density of Molecule A. The charge redistribution  $\delta^A$  shows no singularities and can also be treated nicely. The molecular charge density can now either be treated fully, or suitable charge partitioning schemes can be used.

## Numerical Treatment

**Charge partitioning with Electrostatic potential (ESP) charges:** ESP charges are fitted from the electrostatic potential on a cube grid such, that they reproduce the potential at a certain distance only with point charges at the center of the atoms. In principle this method works for non periodic and periodic systems, but within FHI-aims for periodic calculations, all atoms have to be in the first unit cell. A representative control.in command to get ESP charges is:

```
output esp
esp n_radius 10      # number of shells where potential is calculated
esp radius 1.0 2.    # points within once and twice the vdW radius
esp pbc_method 1     # 1: Ewald summation, 2: Wolf summation
esp R_c 10           # Cutoff Radius for pbc
```

**Using Potential and electron density:** For the parts in equation A5, where the full density or potential is needed, cube files were used. Cube files represent the continuous density and potential in a discretized three-dimensional grid. When evaluating equation A4, this results in a simple summation over all points of two cube files.

FHI-aims commands to get the potential and electron density are:

```
output cube total_density      # get whole electron density on a cube grid
output cube hartree_potential  # short and long-range potential
cube origin 0.0 0.0 0.0       # center of region to be plotted in A
cube edge n dx dy dz          # specification of cube edge and voxel partition
cube edge 200 0.1 0 0         # e.g give an edge with 20 A in x direction
cube edge 200 0 0.1 0         # all three combined specify the cube volume
cube edge 200 0 0 0.1
```

Alternatively the values can also be written out at the integration grid points:

```
output density # electron density
output grid    # integration grid for each species
output v_eff   # local effective potential
output v_hartree # writes electrostatic multipole components separate for atom, l and m
```

Attention: Cube outputs of the geometry and lattice are given in Bohr while the density is given per cubic Angstrom.

The electric permittivity is usually given in Farad per meters, but the energy unit used in this thesis is electron volts, therefore the following conversion is used:

$$1 \text{ C} = \frac{1}{1.602176565 * 10^{-19}} e, 1 \text{ m} = 10^{10} \text{ A}$$

$$\epsilon_0 = 8.8541878176 * 10^{-12} \text{ F/m} = 5.526349599 * 10^{-3} e/AV$$

## E. Minimal Distances

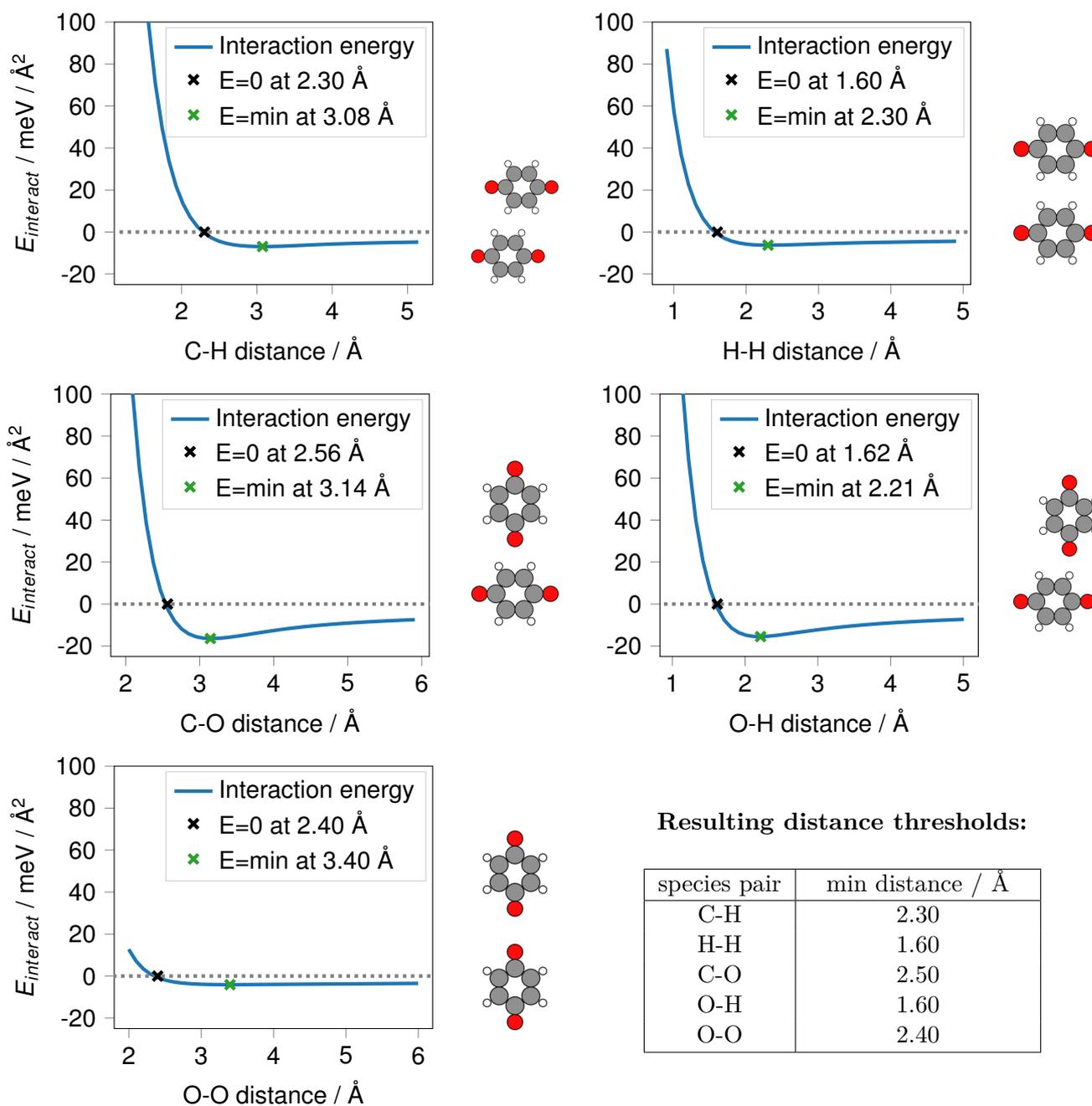


Figure A12.: Interaction energy evolution of five distinct dimers with increasing molecule distance.

## F. Interpolation of the Rotational PES

In the following, the mathematical derivation of the spline interpolation for the rotational dependence of the adsorption energy for single molecules on a surface is given. The first derivative of the energy with respect to the rotation angle is determined via the torque

$$\frac{dE}{d\theta} = \tau = \sum_{i=1}^{N_{atoms}} \mathbf{r}_i * \mathbf{f}_i \quad (\text{A6})$$

As a boundary condition the splines and their first two derivatives need to be continuous at all data points

$$\begin{aligned} y_i(x) &\stackrel{!}{=} y_{i+1}(x) = E_j \\ y'_i(x) &\stackrel{!}{=} y'_{i+1}(x) = \tau_j \\ y''_i(x) &\stackrel{!}{=} y''_{i+1}(x) \end{aligned}$$

Taking now a third degree polynomial and its derivatives of the form

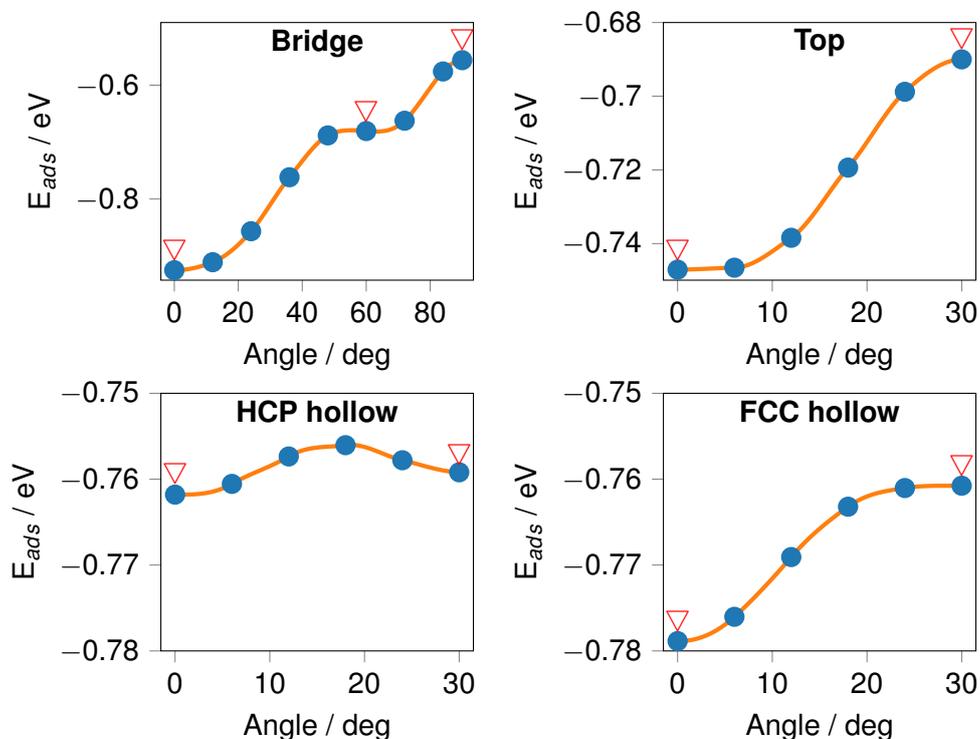
$$\begin{aligned} y_i(x) &= a_i + b_i * x + c_i * x^2 + d_i * x^3 \\ y'_i(x) &= b_i + 2 * c_i * x + 3 * d_i * x^2 \\ y''_i(x) &= 2 * c_i + 6 * d_i * x \end{aligned}$$

At a single support point, the adjoined polynomials and their derivatives need to be continuous, which leads to the following polynomial coefficients

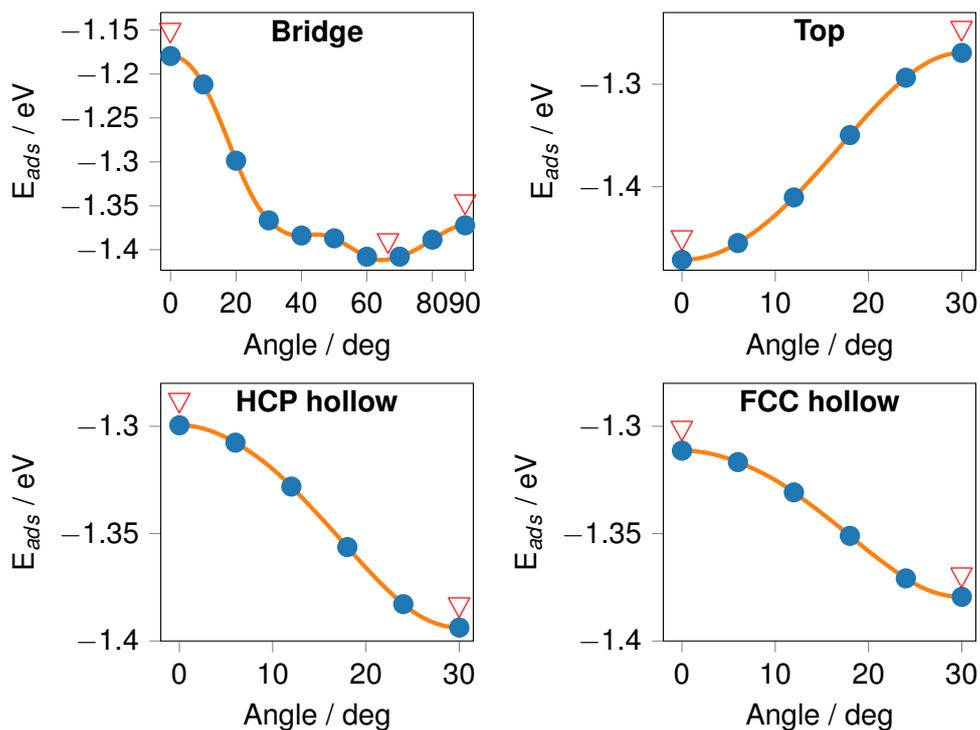
$$\begin{aligned} a_i &= E_j \\ b_i &= \tau_j \\ c_i &= \frac{3 * (E_{j+1} - E_j) - 2 * \tau_j * \Delta x - \tau_{j+1} * \Delta x - \Delta x * \tau_j}{\Delta x^2} \\ d_i &= \frac{\tau_j * \Delta x + 2 * E_j - 2 * E_{j+1} + \tau_{j+1} * \Delta x}{\Delta x^3} \end{aligned}$$

## G. Local Geometry Rotation Sweeps

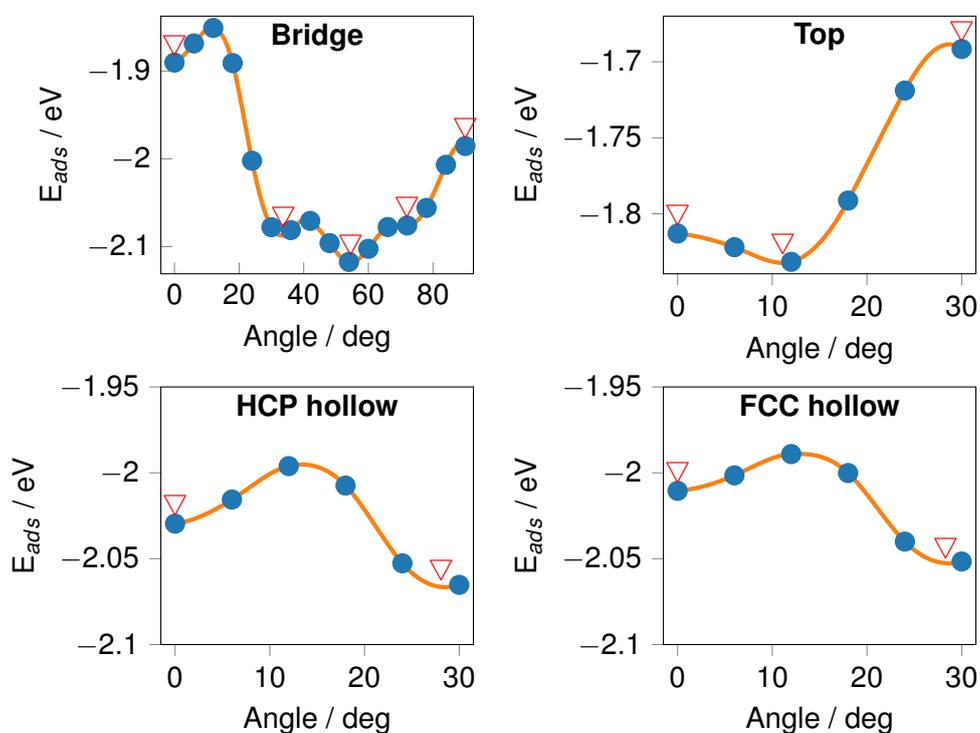
Here the rotation sweeps for the three molecules are shown. Due to the molecular symmetry along the first and second main axis, all points after 30 degree are symmetry equivalent for top and the two hcp. For bridge the symmetry is after 90 degrees. If the molecule had no symmetries, the symmetry should be after 180 degrees.



**Figure A13.:** Adsorption energy of BQ rotated around 4 high symmetry points. The points indicate DFT calculations. The red arrows are starting geometries for a subsequent geometry optimization.



**Figure A14.:** Adsorption energy of AQ rotated around 4 high symmetry points. The points indicate DFT calculations. Red arrows indicate starting points for geometry optimization.



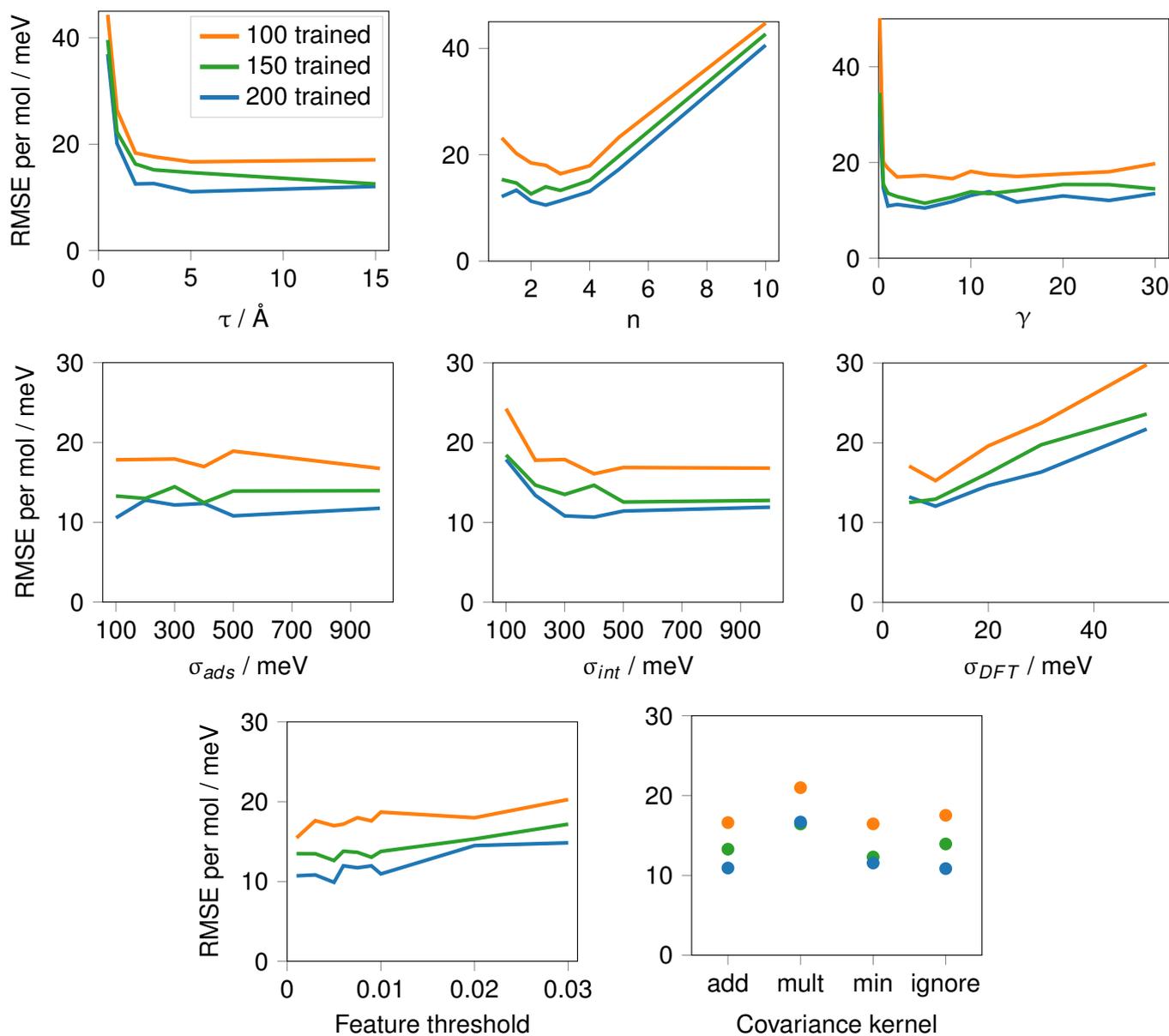
**Figure A15.:** Adsorption energy of PQ rotated around 4 high symmetry points. The points indicate DFT calculations. Red arrows indicate starting points for geometry optimization.

## H. Hyper Parameter Tests

Here the full hyper parameter sweep for benzoquinone and anthraquinone is shown. First the hyper parameters were preoptimized by changing single parameters and then performing the sweep again. For every sweep all other parameters were kept at the reference value shown in table A.3.

**Table A.3.:** Reference parameters used for optimization

system	$\sigma_{ads}$ [meV]	$\sigma_{int}$ [meV]	$\sigma_{DFT}$ [meV]	$\tau$ [Å]	n	$\gamma$	feat. thres.	cov. kernel
BQ	100	300	10	10	3	12	0.010	additive
AQ	100	300	10	10	3	12	0.010	additive



**Figure A16.:** Hyper parameter tests for a gas phase monolayer of benzoquinone.

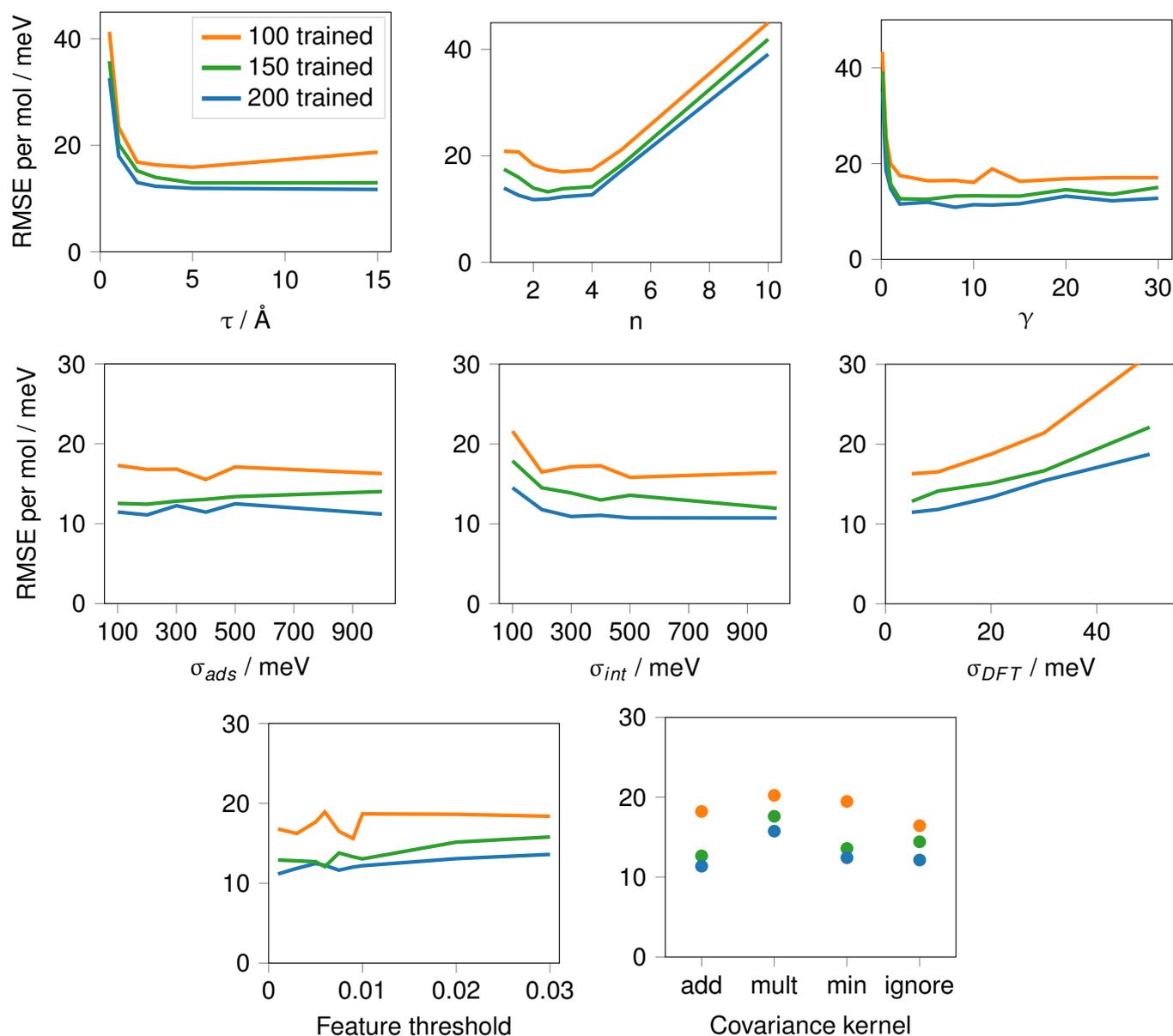
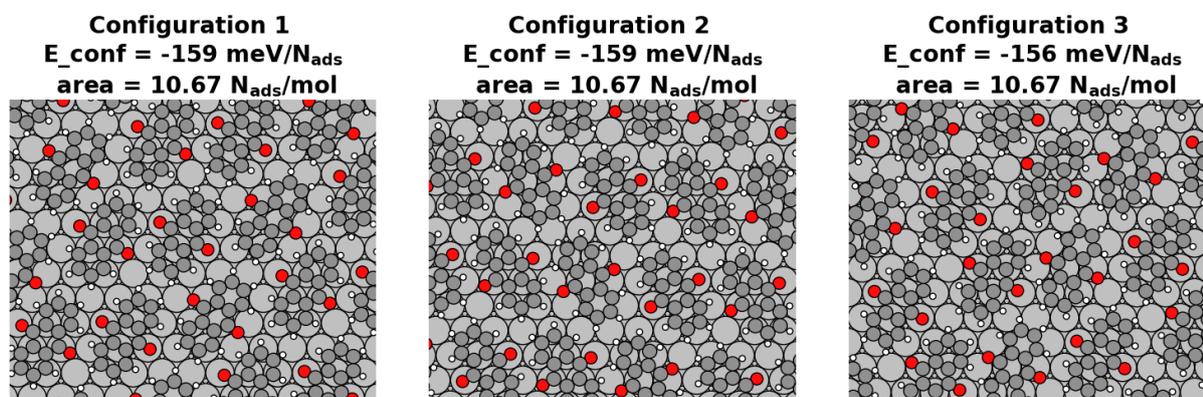
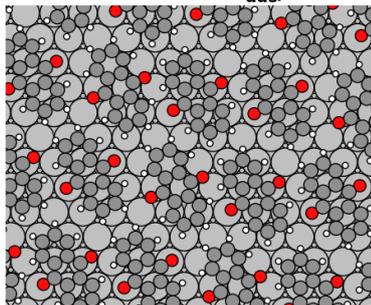


Figure A17.: Hyper parameter tests for a gas phase monolayer of anthraquinone.

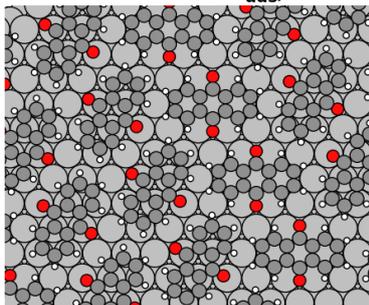
## I. Best Structures of Anthraquinone



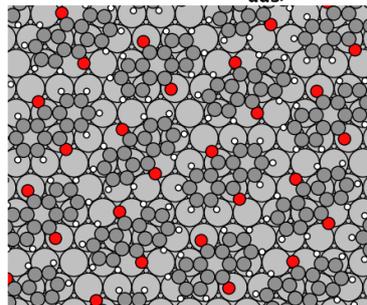
**Configuration 4**  
 $E_{\text{conf}} = -156 \text{ meV}/N_{\text{ads}}$   
 $\text{area} = 10.67 \text{ N}_{\text{ads}}/\text{mol}$



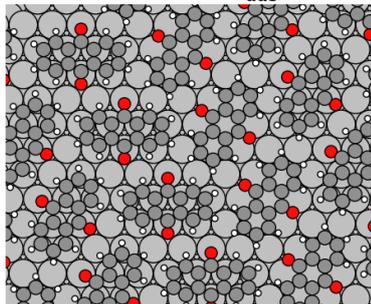
**Configuration 5**  
 $E_{\text{conf}} = -156 \text{ meV}/N_{\text{ads}}$   
 $\text{area} = 11.00 \text{ N}_{\text{ads}}/\text{mol}$



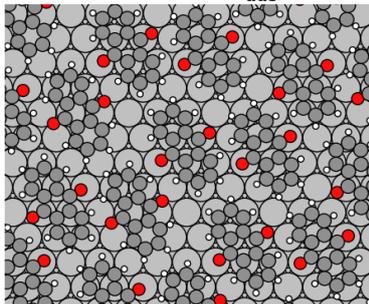
**Configuration 6**  
 $E_{\text{conf}} = -154 \text{ meV}/N_{\text{ads}}$   
 $\text{area} = 11.00 \text{ N}_{\text{ads}}/\text{mol}$



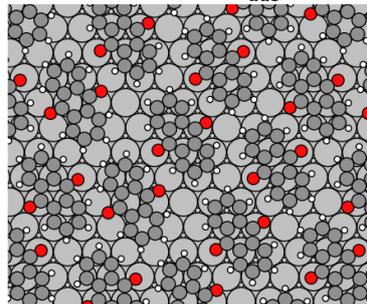
**Configuration 7**  
 $E_{\text{conf}} = -154 \text{ meV}/N_{\text{ads}}$   
 $\text{area} = 11.00 \text{ N}_{\text{ads}}/\text{mol}$



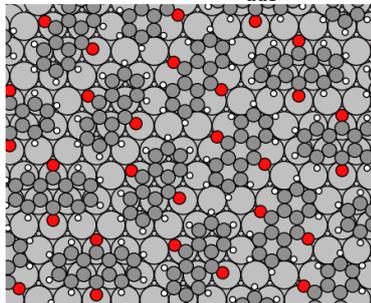
**Configuration 8**  
 $E_{\text{conf}} = -154 \text{ meV}/N_{\text{ads}}$   
 $\text{area} = 11.00 \text{ N}_{\text{ads}}/\text{mol}$



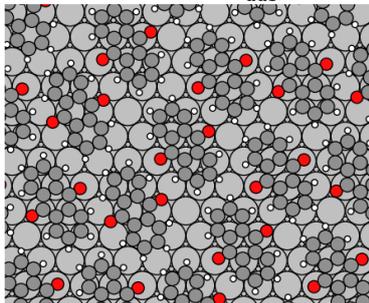
**Configuration 9**  
 $E_{\text{conf}} = -154 \text{ meV}/N_{\text{ads}}$   
 $\text{area} = 11.00 \text{ N}_{\text{ads}}/\text{mol}$



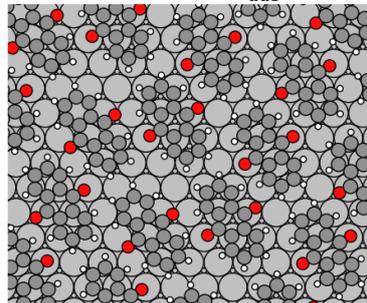
**Configuration 10**  
 $E_{\text{conf}} = -153 \text{ meV}/N_{\text{ads}}$   
 $\text{area} = 11.00 \text{ N}_{\text{ads}}/\text{mol}$



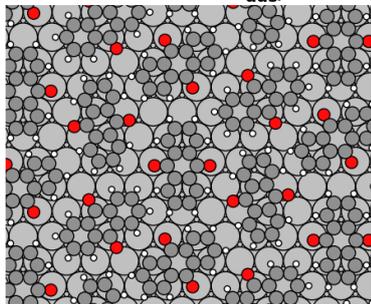
**Configuration 11**  
 $E_{\text{conf}} = -153 \text{ meV}/N_{\text{ads}}$   
 $\text{area} = 11.00 \text{ N}_{\text{ads}}/\text{mol}$



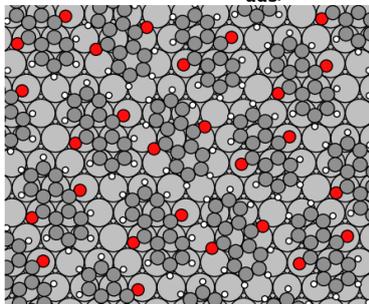
**Configuration 12**  
 $E_{\text{conf}} = -153 \text{ meV}/N_{\text{ads}}$   
 $\text{area} = 11.00 \text{ N}_{\text{ads}}/\text{mol}$



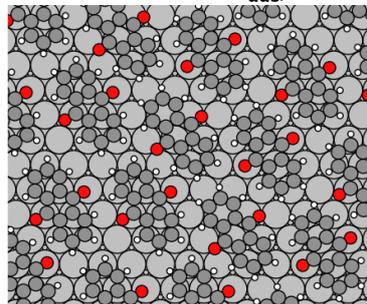
**Configuration 13**  
 $E_{\text{conf}} = -153 \text{ meV}/N_{\text{ads}}$   
 $\text{area} = 11.00 \text{ N}_{\text{ads}}/\text{mol}$



**Configuration 14**  
 $E_{\text{conf}} = -153 \text{ meV}/N_{\text{ads}}$   
 $\text{area} = 11.00 \text{ N}_{\text{ads}}/\text{mol}$



**Configuration 15**  
 $E_{\text{conf}} = -153 \text{ meV}/N_{\text{ads}}$   
 $\text{area} = 11.00 \text{ N}_{\text{ads}}/\text{mol}$



## List of Figures

1.1.	Deviation from straight line condition shown for DFT calculations of a single Cl atom with fractional charges and different XC functionals. . . . .	12
1.2.	fcc lattice in real space and corresponding bravais lattice in reciprocal space . . . . .	13
1.3.	Illustration of the repeated slab approach . . . . .	14
1.4.	Explanation of the different energies used when explaining the interactions of molecules on surfaces . . . . .	15
1.5.	Schematic representation of the first three layers of pentacene on boron nitride . . . . .	16
1.6.	Creation of a configuration with 2 molecules in a 4x4 unit cell . . . . .	20
1.7.	Visualization of the interactions used for the energy model . . . . .	21
1.8.	Visualization of all distances between hydrogen atoms of two benzene molecules . . . . .	22
1.9.	Schematic workflow of the sample method from the calculation of local geometries to the prediction of surface polymorphs . . . . .	24
2.1.	Accuracy requirements for the metal slab layers. . . . .	25
2.2.	Work function evolution and mean SCF-cycle time of an Ag pillar of 10 atoms when gradually replacing the basis set. . . . .	26
2.3.	Evolution of surface relaxation with increasing number of <i>tight</i> Ag layers starting from an 8-layer slab with <i>reallylight</i> Ag . . . . .	27
2.4.	Adsorption energy accuracy dependent on the calculation time for a layer convergence with <i>reallylight</i> Ag and 36 k-points . . . . .	27
2.5.	Layer convergence for pentacenequinone . . . . .	28
2.6.	K-point convergence for PQ on Ag(111) with mixed Ag species. . . . .	28
2.7.	High-symmetry points on the fcc(111) surface. . . . .	29
2.8.	Molecules and surfaces already used for structure searches . . . . .	30
2.9.	Representation of the feature vector for species O and H. . . . .	30
2.10.	Interaction energy evolution of a BQ dimer used to probe the minimal distance between hydrogen atoms. . . . .	31
2.11.	Performance of the different kernels for a BQ gas phase monolayer . . . . .	32
2.12.	Comparison of the different feature vectors for BQ on Ag(111) . . . . .	34
3.1.	Molecules used for structure search on Ag(111). . . . .	35
3.2.	High symmetry points on the fcc(111) surface. . . . .	36
3.3.	Rotational adsorption curve for BQ at the bridge position . . . . .	36
3.4.	Local adsorption geometries of BQ with their adsorption energies . . . . .	37
3.5.	Adsorption heights and energies of the local geometries of BQ. . . . .	38
3.6.	Local adsorption geometries of AQ with their adsorption energies . . . . .	38
3.7.	Adsorption heights and energies of the local geometries of AQ. . . . .	39
3.8.	Adsorption heights of pentacenequinone . . . . .	39
3.9.	Local adsorption geometries of PQ with their adsorption energies . . . . .	40
3.10.	Comparison of vdW and gas phase prior for benzoquinone and anthraquinone . . . . .	42
3.11.	Comparison of the interaction energy change from the van der Waals and the gas phase prior for BQ. . . . .	43
3.12.	Comparison of the interaction energy change from the van der Waals and the gas phase prior for AQ. . . . .	44
3.13.	Prediction versus DFT calculations for BQ and AQ on Ag(111) . . . . .	44

3.14. Prediction of configuration energies for benzoquinone on Ag(111) dependent on coverage and ranked by energy . . . . .	45
3.15. Prediction of configuration energies for anthraquinone on Ag(111) dependent on coverage and ranked by energy . . . . .	46
3.16. Comparison of the energy contributions for the 50 best structures of BQ and AQ. . . . .	47
3.17. Best configurations for BQ on Ag(111) . . . . .	47
3.18. Pairwise interaction plots for all pairs of the two best configuration for BQ. . . . .	48
3.19. The two best configurations for AQ on Ag(111) and the fifth best structure, which contains an adsorption geometry that causes a kink to form . . . . .	49
3.20. Pairwise interaction plots for the relevant pairs of the best configurations for AQ. . . . .	49
A1. Convergence behavior of the total energy of an Ag-bulk with respect to the number of k-points in each direction. . . . .	51
A2. Evolution of total energy of a fcc Ag-bulk with variation of the lattice constant. . . . .	52
A3. Evolution of the surface energy and work function of an Ag(111) slab in dependence of the number of layers and k-points. . . . .	53
A4. Geometry used for parameter convergence tests. . . . .	53
A5. Convergence calculations for several parameters of the tight Ag and tight C species . . . . .	54
A6. Previously used configuration and second adsorption geometry on a 6x6 Ag(111) slab . . . . .	55
A7. Dependence of the adsorption energies and difference of adsorption energies on the number of k-points and number of Ag layers . . . . .	55
A8. Work function evolution and mean SCF-cycle time of an Ag pillar of 8 / 10 atoms when gradually replacing the basis set. . . . .	57
A9. Difference of electron densities to the reference density for substitution with <i>reallylight</i> Ag species . . . . .	57
A10. Convergence behavior of CO adsorption energy and time per SCF cycle with different k point densities. . . . .	58
A11. Dependence of of the adsorption energy on the total number of layers for the mixed basis and the standard tight-only Ag. . . . .	58
A12. Interaction energy evolution of five distinct dimers with increasing molecule distance. . . . .	63
A13. Adsorption energy of BQ rotated around 4 high symmetry points . . . . .	65
A14. Adsorption energy of AQ rotated around 4 high symmetry points . . . . .	66
A15. Adsorption energy of PQ rotated around 4 high symmetry points . . . . .	66
A16. Hyper parameter tests for a gas phase monolayer of benzoquinone. . . . .	67
A17. Hyper parameter tests for a gas phase monolayer of anthraquinone. . . . .	68

## Bibliography

- [1] Frank Jensen. *Introduction to computational chemistry*. John Wiley & Sons, 2017.
- [2] E. Schrödinger. “An Undulatory Theory of the Mechanics of Atoms and Molecules”. In: *Phys. Rev.* 28 (6 1926), pp. 1049–1070.
- [3] John C Slater. “Atomic shielding constants”. In: *Physical Review* 36.1 (1930), p. 57.
- [4] S Francis Boys. “Electronic wave functions-I. A general method of calculation for the stationary states of any molecular system”. In: *Proc. R. Soc. Lond. A*. Vol. 200. 1063. The Royal Society. 1950, pp. 542–554.
- [5] Volker Blum et al. “Ab initio molecular simulations with numeric atom-centered orbitals”. In: *Computer Physics Communications* 180.11 (2009), pp. 2175–2196.
- [6] Pierre Hohenberg and Walter Kohn. “Inhomogeneous electron gas”. In: *Physical review* 136.3B (1964), B864.
- [7] Walter Kohn and Lu Jeu Sham. “Self-consistent equations including exchange and correlation effects”. In: *Physical review* 140.4A (1965), A1133.
- [8] John P Perdew and Karla Schmidt. “Jacob’s ladder of density functional approximations for the exchange-correlation energy”. In: *AIP Conference Proceedings*. Vol. 577. 1. AIP. 2001, pp. 1–20.
- [9] Chengteh Lee, Weitao Yang, and Robert G Parr. “Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density”. In: *Physical review B* 37.2 (1988), p. 785.
- [10] Axel D Becke. “Density-functional exchange-energy approximation with correct asymptotic behavior”. In: *Physical review A* 38.6 (1988), p. 3098.
- [11] John P Perdew, Kieron Burke, and Matthias Ernzerhof. “Generalized gradient approximation made simple”. In: *Physical review letters* 77.18 (1996), p. 3865.
- [12] Jianwei Sun, Adrienn Ruzsinszky, and John P Perdew. “Strongly constrained and appropriately normed semilocal density functional”. In: *Physical review letters* 115.3 (2015), p. 036402.
- [13] Jianmin Tao et al. “Climbing the density functional ladder: Nonempirical meta-generalized gradient approximation designed for molecules and solids”. In: *Physical Review Letters* 91.14 (2003), p. 146401.
- [14] Axel D Becke. “Density-functional thermochemistry. III. The role of exact exchange”. In: *The Journal of chemical physics* 98.7 (1993), pp. 5648–5652.
- [15] John P Perdew, Matthias Ernzerhof, and Kieron Burke. “Rationale for mixing exact exchange with density functional approximations”. In: *The Journal of chemical physics* 105.22 (1996), pp. 9982–9985.
- [16] Hisayoshi Iikura et al. “A long-range correction scheme for generalized-gradient-approximation exchange functionals”. In: *The Journal of Chemical Physics* 115.8 (2001), pp. 3540–3544.
- [17] Ester Livshits and Roi Baer. “A well-tempered density functional theory of electrons in molecules”. In: *Physical Chemistry Chemical Physics* 9.23 (2007), pp. 2932–2941.
- [18] Jochen Heyd, Gustavo E Scuseria, and Matthias Ernzerhof. “Hybrid functionals based on a screened Coulomb potential”. In: *The Journal of chemical physics* 118.18 (2003), pp. 8207–8215.
- [19] Stefan Grimme. “Accurate description of van der Waals complexes by density functional theory including empirical corrections”. In: *Journal of computational chemistry* 25.12 (2004), pp. 1463–1473.

- [20] Alexandre Tkatchenko and Matthias Scheffler. “Accurate molecular van der Waals interactions from ground-state electron density and free-atom reference data”. In: *Physical review letters* 102.7 (2009), p. 073005.
- [21] Victor G Ruiz et al. “Density-functional theory with screened van der Waals interactions for the modeling of hybrid inorganic-organic systems”. In: *Physical review letters* 108.14 (2012), p. 146103.
- [22] Alexandre Tkatchenko et al. “Accurate and efficient method for many-body van der Waals interactions”. In: *Physical review letters* 108.23 (2012), p. 236402.
- [23] Aron J Cohen, Paula Mori-Sánchez, and Weitao Yang. “Challenges for density functional theory”. In: *Chemical Reviews* 112.1 (2011), pp. 289–320.
- [24] Aron J Cohen, Paula Mori-Sánchez, and Weitao Yang. “Insights into current limitations of density functional theory”. In: *Science* 321.5890 (2008), pp. 792–794.
- [25] Aron J Cohen, Paula Mori-Sánchez, and Weitao Yang. *Fractional spins and static correlation error in density functional theory*. 2008.
- [26] Felix Bloch. “Über die quantenmechanik der elektronen in kristallgittern”. In: *Zeitschrift für physik* 52.7-8 (1929), pp. 555–600.
- [27] Hendrik J Monkhorst and James D Pack. “Special points for Brillouin-zone integrations”. In: *Physical review B* 13.12 (1976), p. 5188.
- [28] M Otani and O Sugino. “First-principles calculations of charged surfaces and interfaces: A plane-wave nonrepeated slab approach”. In: *Physical Review B* 73.11 (2006), p. 115407.
- [29] Hisao Ishii et al. “Energy level alignment and interfacial electronic structures at organic/metal and organic/organic interfaces”. In: *Advanced materials* 11.8 (1999), pp. 605–625.
- [30] Slawomir Braun, William R Salaneck, and Mats Fahlman. “Energy-level alignment at organic/-metal and organic/organic interfaces”. In: *Advanced Materials* 21.14-15 (2009), pp. 1450–1472.
- [31] Walter C McCrone. “Polymorphism”. In: *Physics and chemistry of the organic solid state 2* (1965), pp. 725–767.
- [32] Terence L Threlfall. “Analysis of organic polymorphs. A review”. In: *Analyst* 120.10 (1995), pp. 2435–2460.
- [33] Takeshi Matsukawa et al. “Polymorphs of rubrene crystal grown from solution”. In: *Japanese Journal of Applied Physics* 49.8R (2010), p. 085502.
- [34] Etienne Menard et al. “High-Performance n-and p-Type Single-Crystal Organic Transistors with Free-Space Gate Dielectrics”. In: *Advanced materials* 16.23-24 (2004), pp. 2097–2101.
- [35] Hyunjoong Chung and Ying Diao. “Polymorphism as an emerging design strategy for high performance organic electronics”. In: *Journal of Materials Chemistry C* 4.18 (2016), pp. 3915–3933.
- [36] Andrew OF Jones et al. “Substrate-Induced and Thin-Film Phases: Polymorphism of Organic Materials on Surfaces”. In: *Advanced functional materials* 26.14 (2016), pp. 2233–2255.
- [37] Yuhan Zhang et al. “Probing carrier transport and structure-property relationship of highly ordered organic semiconductors at the two-dimensional limit”. In: *Physical review letters* 116.1 (2016), p. 016602.
- [38] J Pannetier et al. “Prediction of crystal structures from crystal chemistry rules by simulated annealing”. In: *Nature* 346.6282 (1990), p. 343.
- [39] Artem R Oganov et al. “First blind test of inorganic crystal structure prediction methods.” In: Wiley Online Library, 2011.
- [40] David J Wales and Jonathan PK Doye. “Global optimization by basin-hopping and the lowest energy structures of Lennard-Jones clusters containing up to 110 atoms”. In: *The Journal of Physical Chemistry A* 101.28 (1997), pp. 5111–5116.

- [41] Chiara Panosetti et al. “Global Materials Structure Search with Chemically Motivated Coordinates”. In: *Nano letters* 15.12 (2015), pp. 8044–8048.
- [42] Stefan Goedecker. “Minima hopping: An efficient search method for the global minimum of the potential energy surface of complex molecular systems”. In: *The Journal of chemical physics* 120.21 (2004), pp. 9911–9917.
- [43] Stefan Goedecker. “Global optimization with the minima hopping method”. In: *Modern Methods of Crystal Structure Prediction* (2011), pp. 131–145.
- [44] Alvin L-S Chua et al. “A genetic algorithm for predicting the structures of interfaces in multi-component systems”. In: *Nature materials* 9.5 (2010), p. 418.
- [45] Peter Siepmann et al. “A genetic algorithm approach to probing the evolution of self-organized nanostructured systems”. In: *Nano letters* 7.7 (2007), pp. 1985–1990.
- [46] Yanchao Wang et al. “Crystal structure prediction via particle-swarm optimization”. In: *Physical Review B* 82.9 (2010), p. 094116.
- [47] Konstantin Krautgasser et al. “Global structure search for molecules on surfaces: Efficient sampling with curvilinear coordinates”. In: *The Journal of chemical physics* 145.8 (2016), p. 084117.
- [48] Milica Todorović et al. “Efficient Bayesian Inference of Atomistic Structure in Complex Functional Materials”. In: *arXiv preprint arXiv:1708.09274* (2017).
- [49] Daniel M Packwood, Patrick Han, and Taro Hitosugi. “Chemical and entropic control on the molecular self-assembly process”. In: *Nature Communications* 8 (2017), p. 14463.
- [50] Andrey Tikhonov. “Solution of incorrectly formulated problems and the regularization method”. In: *Soviet Meth. Dokl.* 4 (1963), pp. 1035–1038.
- [51] Michael Scherbela. “Structure Prediction at Organic/Inorganic Interfaces using Machine Learning”. Master thesis. Institute of Solid State Physics, Graz University of Technology, 2017.
- [52] Veronika Obersteiner et al. “Structure prediction for surface-induced phases of organic monolayers: Overcoming the combinatorial bottleneck”. In: *Nano letters* 17.7 (2017), pp. 4453–4460.
- [53] Michael Scherbela et al. “Leaving the Valley: Charting the Energy Landscape of Metal/Organic Interfaces via Machine Learning”. In: *arXiv preprint arXiv:1709.05417* (2017).
- [54] Lukas Hörmann. “Overcoming the Configurational Explosion for Surface Structure Search by Coarse Grained Modeling and Bayesian Learning”. Master thesis. Institute of Solid State Physics, Graz University of Technology, 2018.
- [55] G Heimel et al. “Charged and metallic molecular monolayers through surface-induced aromatic stabilization”. In: *Nature chemistry* 5.3 (2013), pp. 187–194.
- [56] Francis Birch. “Finite Elastic Strain of Cubic Crystals”. In: *Phys. Rev.* 71 (11 1947), pp. 809–824.
- [57] Wheeler P Davey. “Precision measurements of the lattice constants of twelve common metals”. In: *Physical Review* 25.6 (1925), p. 753.